ARTIN APPROXIMATION

by

Herwig HAUSER and Guillaume ROND

Abstract. — In 1968, M. Artin proved that any formal power series solution of a system of analytic equations may be approximated by convergent power series solutions. Motivated by this result and a similar result of Płoski, he conjectured that this remains true when we replace the ring of convergent power series by a more general ring.

This paper presents the state of the art on this problem, aimed at non-experts. In particular we put a slant on the Artin Approximation Problem with constraints.

Contents

1. Introduction..................................................... 1
2. Artin Approximation............................................ 15
3. Strong Artin Approximation..................................... 30
4. Examples of Applications........................................ 44
5. Approximation with constraints.................................. 46
Appendix A. Weierstrass Preparation Theorem.................... 62
Appendix B. Regular morphisms and excellent rings............... 63
Appendix C. Étale morphisms and Henselian rings................. 65
References.......................................................... 68

1. Introduction

The aim of this paper is to present the Artin Approximation Theorem and some related results. The problem we are interested in is to find analytic solutions of some system of equations when this system admits formal power series solutions and the Artin Approximation Theorem yields a positive answer to this problem. We begin this paper by giving several examples explaining what this sentence means exactly. Then we will present the state of the art on this problem. There is essentially three parts: the first part is dedicated to present the Artin Approximation Theorem and
its generalizations; the second part presents a stronger version of Artin Approximation Theorem; the last part is mainly devoted to explore the Artin Approximation Problem in the case of constraints. An appendix presents the algebraic material used in this paper (Weierstrass Preparation Theorem, excellent rings, étale morphisms and Henselian rings).

We do not give the proofs of all the results presented in this paper but, at least, we always try to outline the proofs and give the main arguments.

**Example 1.** — Let us consider the following curve \( C := \{ (t^3, t^4, t^5), t \in \mathbb{C} \} \) in \( \mathbb{C}^3 \). This curve is an algebraic set which means that it is the zero locus of polynomials in three variables. Indeed, we can check that \( C \) is the zero locus of the polynomials \( f := y^2 - xz, g := yz - x^3 \) and \( h := z^2 - x^2y \). If we consider the zero locus of any two of these polynomials we always get a set larger than \( C \). The complex dimension of the zero locus of one non-constant polynomial in three variables is 2 (such a set is called a hypersurface of \( \mathbb{C}^3 \)). Here \( C \) is the intersection of the zero locus of three hypersurfaces and not of two of them, but its complex dimension is 1.

In fact we can see this phenomenon as follows: we call an algebraic relation between \( n \) hypersurfaces and not of two of them, but its complex dimension is 1.

In this case we have two relations which are \( r_4 := (x^2, -z, y) \) and \( r_5 := (x^2, -z, y) \) and \( r_4 \) and \( r_5 \) cannot be written as \( a_1 r_1 + a_2 r_2 + a_3 r_3 \) with \( a_1, a_2 \) and \( a_3 \in \mathbb{C}[x, y, z] \), which means that \( r_4 \) and \( r_5 \) are not in the sub-\( \mathbb{C}[x, y, z] \)-module of \( \mathbb{C}[x, y, z] \) generated by \( r_1, r_2 \) and \( r_3 \).

On the other hand we can prove that \( \text{Ker}(\varphi) \) is generated by \( r_1, r_2, r_3, r_4 \) and \( r_5 \).

Let \( X \) be the common zero locus of \( f \) and \( g \). If \( (x, y, z) \in X \) and \( x \neq 0 \), then \( h = \frac{yf + zg}{x} = 0 \) thus \((x, y, z) \in C \). If \( (x, y, z) \in X \) and \( x = 0 \), then \( y = 0 \). Geometrically this means that \( X \) is the union of \( C \) and the \( z \)-axis, i.e. the union of two curves.

Now let us denote by \( \mathbb{C}[x, y, z] \) the ring of formal power series with coefficients in \( \mathbb{C} \). We can also consider formal relations between \( f, g \) and \( h \), that is elements of the kernel of the map \( \mathbb{C}[x, y, z] \rightarrow \mathbb{C}[x, y, z] \) induced by \( \varphi \). Any element of the form \( a_1 r_1 + a_2 r_2 + a_3 r_3 + a_4 r_4 + a_5 r_5 \) is a formal relation as soon as \( a_1, \ldots, a_5 \in \mathbb{C}[x, y, z] \).

In fact any formal relation is of this form, i.e. the algebraic relations generate the formal and analytic relations. We can show this as follows: we can assign the weights 3 to \( x \), 4 to \( y \) and 5 to \( z \). In this case \( f, g, h \) are homogeneous polynomials of weights 8, 9 and 10 and \( r_1, r_2, r_3, r_4 \) and \( r_5 \) are homogeneous relations of weights \( (9, 8, 0), (10, 0, 8), (0, 10, 9), (5, 4, 3), (6, 4, 5) \). If \( (a, b, c) \in \mathbb{C}[x, y, z] \) is a formal relation then we can write \( a = \sum_{i=0}^{\infty} a_i \), \( b = \sum_{i=0}^{\infty} b_i \) and \( c = \sum_{i=0}^{\infty} c_i \) where \( a_i, b_i \) and \( c_i \) are homogeneous polynomials of degree \( i \) with respect to the previous weights. Then saying that \( af + bg + ch = 0 \) is equivalent to

\[
a_i f + b_{i-1} g + c_{i-2} h = 0 \quad \forall i \in \mathbb{N}
\]
with the assumption \( b_i = c_i = 0 \) for \( i < 0 \). Thus \( (a_0, 0, 0) \), \( (a_1, b_0, 0) \) and any \( (a_i, b_{i-1}, c_{i-2}) \), for \( 2 \leq i \), are in \( \text{Ker}(\varphi) \), thus are homogeneous combinations of \( r_1, \ldots, r_5 \). Hence \( (a, b, c) \) is a combination of \( r_1, \ldots, r_5 \) with coefficients in \( \mathbb{C}[x, y, z] \).

Now we can investigate the same problem by replacing the ring of formal power series by \( \mathbb{C}\{x, y, z\} \), the ring of convergent power series with coefficients in \( \mathbb{C} \), i.e.

\[
\mathbb{C}\{x, y, z\} := \left\{ \sum_{i,j,k \in \mathbb{N}} a_{i,j,k} x^i y^j z^k / \exists \rho > 0, \sum_{i,j,k} |a_{i,j,k}| \rho^{i+j+k} < \infty \right\}.
\]

We can also consider analytic relations between \( f \), \( g \) and \( h \), i.e. elements of the kernel of the map \( \mathbb{C}\{x, y, z\}^3 \to \mathbb{C}\{x, y, z\} \) induced by \( \varphi \). From the formal case we see that any analytic relation \( r \) is of the form \( a_1 r_1 + a_2 r_2 + a_3 r_3 + a_4 r_4 + a_5 r_5 \) with \( a_i \in \mathbb{C}\{x, y, z\} \) for \( 1 \leq i \leq 5 \). In fact we can prove that \( a_i \in \mathbb{C}\{x, y, z\} \) for \( 1 \leq i \leq 5 \). Let us remark that, saying that \( r = a_1 r_1 + \cdots + a_5 r_5 \) is equivalent to say that \( a_1, \ldots, a_5 \) satisfy a system of three affine equations with analytic coefficients. This is the first example of the problem we are interested in: if we some equations with analytic coefficients have formal solutions do they have analytic solutions? Artin Approximation Theorem yields an answer to this problem. Here is the first theorem proven by M. Artin in 1968:

**Theorem 1.1 (Artin Approximation Theorem).** — [Ar68] Let \( f(x, y) \) be a vector of convergent power series over \( \mathbb{C} \) in two sets of variables \( x \) and \( y \). Assume given a formal power series solution \( \tilde{g}(x) \),

\[
f(x, \tilde{g}(x)) = 0.
\]

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

\[
f(x, g(x)) = 0
\]

which coincides with \( \tilde{g}(x) \) up to degree \( c \),

\[
y(x) \equiv \tilde{g}(x) \modulo (x)^c.
\]

We can define a topology on \( \mathbb{C}[x] \) by saying that two power series are close if their difference is in a high power of the maximal ideal \( (x) \). Thus we can reformulate Theorem 1.1 as: formal power series solutions of a system of analytic equations may be approximated by convergent power series solutions.

**Example 2.** — A special case of Theorem 1.1 and a generalization of Example 1 occurs when \( f \) is linear in \( y \), say \( f(x, y) = \sum f_i(x) y_i \), where \( f_i(x) \) is a vector of convergent power series with \( r \) coordinates for any \( i \). A solution \( y(x) \) of \( f(x, y) = 0 \) is a relation between the \( f_i(x) \). In this case the formal relations are linear combinations of analytic combinations with coefficients in \( \mathbb{C}[x] \). In term of commutative algebra, this is expressed as the flatness of the ring of formal power series over the ring of convergent powers series, a result which can be proven via the Artin-Rees Lemma. It means that if \( \tilde{g}(x) \) is a formal solution of \( f(x, y) = 0 \), then there exist analytic solutions of \( f(x, y) = 0 \) denoted by \( \hat{g}_i(x) \), \( 1 \leq i \leq s \), and formal power series \( \hat{b}_i(x) \),..
\(\hat{b}_i(x)\), such that \(\hat{g}(x) = \sum \hat{b}_i(x)\hat{y}_i(x)\). Thus, by replacing in the previous sum the \(\hat{b}_i(x)\) by their truncation at order \(c\), we obtain an analytic solution of \(f(x,y) = 0\) coinciding with \(\hat{g}(c)\) up to degree \(c\).

If the \(f_i(x)\) are vectors of polynomials then the formal relations are also linear combinations of algebraic relations since the ring of formal power series is flat over the ring of polynomials, and Theorem 1.1 remains true if \(f(x,y)\) is linear in \(y\) and \(\mathbb{C}\{x\}\) is replaced by \(\mathbb{C}[x]\).

**Example 3.** — A slight generalization of the previous example is when \(f(x,y)\) is a vector of polynomials in \(y\) of degree one with coefficients in \(\mathbb{C}\{x\}\) (resp. \(\mathbb{C}[x]\)), say

\[
f(x,y) = \sum_{i=1}^{m} f_i(x)y_i + b(x)
\]

where the \(f_i(x)\) and \(b(x)\) are vectors of convergent power series (resp. polynomials).

Here \(x\) and \(y\) are multi-variables If \(\hat{g}(x)\) is a formal power series solution of \(f(x,y) = 0\), then \((\hat{g}(x),1)\) is a formal power series solution of \(g(x,y,z) = 0\) where

\[
g(x,y,z) := \sum_{i=1}^{m} f_i(x)y_i + b(x)z
\]

and \(z\) is a single variable. Thus using the flatness of \(\mathbb{C}[x]\) over \(\mathbb{C}\{x\}\) (resp. \(\mathbb{C}[x]\)) (Example 2), we can approximate \((\hat{g}(x),1)\) by a convergent power series (resp. polynomial) solution \((\tilde{g}(x),\tilde{z}(x))\) which coincides with \((\hat{g}(x),1)\) up to degree \(c\). In order to obtain a solution of \(f(x,y) = 0\) we would like to be able to divide \(\hat{g}(x)\) by \(\tilde{z}(x)\) since \(\hat{g}(x)\tilde{z}(x)^{-1}\) would be a solution of \(f(x,y) = 0\) approximating \(\tilde{g}(x)\). We can remark that, if \(c \geq 1\), then \(\tilde{z}(0) = 1\) thus \(\tilde{z}(x)\) is not in the ideal \((x)\). But \(\mathbb{C}\{x\}\) is a local ring. We call a local ring any ring \(A\) that has only one maximal ideal. This is equivalent to say that \(A\) is the disjoint union of one ideal (its only maximal ideal) and of the set of units in \(A\). In particular \(\tilde{z}(x)^{-1}\) is invertible in \(\mathbb{C}\{x\}\), hence we can approximate formal power series solutions of \(f(x,y) = 0\) by convergent power series solutions.

In the case \((\hat{g}(x),\tilde{z}(x))\) is a polynomial solution of \(g(x,y,z) = 0\), \(\tilde{z}(x)\) is not invertible in general in \(\mathbb{C}[x]\) since it is not a local ring. For instance set

\[
f(x,y) := (1-x)y - 1
\]

where \(x\) and \(y\) are single variables. Then \(y(x) := \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}\) is the only formal power series solution of \(f(x,y) = 0\), but \(y(x)\) is not a polynomial. Thus we cannot approximate the roots of \(f\) in \(\mathbb{C}[x]\) by roots of \(f\) in \(\mathbb{C}[x]\).

But instead of working in \(\mathbb{C}[x]\) we can work in \(\mathbb{C}[x](x)\) which is the ring of rational functions whose denominator does not vanish at 0. This ring is a local ring. Since \(\tilde{z}(0) \neq 0\), then \(\hat{y}(x)\tilde{z}(x)^{-1}\) is a vector of rational function of \(\mathbb{C}[x](x)\). In particular any system of polynomial equations of degree one with coefficients in \(\mathbb{C}[x]\) which has solutions in \(\mathbb{C}[x]\) has solutions in \(\mathbb{C}[x](x)\).
In term of commutative algebra, the fact that degree 1 polynomial equations satisfy Theorem 3.3 is expressed as the faithful flatness of the ring of formal power series over the ring of convergent powers series, a result that follows from the flatness and the fact that the ring of convergent power series is a local ring.

Example 4. — The next example we are looking at is the following: set $f \in A$ where $A = \mathbb{C}[x]$ or $\mathbb{C}[x,y]$ or $\mathbb{C}[x]$. When do there exist $g, h \in A$ such that $f = gh$? First of all, we can take $g = 1$ and $h = f$ or, more generally, $g$ a unit in $A$ and $h = g^{-1}f$. These are trivial cases and thus we are looking for non units $g$ and $h$.

Of course, if there exist non units $g$ and $h$ in $A$ such that $f = gh$, then $f = (\hat{u}g)(\hat{u}^{-1}h)$ for any unit $\hat{u} \in \mathbb{C}[x]$. But is the following true: let us assume that there exist $\hat{g}, \hat{h} \in \mathbb{C}[x]$ such that $f = \hat{g}\hat{h}$. Then do there exist non units $\hat{g}, \hat{h} \in A$ such that $f = gh$?

Let us remark that this question is equivalent to the following: if $\frac{\mathbb{C}[x]}{(f)}$ is an integral domain, is $\frac{\mathbb{C}[x]}{(f)}$ still an integral domain?

The answer to this question is no in general: set $A := \mathbb{C}[x,y]$ and set $f := x^2 - y^2(1+y)$. Then $f$ is irreducible as a polynomial since $y^2(1+y)$ is not a square in $\mathbb{C}[x,y]$. But $f = (x+y\sqrt{1+y})(x-y\sqrt{1+y})$ where $\sqrt{1+y}$ is a formal power series such that $\sqrt{1+y}^2 = 1 + y$. Thus $f$ is not irreducible in $\mathbb{C}[x,y]$ nor in $\mathbb{C}[x,y][x,y]$ but it is irreducible in $\mathbb{C}[x,y]$ or $\mathbb{C}[x,y][x,y]$.

In fact it is easy to see that $x + y\sqrt{1+y}$ and $x - y\sqrt{1+y}$ are power series which are algebraic over $\mathbb{C}[x,y]$, i.e. they are roots of polynomials with coefficients in $\mathbb{C}[x,y]$. The set of such algebraic power series is a subring of $\mathbb{C}[x,y]$ and it is denoted by $\mathbb{C}(x,y)$. In general if $x$ is a multivariable the ring of algebraic power series $\mathbb{C}(x)$ is the following:

$$\mathbb{C}(x) := \{ f \in \mathbb{C}[x] / \exists P(z) \in \mathbb{C}[x][z], \ P(f) = 0 \} .$$

It is not difficult to prove that the ring of algebraic power series is a subring of the ring of convergent power series and is a local ring. In 1969, M. Artin proved an analogue of Theorem 2.1 for the rings of algebraic power series [At69]. Thus if $f \in \mathbb{C}(x)$ (or $\mathbb{C}(x)$) is irreducible then it remains irreducible in $\mathbb{C}[x]$, this is a consequence of Artin Approximation Theorem. From this theorem we can also deduce that if $f \in \mathbb{C}[x]$ (or $\mathbb{C}(x)$), for some ideal $I$, is irreducible, then it remains irreducible in $\mathbb{C}[x]/I$.

Example 5. — Let us strengthen the previous question. Let us assume that there exist $\hat{g}, \hat{h} \in \mathbb{C}[x]$ such that $f = \hat{g}\hat{h}$ with $f \in A$ with $A = \mathbb{C}[x]$ or $\mathbb{C}[x]$. Then does there exist a unit $\hat{u} \in \mathbb{C}[x]$ such that $\hat{u}g \in A$ and $\hat{u}^{-1}h \in A$?

The answer to this question is positive if $A = \mathbb{C}[x]$ or $\mathbb{C}[x]$, this is a non trivial corollary of Artin Approximation Theorem (see Corollary 4.4). But it is negative in general for $\mathbb{C}(x)$ or $\mathbb{C}(x)$ if $I$ is an ideal. The following example is due to S. Izumi [Iz92].
As before, the answer to the first question is positive for 
\[ u = f \]
Nevertheless, the answer to the second question is positive in the cases
\[ g \in \mathbb{C}[x,y,z] \]

Example 6. — A similar question is the following: if \( f \in A \) with \( A = \mathbb{C}[x], \mathbb{C}[x,y] \), \( \mathbb{C}[x] \) or \( \mathbb{C}[x,y,z] \) and if there exist a non unit \( \widehat{g} \in \mathbb{C}[x] \) and an integer \( m \in \mathbb{N} \) such that \( \widehat{g}^m = f \), does there exist a non unit \( g \in A \) such that \( g^m = f \)?

A weaker question is the following: if \( \frac{A}{(f)} \) is reduced, is \( \frac{\mathbb{C}[x]}{(\widehat{f})} \) still reduced? Indeed, if \( \widehat{g}^m = f \) for some non unit \( \widehat{g} \) then \( \frac{\mathbb{C}[x]}{(\widehat{f})} \) is not reduced. Thus, if the answer to the second question is positive, then there exists a non unit \( g \in A \) and a unit \( u \in A \) such that \( u g^k = f \) for some integer \( k \).

As before, the answer to the first question is positive for \( A = \mathbb{C}[x,y] \) and \( A = \mathbb{C}[x] \) by Artin Approximation Theorem.

If \( A = \mathbb{C}[x] \) or \( \mathbb{C}[x,y,z] \), the answer to this question is negative. Indeed let us consider \( f = x^m + x^{m+1} \). Then \( f = \widehat{g}^m \) with \( \widehat{g} := x \sqrt[3]{1 + x} \) but there is no \( g \in A \) such that \( g^m = f \).

Nevertheless, the answer to the second question is positive in the cases \( A = \mathbb{C}[x] \) or \( \mathbb{C}[x,y,z] \). This deep result is due to D. Rees (see [HIS06] for instance).

Example 7. — Using the same notation as in Example 4 we can ask a stronger question: set \( A = \mathbb{C}[x] \) or \( \mathbb{C}[x,y] \) and let \( f \) be in \( A \). If there exist \( \overline{g} \) and \( \overline{h} \in \mathbb{C}[x] \), vanishing at 0, such that \( f = \overline{g} \overline{h} \) modulo a large power of the ideal \( (x) \), do there exist \( g \) and \( h \) in \( A \) such that \( f = gh \)? By example 4 there is no hope, if \( g \) and \( h \) exist, to expect that \( g \) and \( h \) belong to \( \mathbb{C}[x] \).

We have the following theorem:
Theorem 1.2 (Strong Artin Approximation Theorem) [Ar69] Let \( f(x, y) \) be a vector of polynomials over \( \mathbb{C} \) in two sets of variables \( x \) and \( y \). Then there exists a function \( \beta : \mathbb{N} \rightarrow \mathbb{N} \), such that for any integer \( c \) and any given approximate solution \( y(x) \) at order \( \beta(c) \),

\[
 f(x, y(x)) \equiv 0 \text{ modulo } (x)^{\beta(c)},
\]

there exists an algebraic power series solution \( y(x) \),

\[
 f(x, y(x)) = 0
\]

which coincides with \( y(x) \) up to degree \( c \),

\[
 y(x) \equiv y(x) \text{ modulo } (x)^c.
\]

In particular, if \( y(x) - f(x) \equiv 0 \) modulo \( (x)^{\beta(1)} \), where \( \beta \) is the function of the previous theorem for the polynomial \( y_1y_2 - f \), and if \( y(0) = f(0) = 0 \), then there exist non units \( g \) and \( h \) in \( \mathbb{C}(x) \) such that \( gh - f = 0 \).

A natural question is: given \( f \in \mathbb{C}[x] \) how to compute \( \beta \) or, at least, \( \beta(1) \)? That is, up to what order do we have to check that the equation \( y_1y_2 - f = 0 \) has an approximate solution in order to be sure that this equation has solutions? For instance, if \( f := x_1x_2 - x_3^2 \) then \( f \) is irreducible but \( x_1x_2 - f \equiv 0 \) modulo \( (x)^d \) for any \( d \in \mathbb{N} \), so obviously \( \beta(1) \) really depends on \( f \).

In fact, in Theorem 1.2 M. Artin proved that \( \beta \) can be chosen independently of the degree of the components of the the vector \( f(x, y) \). But it is still an open problem to find effective bounds on \( \beta \) (see Section 3.4).

Example 8 (Ideal Membership Problem). — Set \( f_1, \ldots, f_r \in \mathbb{C}[x] \) where \( x = (x_1, \ldots, x_n) \). Let us denote by \( I \) the ideal of \( \mathbb{C}[x] \) generated by \( f_1, \ldots, f_r \). If \( g \) is a power series, how can we detect that \( g \in I \) or \( g \notin I \)? Since a power series is determined by its coefficients, saying that \( g \in I \) will depend in general on an infinite number of conditions and it will not be possible to check that all these conditions are satisfied in finite time. Another problem is to find canonical representatives of power series modulo the ideal \( I \) that will help us to make computations in the quotient ring \( \mathbb{C}[x]/I \).

One way to solve these problems is the following. Let us consider the following order on \( \mathbb{N}^n \): for any \( \alpha, \beta \in \mathbb{N}^n \), we say that \( \alpha \leq \beta \) if \( (|\alpha|, \alpha_1, \ldots, \alpha_n) \leq_{\text{lex}} (|\beta|, \beta_1, \ldots, \beta_n) \) where \( |\alpha| := \alpha_1 + \cdots + \alpha_n \) and \( \leq_{\text{lex}} \) is the lexicographic order. For instance

\[
 (1, 1, 1) \leq (1, 2, 3) \leq (2, 2, 2) \leq (3, 2, 1) \leq (2, 2, 3).
\]

This order induces an order on the sets of monomials \( x_1^{\alpha_1} \ldots x_n^{\alpha_n} \): we say that \( x^\alpha \leq x^\beta \) if \( \alpha \leq \beta \). Thus

\[
 x_1x_2x_3 \leq x_1x_2^2x_3 \leq x_1^2x_2^2x_3 \leq x_1^3x_2^2x_3 \leq x_1^2x_2x_3^2.
\]
If \( f := \sum_{\alpha \in \mathbb{N}^n} f_\alpha x^\alpha \in \mathbb{C}[x] \), the initial exponent of \( f \) with respect to the previous order is

\[
\exp(f) := \min\{\alpha \in \mathbb{N}^n / f_\alpha \neq 0\} = \inf \text{Supp}(f)
\]

where the support of \( f \) is \( \text{Supp}(f) := \{\alpha \in \mathbb{N}^n / f_\alpha \neq 0\} \). The initial term of \( f \) is \( f_{\exp(f)} x^{\exp(f)} \). This is the smallest non zero monomial in the Taylor expansion of \( f \) with respect to the previous order.

If \( I \) is an ideal of \( \mathbb{C}[x] \), we define \( \Gamma(I) \) to be the subset of \( \mathbb{N}^n \) of all the initial exponents of elements of \( I \). Since \( I \) is an ideal, for any \( \beta \in \mathbb{N}^n \) and any \( f \in I \), \( x^\beta f \in I \). This means that \( \Gamma(I) + \mathbb{N}^n = \Gamma(I) \). Then we can prove that there exists a finite number of elements \( g_1, \ldots, g_s \in I \) such that

\[
\{\exp(g_1), \ldots, \exp(g_s)\} + \mathbb{N}^n = \Gamma(I).
\]

Set

\[
\Delta_1 := \exp(g_1) + \mathbb{N}^n \quad \text{and} \quad \Delta_i = (\exp(g_i) + \mathbb{N}^n) \setminus \bigcup_{1 \leq j < i} \Delta_j, \quad \text{for} \ 2 \leq i \leq s.
\]

Finally, set

\[
\Delta_0 := \mathbb{N}^n \setminus \bigcup_{i=1}^s \Delta_i.
\]

For instance, if \( I \) is the ideal of \( \mathbb{C}[x_1, x_2] \) generated by \( g_1 := x_1x_2^3 \) and \( g_2 := x_1^2x_2^2 \), we can check that

\[
\Gamma(I) = \{(1, 3), (2, 2)\} + \mathbb{N}_2.
\]

Set \( g \in \mathbb{C}[\![x]\!] \). Then, by Galligo-Grauert-Hironaka Division Theorem [Gal79], there exist unique power series \( q_1, \ldots, q_s, r \in \mathbb{C}[x] \) such that

\[
g = g_1 q_1 + \cdots + g_s q_s + r
\]

\[
\exp(g_i) + \text{Supp}(q_i) \subset \Delta_i \quad \text{and} \quad \text{Supp}(r) \subset \Delta_0.
\]
The uniqueness of the division comes from the fact the $\Delta_i$ are disjoint subsets of $\mathbb{N}^n$. The existence of such decomposition is proven through the division algorithm:

Set $\alpha := \exp(g)$. Then there exists an integer $i_1$ such that $\alpha \in \Delta_{i_1}$.
- If $i_1 = 0$, then set $r^{(1)} := \text{in}(g)$ and $q_i^{(1)} := 0$ for all $i$.
- If $i_1 > 1$, then set $r^{(1)} := 0$, $q_i^{(1)} := 0$ for $i \neq i_1$ and $q_i^{(1)} := \frac{\text{in}(g) - \exp(g)}{\exp(g)}$.

Finally set $g^{(1)} := g - \sum_{i=1}^{s} g_i q_i^{(1)}$. Thus we have $\exp(g^{(1)}) > \exp(g)$. Then we replace $g$ by $g^{(1)}$ and the repeat the preceding process.

In this way we construct a sequence $(g^{(k)})_k$ of power series such that, for any $k \in \mathbb{N}$, $\exp(g^{(k+1)}) > \exp(g^{(k)})$ and $g^{(k)} = g - \sum_{i=1}^{s} g_i q_i^{(k)} - r^{(k)}$ with

$$\exp(g_i) + \text{Supp}(q_i^{(k)}) \subset \Delta_i$$

At the limit $k \to \infty$ we obtain the desired decomposition.

In particular since $\{\exp(g_1),...,\exp(g_s)\} + \mathbb{N}^n = \Gamma(I)$ we deduce from this that $I$ is generated by $g_1,...,g_s$.

This algorithm means that for any $g \in \mathbb{C}[x]$ there exists a unique power series $r$ whose support is included in $\Delta$ and such that $g - r \in I$ and the division algorithm yields a way to obtain this representative $r$.

Moreover, saying that $g \not\in I$ is equivalent to $r \neq 0$ and this is equivalent to say that, for some integer $k$, $r^{(k)} \neq 0$. But $g \in I$ is equivalent to $r = 0$ which is equivalent to $r^{(k)} = 0$ for all $k \in \mathbb{N}$. Thus applying the division algorithm, if for some integer $k$, $r^{(k)} \neq 0$, then we can conclude that $g \not\in I$. But this algorithm will not help us to determine if $g \in I$ since we would have to make a infinite number of computations.

Now a natural question is, what happens if we replace $\mathbb{C}[x]$ by $A := \mathbb{C}(x)$ or $\mathbb{C}\{x\}$? Of course we can proceed with the division algorithm but we do not know if $q_1,...,q_s, r \in A$. In fact by controlling the size of the coefficients of $q_i^{(k)}, q_s^{(k)}, r^{(k)}$ at each step of the division algorithm, we can prove that if $g \in \mathbb{C}\{x\}$ then $q_1,...,q_s$ and $r$ remain in $\mathbb{C}\{x\}$ ([Hir64], [Gra72], [Gal79] and [dJP00]). But if $g \in \mathbb{C}(x)$ then it may happen that $q_1,..., q_s$ and $r$ are not in $\mathbb{C}(x)$ (see Example 5.4 of Section 5).

**Example 9 (Arcs Space and Jets Spaces).** — Let $X$ be an affine algebraic subset of $\mathbb{C}^n$, i.e. $X$ is the zero locus of some polynomials in $m$ variables: $f_1,...,f_r \in \mathbb{C}[y_1,...,y_m]$. Let $t$ be a single variable. For any integer $n$, let us define $X_n$ to be the set of vectors $y(t)$ whose coordinates are polynomials of degree $\leq n$ and such that $f(y(t)) \equiv 0$ modulo $(t)^{n+1}$. The elements of $X_n$ are called $n$-jets on $X$.

If $y_i(t) = y_{i,0} + y_{i,1}t + \cdots + y_{i,n}t^n$ and if we consider each $y_{i,j}$ has one indeterminate, saying that $f(y(t)) \in (t)^{n+1}$ is equivalent to the vanishing of $r(n+1)$ polynomials...
For instance, if \(10\) equations involving the \(y_{i,j}\) . This shows that the jets spaces of \(X\) are algebraic sets. For instance, if \(X\) is a cusp, i.e. the plane curve defined by \(X := \{y_1^2 - y_2^3 = 0\}\), then
\[
X_0 := \{(a_0, b_0) \in \mathbb{C}^2 / a_0^2 - b_0^3 = 0\} = X.
\]

We have
\[
X_1 = \{(a_0, a_1, b_0, b_1) \in \mathbb{C}^4 / (a_0 + a_1 t)^2 - (b_0 + b_1 t)^3 \equiv 0 \mod t^2\}
= \{(a_0, a_1, b_0, b_1) \in \mathbb{C}^4 / a_0^2 - b_0^3 = 0 \text{ and } 2a_0 a_1 - 3b_0 b_1 = 0\}.
\]
The morphisms \(\mathbb{C}[t] \rightarrow \mathbb{C}[\frac{t}{(t)^{n+1}}]\), for \(k \geq n\), induce truncation maps \(\pi_n^k : X_k \rightarrow X_n\) by reducing \(k\)-jets modulo \((t)^{n+1}\). In the example we are considering, the fibre of \(\pi_1^1\) over the point \((a_0, b_0) \neq (0, 0)\) is the line in the \((a_1, b_1)\)-plane whose equation is
\[
2a_0 a_1 - 3b_0^2 b_1 = 0.
\]
This line is exactly the tangent space at \(X\) at the point \((a_0, b_0)\).

The tangent space at \(X\) in \((0, 0)\) is the whole plane since this point is a singular point of the plane curve \(X\). This corresponds to the fact that the fibre of \(\pi_1^1\) over \((0, 0)\) is the whole plane.

On this example we show that \(X_1\) is isomorphic to the tangent bundle of \(X\), which is a general fact.

We can easily see that \(X_2\) is given by the following equations:
\[
\begin{cases}
a_0^2 - b_0^3 = 0 \\
2a_0 a_1 - 3b_0^2 b_1 = 0 \\
a_1^2 + 2a_0 a_2 - 3b_0 b_1^2 - 3b_0^2 b_2 = 0
\end{cases}
\]
In particular, the fibre of \(\pi_2^2\) over \((0, 0)\) is the set of points of the form \((0, 0, a_2, 0, b_1, b_2)\) and the image of this fibre by \(\pi_2^1\) is the line \(a_1 = 0\). This shows that \(\pi_2^1\) is not surjective.

But, we can show that above the smooth part of \(X\), the maps \(\pi_n^{n+1}\) are surjective and the fibres are isomorphic to \(\mathbb{C}\).

The space of arcs on \(X\), denoted by \(X_\infty\), is the set of vectors \(y(t)\) whose coordinates are formal power series satisfying \(f(y(t)) = 0\). For such a general vector of formal power series \(y(t)\), saying that \(f(y(t)) = 0\) is equivalent to say that the coefficients of all the powers of \(t\) in the Taylor expansion of \(f(y(t))\) are equal to zero. This shows that \(X_\infty\) may be defined by a countable number of equations in a countable number of variables. For instance, in the previous example, \(X_\infty\) is the subset of \(\mathbb{C}^n\) with coordinates \((a_0, a_1, a_2, \ldots, b_0, b_1, b_2, \ldots)\) defined by the infinite following equations:
\[
\begin{cases}
a_0^2 - b_0^3 = 0 \\
2a_0 a_1 - 3b_0^2 b_1 = 0 \\
a_1^2 + 2a_0 a_2 - 3b_0 b_1^2 - 3b_0^2 b_2 = 0 \\
\ldots \ldots
\end{cases}
\]

The morphisms \(\mathbb{C}[t] \rightarrow \mathbb{C}[\frac{t}{(t)^{n+1}}]\) induce truncations maps \(\pi_n : X_\infty \rightarrow X_n\) by reducing arcs modulo \((t)^{n+1}\).
In general it is a difficult problem to compare $\pi_n(X_{\infty})$ and $X_n$. It is not even clear if $\pi_n(X_{\infty})$ is finitely defined. But we have the following theorem due to Greenberg which is a particular case of Theorem \ref{thm:artin} in which $\beta$ is bounded by an affine function:

**Theorem 1.3 (Greenberg Theorem).** — \cite{Gre66} Let $f(y)$ be a vector of polynomials in $m$ variables and let $t$ be a single variable. Then there exist two positive integers $a$ and $b$, such that for any polynomial solution $y(t)$ modulo $(t)^{an+b}$,

$$f(y(t)) \equiv 0 \mod (t)^{an+b+1},$$

there exists a formal power series solution $\tilde{y}(t)$,

$$f(\tilde{y}(t)) = 0$$

which coincides with $y(t)$ up to degree $n + 1$,

$$y(t) \equiv \tilde{y}(t) \mod (t)^{n+1}.$$

We can reinterpret this result as follows: let $X$ be the zero locus of $f$ and let $y(t)$ be a $(an + b)$-jet on $X$. Then the truncation of $y(t)$ modulo $(t)^{n+1}$ is the truncation of a formal power series solution of $f = 0$. Thus we have

$$\pi_n(X_{\infty}) = \pi_{an+b}(X_{an+b}), \quad \forall n \in \mathbb{N}.$$ 

A constructible subset of $\mathbb{C}^n$ is a set defined by the vanishing of some polynomials and the non-vanishing of other polynomials, i.e. a set of the form

$$\{x \in \mathbb{C}^n / f_1(x) = \cdots = f_r(x) = 0, g_1(x) \neq 0, \ldots, g_s(x) \neq 0\}$$

for some polynomials $f_i, g_j$. In particular algebraic sets are constructible sets. Since a theorem of Chevalley asserts that the projection of an algebraic subset of $\mathbb{C}^{n+k}$ onto $\mathbb{C}^k$ is a constructible subset of $\mathbb{C}^n$, Theorem \ref{thm:artin} asserts that $\pi_n(X_{\infty})$ is a constructible subset of $\mathbb{C}^n$ since $X_{an+b}$ is an algebraic set. In particular $\pi_n(X_{\infty})$ is finitely defined, i.e. it is defined by a finite number of data (see \cite{GoLJ96} for an introduction to the study of these sets).

A difficult problem in singularity theory is to understand the behaviors of $X_n$ and $\pi_n(X_{\infty})$ and to relate them to the geometry of $X$. One way to do this is to define the (motivic) measure of a constructible subset of $\mathbb{C}^n$, that is an additive map $\chi$ from the set of constructible sets to a commutative ring $R$, such that:

- $\chi(X) = \chi(Y)$ as soon as $X$ and $Y$ are isomorphic algebraic sets,
- $\chi(X \cup Y) + \chi(U) = \chi(X)$ as soon as $U$ is an open set of an algebraic set $X$,
- $\chi(X \times Y) = \chi(X).\chi(Y)$ for any algebraic sets $X$ and $Y$.

Then we are interested to understand the following formal power series:

$$\sum_{n \in \mathbb{N}} \chi(X_n)T^n \quad \text{and} \quad \sum_{n \in \mathbb{N}} \chi(\pi_n(X_{\infty}))T^n \in R[T].$$

The reader may consult \cite{DeLo99}, \cite{Lo00}, \cite{Ve06} for instance.
Let us assume that $X$ by as we have done in the previous example, for any sets $k^\geq$, $\geq k$ that $k$, $g$ is algebraically closed field and if $w$ is uncountable. As in the previous example let us define the following sets:

$$X_i := \{ y(x) \in k[x]^m / f_i(x, y(x)) \in (x)^{l+1} \ \forall i \}.$$ 

As we have done in the previous example, for any $l$ there exists an integer $N(l) \in \mathbb{N}$ such that $X_l \subset k^{N(l)}$. Moreover $X_l$ is an algebraic subset of $K^{N(l)}$ and the morphisms $\pi^k_l : X_k \to X_l$ for $k \leq l$ induce truncations maps $\pi^k_l : X_k \to X_l$ for any $k \geq l$.

By a theorem of Chevalley, for any $l \in \mathbb{N}$, the sequence $(\pi^k_l(X_k))_k$ is a decreasing sequence of constructible subsets of $X_l$. Thus the sequence $(\pi^k_l(X_k))_k$ is a decreasing sequence of algebraic subsets of $X_l$, where $\overline{Y}$ denotes the Zariski closure of a subset $Y$, i.e. the smallest algebraic set containing $Y$. By Noetherianity this sequence stabilizes: $\pi^k_l(X_k) = \pi^{k'}_l(X_k)$ for all $k$ and $k'$ large enough (say for any $k, k' \geq k_1$). Let us denote by $F_l$ this algebraic set.

Let us assume that $X_k \neq \emptyset$ for any $k \in \mathbb{N}$. This implies that $F_l \neq \emptyset$. Set $C_{k,l} := \pi^k_l(X_k)$. It is a constructible set whose Zariski closure is $F_l$ for any $k \geq k_1$. Thus $C_{k,l}$ has the form $F_l \setminus V_k$ where $V_k$ is an algebraic proper subset of $F_l$, for any $k \geq k_1$. Since $k$ is uncountable the set $U_l := \bigcap_k C_{k,l} = \bigcap_k F_l \setminus V_k$ is not empty. By construction $U_l$ is exactly the set of points of $X_l$ that can be lifted to points of $X_k$ for any $k \geq l$. In particular $\pi^k_l(U_k) = U_l$. If $x_0 \in U_0$ then $x_0$ may be lifted to $U_1$, i.e. there exists $x_1 \in U_1$ such that $\pi^0_1(x_1) = x_0$. By induction we may construct a sequence of points $x_l \in U_l$ such that $\pi^k_{l+1}(x_{l+1}) = x_l$ for any $l \in \mathbb{N}$. At the limit we obtain a point $x_\infty$ in $X_\infty$, i.e. a power series solution $y(x) \in k[[x]]$ solution of $f(x, y) = 0$.

We have proved here the following result similar to Theorem 1.2 if $k$ is an uncountable algebraically closed field and if $f(x, y) = 0$ has solutions modulo $(x)^k$ for any $k \in \mathbb{N}$, then there exists a power series solution $y(x)$:

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

This kind of argument using asymptotic contructions (here the Noetherianity is the key point of the proof) may be nicely formalized using ultraproducts. Ultraproducts methods can be used to prove easily stronger results as Theorem 1.2 (See Part 3.3 and Proposition 3.25).

Example 11 (Linearization of germs of diffeomorphisms)

Given $f \in \mathbb{C}[x]$, $x$ being a single variable, let us assume that $f'(0) = \lambda \neq 0$. Then $f$ defines an analytic diffeomorphism from a neighborhood of $0$ in $\mathbb{C}$ onto a neighborhood of $0$ in $\mathbb{C}$ preserving the origin. The linearization problem, firstly investigated by C. L. Siegel, is the following: is $f$ conjugated to its linear part? That is: does there exist $g(x) \in \mathbb{C}[x]$, with $g(0) \neq 0$, such that $f(g(x)) = g(\lambda x)$ or $g^{-1} \circ f \circ g(x) = \lambda x$ (in this case we say that $f$ is analytically linearizable)? This problem is difficult and the following cases may occur: $f$ is not linearizable, $f$ is formally linearizable but not analytically linearizable (i.e. $g$ exists but $g(x) \in \mathbb{C}[x] \setminus \mathbb{C}[x]$, $f$ is analytically linearizable (see Ce91)).
Let us assume that $f$ is formally linearizable, i.e. there exists $\hat{g}(x) \in \mathbb{C}[x]$ such that $f(\hat{g}(x)) - \hat{g}(\lambda x) = 0$. By considering the Taylor expansion of $\hat{g}(\lambda x)$:

$$\hat{g}(\lambda x) = \hat{g}(y) + \sum_{n=1}^{\infty} \frac{(y - \lambda x)^n}{n!} f^{(n)}(y)$$

we see that there exists $\hat{h}(x, y) \in \mathbb{C}[x, y]$ such that $\hat{g}(\lambda x) = \hat{g}(y) + (y - \lambda x)\hat{h}(x, y)$. Thus $f$ is formally linearizable if and only if there exists $\hat{h}(x, y) \in \mathbb{C}[x, y]$ such that

$$f(\hat{g}(x)) - \hat{g}(y) + (y - \lambda x)\hat{h}(x, y) = 0.$$ 

This former equation is equivalent to the existence of $\hat{k}(y) \in \mathbb{C}[y]$ such that

$$\begin{cases} f(\hat{g}(x)) - \hat{k}(y) + (y - \lambda x)\hat{h}(x, y) = 0 \\ \hat{k}(y) - \hat{g}(y) = 0 \end{cases}$$

Using the same trick as before (Taylor expansion), this is equivalent to the existence of $\tilde{l}(x, y, z) \in \mathbb{C}[x, y, z]$ such that

$$\begin{cases} f(\hat{g}(x)) - \hat{k}(y) + (y - \lambda x)\hat{h}(x, y) = 0 \\ \hat{k}(y) - \hat{g}(y) = 0 \\ \hat{l}(y) - \hat{g}(x) + (x - y)\tilde{l}(x, y) = 0 \end{cases}$$

Hence, we see that, if $f$ is formally linearizable, there exists a formal solution

$$\left( \hat{g}(x), \hat{k}(z), \hat{h}(x, y), \tilde{l}(x, y, z) \right)$$

of the system (1). Such a solution is called a solution with constraints. On the other hand, if the system (1) has a convergent solution $(g(x), k(z), h(x, y), l(x, y, z))$, then $f$ is analytically linearizable.

We see that the problem of linearizing analytically $f$ when $f$ is formally linearizable is equivalent to find convergent power series solutions of the system (1) with constraints. Since it happens that $f$ may be analytically linearizable but not formally linearizable, such a system (1) may have formal solutions with constraints but no analytic solutions with constraints.

In Section 5 we will give some results about the Artin Approximation Problem with constraints.

Example 12. — Another related problem is the following: if a differential equation with convergent power series coefficients has a formal power series solution, does it have convergent power series solutions? We can ask the same question by replacing "convergent" by "algebraic".

For instance let us consider the (divergent) formal power series $\hat{g}(x) := \sum_{n=0}^{\infty} n!x^{n+1}$.

It is straightforward to check that it is a solution of the equation

$$x^2 y' - y + x = 0 \text{ (Euler Equation)}.$$
On the other hand if \( \sum_n a_n x^n \) is a solution of the Euler Equation then the sequence \((a_n)_n\) satisfies the following recursion:

\[
\begin{align*}
a_0 &= 0, & a_1 &= 1 \\
a_{n+1} &= na_n & \forall n \geq 1.
\end{align*}
\]

Thus \( a_{n+1} = (n + 1)! \) for any \( n > 0 \) and \( \hat{y}(x) \) is the only solution of the Euler Equation. Hence we have an example of a differential equation with polynomials coefficients with a formal power series solution but without convergent power series solution. We will discuss in Section 5 how to relate this phenomenon to an Artin Approximation problem for polynomial equations with constraints (see Example 5.2).
If $A$ is a local ring, then $\mathfrak{m}_A$ will denote its maximal ideal. For any $f \in A$, $f \neq 0$,
\[ \text{ord}(f) := \max\{n \in \mathbb{N} \mid f \in \mathfrak{m}_A^n\}. \]
If $A$ is an integral domain, $\text{Frac}(A)$ denotes its field of fractions.
If no other indication is given the letters $x$ and $y$ will always denote multivariables, $x := (x_1, ..., x_n)$ and $y := (y_1, ..., y_m)$, and $t$ will denote a single variable.
If $f(y)$ is a vector of polynomials with coefficients in a ring $A$,
\[ f(y) := (f_1(y), ..., f_r(y)) \in A[y]^r, \]
if $I$ is an ideal of $A$ and $\overline{y} \in A^m$, then $f(\overline{y}) \in I$ (resp. $f(\overline{y}) = 0$) means $f_i(\overline{y}) \in I$ (resp. $f_i(\overline{y}) = 0$) for $1 \leq i \leq r$.

2. Artin Approximation

In this first part we review the main results concerning the Artin Approximation Property. We give four results that are the most characteristic in the story: the classical Artin Approximation Theorem in the analytic case, its generalization by A. Płoski, a result of J. Denef and L. Lipshitz concerning rings with the Weierstrass Division Property and, finally, Popescu Approximation Theorem.

2.1. The analytic case. — In the analytic case, the first result is due to Michael Artin in 1968 [Ar68]. His result asserts that the set of convergent solutions is dense in the set of formal solutions of a system of implicit analytic equations. This result is particularly useful, since if you have some analytic problem that you can express in a system of analytic equations, in order to find solutions of this problem you only need to find formal solutions and this may be done in general by an inductive process. Another way to use this result is the following: let us assume that you have some algebraic problem and that you are working over a ring of the form $A := k[[x]]$, where $x := (x_1, ..., x_n)$ and $k$ is a characteristic zero field. If the problem involves only a countable number of data (which is often the case in this context), since $\mathbb{C}$ is algebraically closed and the transcendence degree of $\mathbb{Q} \rightarrow \mathbb{C}$ is uncountable, you may assume that you work over $\mathbb{C}[x]$. Using Theorem 2.1, you may, in some cases, reduce the problem to $A = \mathbb{C}\{x\}$. Then you can use powerful methods of complex analytic geometry to solve the problem. This kind of method is used, for instance, in the recent proof of the Nash Conjecture for algebraic surfaces (see Theorem A of [FB12] and the crucial use of this theorem in [FBPP1]) or in the proof of the Abhyankar-Jung Theorem given in [PR12]. Let us mention that C. Chevalley had apparently proven this theorem some years before M. Artin but he did not publish it because he did not find applications of it [Ra].

2.1.1. Artin result. —

**Theorem 2.1.** — [Ar68] Let $k$ be a valued field of characteristic zero 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 \( \tilde{g}(x) \) vanishing at 0,
\[
 f(x, \tilde{g}(x)) = 0.
\]
Then there exists, for any \( c \in \mathbb{N} \), a convergent power series solution \( \tilde{g}(x) \),
\[
 f(x, \tilde{g}(x)) = 0
\]
which coincides with \( \tilde{g}(x) \) up to degree \( c \),
\[
 \tilde{g}(x) \equiv \hat{g}(x) \mod (x)^c.
\]

**Remark 2.2.** — This theorem has been conjectured by S. Lang in [Lan54] (last paragraph p. 372) when \( \mathbb{k} = \mathbb{C} \).

**Remark 2.3.** — The ideal \( (x) \) defines a topology on \( \mathbb{k}[x] \) called the Krull topology induced by the following norm: \( |a(x)| := e^{-\text{ord}(a(x))} \). In this case small elements of \( \mathbb{k}[x] \) are elements of high order. Thus Theorem 2.1 asserts that the set of solutions in \( \mathbb{k}\{x\}^m \) of \( f(x, y) = 0 \) is dense in the set of solutions in \( \mathbb{k}\{x\}^m \) of \( f(x, y) = 0 \) for the Krull topology.

**Proof of Theorem 2.1** — Let us first give the main ideas of the proof. The proof is done by induction on \( n \), the case \( n = 0 \) being obvious.

The first step is to reduce the problem to the case the ideal \( I \) generated by \( f_1, \ldots, f_r \) is a prime ideal by adding to \( I \) all the elements \( g(x, y) \) such that \( g(x, \tilde{g}(x)) = 0 \). Let us denote by \( X \) the analytic set defined by \( I \).

The next step is to reduce to the case \( X \) is complete intersection, this means that \( I \) is generated by \( r \) elements where \( r \) is equal to the codimension of \( X \) in \( \mathbb{k}^{n+m} \).

After these reductions, the proper proof starts. The key ingredient is a suitable minor \( \delta \) of the Jacobian matrix \( \left( \frac{\partial f}{\partial y} \right) \) of \( f \), namely one which is not identically zero on \( X \). The existence of such a minor is ensured by the Jacobian Criterion: at a smooth point of \( X \), the rank of the Jacobian matrix is the codimension of \( X \) at this point. Since the set of smooth points is dense, the assertion follows.

We denote by \( \hat{\delta}(x) := \delta(x, \tilde{g}(x)) \) the evaluation of \( \delta \) at our given formal solution.

Then, the idea is the following: instead of trying to solve \( f(x, y) = 0 \) with a convergent solution, we aim at finding a convergent power series vector \( \hat{y}(x) \) such that \( \hat{\delta}^2(x, \hat{y}(x)) \) divides \( f(x, \hat{y}(x)) \). Since \( f(x, \hat{y}(x)) = 0 \), then \( \hat{\delta}^2(x, \hat{y}(x)) \) already divides \( f(x, \hat{y}(x)) \), we will reformulate the statement \( \hat{\delta}^2(x, y(x))^2 \) divides \( f(x, y(x)) \) as the vanishing of analytic equations defined over \( \mathbb{k}\{x_1, \ldots, x_{n-1}\} \).

By a linear change of coordinates in \( x \) we may transform \( \hat{\delta}(x)^2 \) into a \( x_n \)-regular series of order \( d \). Thus \( \hat{\delta}(x)^2 \) is, up to multiplication by a unit, a monic polynomial in \( x_n \) of degree \( d \) with coefficients in \( \mathbb{k}[x'] \) where \( x' \) denotes the first \( n-1 \) variables \( x_1, \ldots, x_{n-1} \) (by the Weierstrass Preparation Theorem, see Section [A]). We first divide \( \hat{\delta}(x)^2 \) by \( \hat{\delta}(x)^2 \) and work with the remainder of this division. So write \( \hat{y}(x) \equiv \hat{z}(x) \mod (\hat{\delta}(x)^2) = \hat{\delta}(x)^2 \) with \( \hat{z}(x) \) a vector of polynomials in \( x_n \) of degree < \( d \) with coefficients in \( \mathbb{k}[x'] \). A short but technical computation shows that the divisibility of \( f(x, y(x)) \) by \( \hat{\delta}(x, y(x))^2 \) is equivalent to solving a finite system of analytic equations for the coefficients of a vector \( z(x) \) of polynomials in \( x_n \) of degree < \( d \) with coefficients in
Let us now explain the proof in more details. Let us assume that the theorem is proven for \( n \) and let us prove it for \( n + 1 \).

Let \( I \) be the ideal of \( k[x, y] \) generated by \( f_1(x, y), \ldots, f_r(x, y) \). Let \( \varphi \) be the \( k[x]-\)morphism \( k[x, y] \to k[x] \) sending \( y_i \) onto \( \hat{y}_i(x) \). Then \( \text{Ker}(\varphi) \) is a prime ideal containing \( I \) and if the theorem is true for generators of \( \text{Ker}(\varphi) \) then it is true for \( f_1, \ldots, f_r \). Thus we can assume that \( I = \text{Ker}(\varphi) \).

The local ring \( k[x, y]_I \) is regular by a theorem of Serre (see Theorem 19.3 [Mat80]).

Set \( h := \text{height}(I) \). Thus, from the Jacobian Criterion, there exists a \( h \times h \) minor of the Jacobian matrix \( \frac{\partial(f_1, \ldots, f_r)}{\partial(x, y)} \), denoted by \( \delta(x, y) \), such that \( \delta \notin I = \text{Ker}(\varphi) \). In particular we have \( \delta(x, \hat{y}(x)) \neq 0 \).

By considering the partial derivative of \( f_i(x, \hat{y}(x)) = 0 \) with respect to \( x_j \) we get

\[
\frac{\partial f_i}{\partial x_j}(x, \hat{y}(x)) = -\sum_{k=1}^r \frac{\partial y_k}{\partial x_j} \frac{\partial f_i}{\partial y_k}(x, \hat{y}(x)).
\]

Thus there exists a \( h \times h \) minor of the Jacobian matrix \( \frac{\partial(f_1, \ldots, f_r)}{\partial(y_1, \ldots, y_h)} \), still denoted by \( \delta(x, y) \), such that \( \delta(x, \hat{y}(x)) \neq 0 \). In particular \( \delta \notin I \). From now on we will assume that \( \delta \) is the determinant of \( \frac{\partial(f_1, \ldots, f_h)}{\partial(y_1, \ldots, y_h)} \).

If we denote \( J := (f_1, \ldots, f_h) \), then \( \text{ht}(Jk[x, y]_I) \leq h \). On the other hand we have \( \text{ht}(Jk[x, y]_I) \geq \text{rk}(\frac{\partial(f_1, \ldots, f_h)}{\partial(y_1, \ldots, y_h)}) \bmod. I \), and \( h \leq \text{rk}(\frac{\partial(f_1, \ldots, f_h)}{\partial(y_1, \ldots, y_h)}) \bmod. I \) since \( \delta(x, \hat{y}(x)) \neq 0 \). Thus \( \text{ht}(Jk[x, y]_I) = h \) and \( \sqrt{\text{J}(k[x, y]_I)} = \text{Ik}[x, y]_I \). This means that there exists \( q \in k[x, y] \), \( q \notin I \), and \( e \in \mathbb{N} \) such that \( q f_i^e \in J \) for \( h + 1 \leq i \leq m \). In particular \( g(x, \hat{y}(x)) \neq 0 \). We will use this fact later.

Then we will use the following lemma with \( g := \delta^2 \).

**Lemma 2.4.** — Let us assume that Theorem 2.1 is true for an integer \( n - 1 \). Let \( g(x, y) \) be a convergent power series and let \( f(x, y) \) be a vector of convergent power series.

Let \( \bar{y}(x) \) be in \( (x)k[x]^m \) such that \( g(x, \bar{y}(x)) \neq 0 \) and \( f(x, \bar{y}(x)) = 0 \bmod. g(x, \bar{y}(x)) \).

Let \( e \) be an integer. Then there exists \( \bar{y}(x) \in (x)k[x]^m \) such that \( f(x, \bar{y}(x)) = 0 \bmod. g(x, \bar{y}(x)) \) and \( \bar{y}(x) \neq \bar{y}(x) \in (x)^c \).

**Proof of Lemma 2.4.** — If \( g(x, \bar{y}(x)) \) is invertible, the result is obvious (just take for \( \bar{y}_i(x) \) any truncation of \( \bar{y}_i(x) \)). Thus let us assume that \( g(x, \bar{y}(x)) \) is not invertible. By making a linear change of variables we may assume that \( g(x, \bar{y}(x)) \) is regular with respect to \( x_n \) and by Weierstrass Preparation Theorem \( g(x, \bar{y}(x)) = \bar{a}(x) \times \text{unit} \) where

\[
\bar{a}(x) := x_n^d + \bar{a}_1(x)x_n^{d-1} + \cdots + \bar{a}_d(x)
\]
where \( x' := (x_1, \ldots, x_{n-1}) \) and \( a_i(x') \in (x')k[[x']] \), \( 1 \leq i \leq d \).

Let us perform the Weierstrass division of \( \tilde{g}_i(x) \) by \( \tilde{a}(x) \):

\[
\tilde{g}_i(x) = \tilde{a}(x)\tilde{w}_i(x) + \sum_{j=0}^{d-1} \tilde{g}_{i,j}(x')x_n^j
\]

for \( 1 \leq i \leq m \). Let us denote

\[
\tilde{g}_i^*(x) := \sum_{j=0}^{d-1} \tilde{g}_{i,j}(x')x_n^j, \quad 1 \leq i \leq m.
\]

Then \( g(x, \tilde{g}(x)) = g(x, \tilde{g}^*(x)) \) mod. \( \tilde{a}(x) \) and \( f_k(x, \tilde{g}(x)) = f_k(x, \tilde{g}^*(x)) \) mod. \( \tilde{a}(x) \)

for \( 1 \leq k \leq r \).

Let \( y_{i,j}, 1 \leq i \leq m, 1 \leq j \leq d-1 \), be new variables. Let us denote \( y_{i}^* := \sum_{j=1}^{d-1} y_{i,j}x_n^j \), \( 1 \leq i \leq m \). Let us denote the polynomial

\[
A(a_i, x_n) := x_n^d + a_1x_n^{d-1} + \cdots + a_d \in k[x_n, a_1, \ldots, a_d]
\]

where \( a_1, \ldots, a_d \) are new variables. Let us perform the Weierstrass division of \( g(x, y^*) \) and \( f_i(x, y^*) \) by \( A \):

\[
g(x, y^*) = AQ + \sum_{l=1}^{d-1} G_lx_n^l
\]

\[
f_k(x, y^*) = AQ_k + \sum_{l=1}^{d-1} F_{k,l}x_n^l, \quad 1 \leq k \leq r
\]

where \( Q, Q_k \in k\{x, y_{i,j}, a_p\} \) and \( G_l, F_{k,l} \in k\{x', y_{i,j}, a_p\} \).

Then we have

\[
g(x, \tilde{g}^*(x)) = \sum_{l=1}^{d-1} G_l(x', \tilde{y}_{i,j}(x'), \tilde{a}_p(x'))x_n^l \text{ mod. } (\tilde{a}(x))
\]

\[
f_k(x, \tilde{g}^*(x)) = \sum_{l=1}^{d-1} F_{k,l}(x', \tilde{y}_{i,j}(x'), \tilde{a}_p(x'))x_n^l \text{ mod. } (\tilde{a}(x)), \quad 1 \leq k \leq r.
\]

This proves that \( G_l(x', \tilde{y}_{i,j}(x'), \tilde{a}_p(x')) = 0 \) and \( F_{k,l}(x', \tilde{y}_{i,j}(x'), \tilde{a}_p(x')) = 0 \) for all \( k \) and \( l \). By the inductive hypothesis, there exists \( \tilde{y}_{i,j}(x' \in k\{x' \} \) and \( \tilde{a}_p(x') \in k\{x' \} \)

for all \( i, j \) and \( s \), such that \( G_l(x', \tilde{y}_{i,j}(x'), \tilde{a}_p(x')) = 0 \) and \( F_{k,l}(x', \tilde{y}_{i,j}(x'), \tilde{a}_p(x')) = 0 \) for all \( k \) and \( l \) and \( \tilde{y}_{i,j}(x') - \tilde{y}_{i,j}(x'), \tilde{a}_p(x') - \tilde{a}_p(x') \in (x')^c \) for all \( i, j \) and \( p \) (Formally in order to apply the induction hypothesis we should have \( \tilde{y}_{i,j}(0) = 0 \) and \( \tilde{a}_p(0) = 0 \) which is not necessarily the case here. We can remove the problem by replacing \( \tilde{y}_{i,j}(x') \) and \( \tilde{a}_p(x') \) by \( \tilde{y}_{i,j}(x') - \tilde{y}_{i,j}(0) \) and \( \tilde{a}_p(x') - \tilde{a}_p(0) \), and \( G_l(x', y_{i,j}, a_p) \) by \( G(x', y_{i,j} + \tilde{y}_{i,j}(0), a_p + \tilde{a}_p(0)) - \text{idem for } F_{k,l} \)).

Let us denote

\[
\tilde{\alpha}(x) := x_n^d + \tilde{\alpha}_1(x')x_n^{d-1} + \cdots + \tilde{\alpha}_d(x')
\]
\[
\gamma_j(x) := \gamma(x)\bar{\gamma}_i(x) + \sum_{j=0}^{d-1} \gamma_{i,j}(x')x_n^j
\]
for some \(\bar{\gamma}_i(x) \in \mathbb{k}\{x\}\) such that \(\bar{\gamma}_i(x) - \gamma_i(x) \in (x)^c\) for all \(i\). It is straightforward to check that \(f_i(x, \bar{\gamma}(x)) = 0\) mod. \(g(x, \bar{\gamma}(x))\) for \(1 \leq i \leq r\) and \(\gamma_j(x) - \bar{\gamma}_j(x) \in (x)^c\) for \(1 \leq j \leq m\).

We can apply this lemma to \(g(x, y) := \delta(x, y)\) with \(c' := c + d + 1\) and \(d := \text{ord}(\delta^2(x, \bar{\gamma}(x)))\). Thus we may assume that there is \(\gamma_i(x) \in \mathbb{k}\{x\}\), \(1 \leq i \leq m\), such that \(f(x, y) \in \delta^2(x, y)\) and \(\gamma_i(x) - \bar{\gamma}_i(x) \in (x)^{c+d+1}\), \(1 \leq i \leq m\). Since \(\text{ord}(\delta^2(x, y)) = d\), then we have \(f(x, \bar{\gamma}) \in \delta^2(x, \bar{\gamma})(x)^c\). Then we use the following generalization of the Implicit Function Theorem to show that there exists \(\bar{\gamma}(x) \in \mathbb{k}\{x\}^m\) with \(\bar{\gamma}(0) = 0\) such that \(\gamma_j(x) - \bar{\gamma}_j(x) \in (x)^c\), \(1 \leq j \leq m\), and and \(f_i(x, \gamma_i(x)) = 0\) for \(1 \leq i \leq h\).

**Theorem 2.5 (Tougeron Implicit Function Theorem)**

[To72] Let \(f(x, y)\) be a vector of \(\mathbb{k}\{x, y\}^h\) with \(m \geq h\), and let \(\delta(x, y)\) be a \(h \times h\) minor of the Jacobian matrix \(\frac{\partial(f_1, \ldots, f_h)}{\partial(y_1, \ldots, y_m)}\). Let us assume that there exists \(y(x) \in \mathbb{k}\{x\}^m\) such that

\[
f(x, y(x)) \in (\delta(x, y(x)))(x)^c\text{ for all }1 \leq i \leq h
\]

and for some \(c \in \mathbb{N}\). Then there exists \(\gamma(x) \in \mathbb{k}\{x\}^m\) such that

\[
f_i(x, \gamma(x)) = 0 \text{ for all }1 \leq i \leq h,
\]

\[
\gamma(x) - y(x) \in (\delta(x, y(x)))(x)^c.
\]

Moreover \(\gamma(x)\) is unique if we impose \(\gamma_j(x) = y_j(x)\) for \(h < j \leq m\).

If \(c > \text{ord}(q(x, \gamma(x)))\), then \(q(x, \gamma(x)) \neq 0\). Since \(q_j^c \in J\) for \(h + 1 \leq i \leq r\), this proves that \(f_i(x, \gamma(x)) = 0\) for all \(i\).

**Proof of Theorem 2.5** — We may assume that \(\delta\) is the first \(r \times r\) minor of the Jacobian matrix. If we add the equations \(f_{k+1} := y_{k+1} - \bar{y}_{k+1}(x) = 0, \ldots, f_m := y_m - \bar{y}_m(x) = 0\), we may assume that \(m = h\) and \(\delta\) is the determinant of the Jacobian matrix \(J(x, y) := \frac{\partial(f_1, \ldots, f_h)}{\partial(y_1, \ldots, y_m)}\). We have

\[
f(x, y(x) + \delta(x, y(x))z) = f(x, y(x)) + \delta(x, y)zJ(x, y(x)) + \delta(x, y)^2H(x, y(x), z)
\]

where \(z := (z_1, \ldots, z_m)\) and \(H(x, y(x), z) \in \mathbb{k}\{x, y(x), z\}^m\) is of order at least 2 in \(z\). Let us denote by \(J'(x, y(x))\) the comatrix of \(J(x, y(x))\). Let \(\varepsilon(x)\) be in \((x)^\ell\mathbb{k}\{x\}'\) such that \(f(x, y(x)) = \delta^2(x, y(x))\varepsilon(x)\). Then we have

\[
f(x, y(x) + \delta(x, y(x))z) =
\]

\[
= \delta(x, y(x))(\varepsilon(x)J'(x, y(x)) + z + H(x, y(x), z)J'(x, y(x))) J(x, y(x)).
\]

Let us denote

\[
g(x, z) := \varepsilon(x)J'(x, y(x)) + z + H(x, y(x), z)J'(x, y(x)).
\]
Then \( g(0,0) = 0 \) and the matrix \( \frac{\partial g(x,z)}{\partial z}(0,0) \) is the Identity matrix. Thus, by the Implicit Function Theorem, there exists a unique \( z(x) \in k[x]^m \) such that \( f(x,y(x) + \delta(x,y(x))z(x)) = 0 \). This proves the theorem.

\[ \square \]

**Remark 2.6.** We can do the following remarks about the proof of Theorem 2.1:

i) In the case \( n = 1 \) i.e. \( x \) is a single variable, set \( e := \text{ord}(\delta(x,\hat{g}(x))) \). If \( \overline{g}(x) \in k[x]^m \) satisfies \( \hat{g}(x) = g(x) \in (x)^{2e+c} \), then we have

\[
\text{ord}(f(x,\overline{g}(x))) \geq 2e + c
\]

and

\[
\delta(x,\overline{g}(x)) = \delta(x,\hat{g}(x)) \mod. (x)^{2e+c},
\]

thus \( \text{ord}(\delta(x,\overline{g}(x))) = \text{ord}(\delta(x,\hat{g}(x))) = e \). Hence we have automatically

\[
f(x,\overline{g}(x)) \in (\delta(x,\overline{g}(x)))^3(x)^c
\]

since \( k[x] \) is a discrete valuation ring (i.e. if \( \text{ord}(a(x)) \leq \text{ord}(b(x)) \) then \( a(x) \) divides \( b(x) \) in \( k[x] \)).

Thus Lemma 2.4 is not necessary in this case and the proof is quite simple. This fact will be general: approximation results will be easier to obtain, and sometimes stronger, in discrete valuation rings than in more general rings.

ii) In fact, we did not use that \( k \) is a field of characteristic zero, we just need \( k \) to be a perfect field in order to use the Jacobian Criterion. But the use of the Jacobian Criterion is more delicate for non perfect fields. This also will be general: approximation results will be more difficult to prove in positive characteristic. For instance M. André proved Theorem 2.1 in the case \( k \) is a complete field of positive characteristic and replace the use of the Jacobian Criterion by the homology of commutative algebras [An75].

iii) For \( n \geq 2 \), the proof of Theorem [Ar68] uses an induction on \( n \). In order to do it we use the Weierstrass Preparation Theorem. But to apply the Weierstrass Preparation Theorem we need to do a linear change of coordinates in \( k[x] \), in order to transform \( g(x,\hat{g}(x)) \) into a power series \( h(x) \) such that \( h(0,...,0,x_n) \neq 0 \). Then the proof does not adapt to prove similar results in the case of constraints: for instance if \( \hat{y}_1(x) \) depends only on \( x_1 \) and \( \hat{y}_2(x) \) depends only on \( x_2 \), can we find a convergent solution such that \( \hat{y}_1(x) \) depends only on \( x_1 \), and \( \hat{y}_2(x) \) depends only on \( x_2 \)?

Moreover, even if we can use a linear change of coordinates without modifying the constrains, the use of the Tougeron Implicit Function Theorem may remove the constrains. We will discuss these problems in Section 5.

**Corollary 2.7.** Let \( k \) be a valued field of characteristic zero and let \( I \) be an ideal of \( k[x] \). If \( f(y) \in \left( \frac{k(x)}{I_k(x)} \right)^r \), let \( \hat{y} \in \left( \frac{k[x]}{I_k[x]} \right)^m \) be a solution of \( f = 0 \) such that \( \hat{y} \equiv 0 \) modulo \( I + (x) \). Then there exists a solution of \( f = 0 \) in \( \left( \frac{k[x]}{I_k[x]} \right)^m \) denoted by \( \tilde{y} \) such that \( \tilde{y} \equiv 0 \) modulo \( I + (x) \) and \( \tilde{y} - \hat{y} \in (x)^c \left( \frac{k[x]}{I_k[x]} \right)^m \).
Proof. — Set $F_i(x, y) \in \mathbb{k}[x, y]$ such that $F_i(x, y) = f_i(y) \mod I$ for $1 \leq i \leq r$. Let $a_1, \ldots, a_s \in \mathbb{k}[x]$ be generators of $I$. Set $\tilde{w}(x) \in \mathbb{k}[x]^m$ such that $\tilde{w}_j(x) = \hat{y}_j \mod I$ for $1 \leq j \leq m$. Since $f_i(\hat{y}) = 0$ then there exists $\tilde{z}_{i,k}(x) \in \mathbb{k}[x]$, $1 \leq i \leq r$ and $1 \leq k \leq s$, such that

$$F_i(x, \tilde{w}(x)) + a_1\tilde{z}_{i,1}(x) + \cdots + a_s\tilde{z}_{i,s}(x) = 0 \ \forall i.$$  

After Theorem 2.1 there exist $\tilde{w}_j(x)$, $\tilde{z}_{i,k}(x) \in \mathbb{k}[x]$ such that

$$F_i(x, \tilde{w}(x)) + a_1\tilde{z}_{i,1}(x) + \cdots + a_s\tilde{z}_{i,s}(x) = 0 \ \forall i$$

and $\tilde{w}_j(x) - \tilde{w}_j(x) \in (x)^e$ for $1 \leq j \leq m$. Then the images of the $\tilde{w}_j(x)$ in $\frac{\mathbb{k}[x]}{I}$ satisfy the conclusion of the corollary.

2.1.2. Płoski result. — A few years after M. Artin result, A. Płoski strengthened Theorem 2.1 by a careful analysis of the proof. His result yields an analytic parametrization of a piece of the set of solutions of $f = 0$ such that the formal solution $\hat{y}(x)$ is a formal point of this parametrization.

**Theorem 2.8.** — [Pł74] Let $k$ be a valued field of characteristic zero and let $f(x, y)$ be a vector of power series in two in $\mathbb{k}[x, y]^r$. Let $\hat{y}(x)$ be a formal power series solution such that $\hat{y}(0) = 0$,

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

Then there exists a convergent power series solution $y(x, z) \in \mathbb{k}[x, z]^m$, where $z = (z_1, \ldots, z_s)$ are new variables,

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

and a vector of formal power series $\tilde{z}(x) \in \mathbb{k}[x]^s$ with $\tilde{z}(0) = 0$ such that

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

This result obviously implies Theorem 2.1 since we can choose convergent power series $\tilde{z}_i(x), \ldots, \tilde{z}_s(x) \in \mathbb{k}[x]$ such that $\tilde{z}_j(x) - \tilde{z}_j(x) \in (x)^e$ for $1 \leq j \leq s$. Then, by denoting $\hat{y}(x) := y(x, \tilde{z}(x))$, we get the conclusion of Theorem 2.1.

**Sketch of proof of Theorem 2.8** — The proof is very similar to the proof of Theorem 2.1. It is also an induction on $n$. The beginning of the proof is the same, so we can assume that $r = h$ and we need to prove an analogue of Lemma 2.4 with parameters for $g = \delta^2$. But in order to prove it we need to make a slight modification in the proof. Here we will make a linear change of variables and assume that $\delta(x, \hat{y}(x))$ is regular with respect to $x_n$, i.e.

$$\delta(x, \hat{y}(x)) = (x_n^d + \tilde{a}_1(x')x_n^{d-1} + \cdots + \tilde{a}_d(x')) \times \text{unit}.$$  

We will denote

$$\tilde{a}(x) := x_n^d + \tilde{a}_1(x')x_n^{d-1} + \cdots + \tilde{a}_d(x').$$
(in the proof of Theorem 2.1 $\widehat{a}(x)$ denotes the square of $(x^d_n + \widehat{a}_1(x)x^d_{n-1} + \cdots + \widehat{a}_d(x'))$)

Then we divide $\widehat{y}_i(x)$ by $\widehat{a}(x)$ for $1 \leq i \leq h$ and by $\widehat{a}(x)^2$ for $h < i \leq m$:

$$\widehat{y}_i(x) = \widehat{a}(x)\widehat{w}_i(x) + \sum_{j=0}^{d-1} \widehat{y}_{i,j}(x')x^j_n, \ 1 \leq i \leq h,$$

$$\widehat{y}_i(x) = \widehat{a}(x)^2\widehat{z}_i(x) + \sum_{j=0}^{2d-1} \widehat{y}_{i,j}(x')x^j_n, \ h < i \leq m.$$ 

Let us denote

$$\widehat{y}_i^* := \sum_{j=0}^{d-1} \widehat{y}_{i,j}(x')x^j_n, \ 1 \leq i \leq h,$$

$$\widehat{y}_i^* := \sum_{j=0}^{2d-1} \widehat{y}_{i,j}(x')x^j_n, \ h < i \leq m.$$ 

Let $M(x, y)$ be the adjoint matrix of $\frac{\partial(f_1, \ldots, f_h)}{\partial(y_1, \ldots, y_h)}$,

$$M(x, y) \frac{\partial(f_1, \ldots, f_h)}{\partial(y_1, \ldots, y_h)} = \frac{\partial(f_1, \ldots, f_h)}{\partial(y_1, \ldots, y_h)} M(x, y) = \delta(x, y)I_h$$

where $I_h$ is the identity matrix of size $h \times h$. Then we denote

$$g(x, y) := M(x, y)f(x, y) = (g_1(x, y), \ldots, g_h(x, y))$$

where $g$ and $f$ are considered as column vectors. We have

$$0 = f(x, \widehat{y}(x)) = f(x, \widehat{y}^*_1(x) + \widehat{a}(x)\widehat{w}_1(x), \ldots, \widehat{y}^*_m(x) + \widehat{a}(x)\widehat{w}_h(x),$$

$$\widehat{y}^*_{h+1}(x) + \widehat{a}(x)^2\widehat{z}_{h+1}(x), \ldots, \widehat{y}^*_m(x) + \widehat{a}(x)^2\widehat{z}_m(x)) =$$

$$= f(x, \hat{y}^*(x)) + \widehat{a}(x)\frac{\partial(f_1, \ldots, f_h)}{\partial(y_1, \ldots, y_h)}(x, \hat{y}^*(x)) \begin{pmatrix} \widehat{w}_1(x) \\ \vdots \\ \widehat{w}_h(x) \end{pmatrix} +$$

$$+ \widehat{a}(x)^2\frac{\partial(f_1, \ldots, f_h)}{\partial(y_{h+1}, \ldots, y_m)}(x, \hat{y}^*(x)) \begin{pmatrix} \widehat{z}_{h+1}(x) \\ \vdots \\ \widehat{z}_m(x) \end{pmatrix} + \hat{a}(x)^2Q(x)$$

for some $Q(x) \in k[[x]]^h$. Hence $g_k(x, \hat{y}^*(x)) \in (\hat{a}(x)^2)$. As in the proof of Theorem 2.1 we have $\delta(x, \hat{y}^*(x)) \in (\hat{a}(x))$.

Assuming Płoski Theorem for $n - 1$, with the notation of the proof of Lemma 2.4 the solutions $\overline{y}_{i,j}(x')$ and $\pi_p(x')$ are replaced by $\overline{y}_{i,j}(x', t), \pi_p(x', t) \in k\{x, t\}$, $t = (t_1, \ldots, t_s)$, such that $\overline{y}_{i,j}(x') = \overline{y}_{i,j}(x', \tilde{t}(x'))$ and $\pi_p(x') = \pi_p(x', \tilde{t}(x'))$ for some $\tilde{t}(x') \in k[x'][t]$ and

$$g(x, \hat{y}^*(x, t)) \in (\pi(x, t)^2)$$

$$f(x, \hat{y}^*(x, t)) \in (g(x, \hat{y}^*(x, t)))$$
where the entries of the vector $Q(x, z'', u)$ are in $(z'', u)^2$. By multiplying on the left this equality by $M(x, \vec{y}^*(x))$ we obtain

$$M(x, \vec{y}^*(x))F(x, z'', u) = \delta^2(x, \vec{y}^*(x))G(x, z'', u)$$
where the entries of the vector \( G(x, z'', u) \) are convergent power series and \( G(0, 0, 0) = 0 \). By differentiation this equality yields
\[
M(x, \overline{y}^*(x, t)) \frac{\partial (F_1, \ldots, F_h)}{\partial (u_1, \ldots, u_h)} (x, z'', u) = \delta^2(x, \overline{y}^*(x, t)) \frac{\partial (G_1, \ldots, G_h)}{\partial (u_1, \ldots, u_h)} (x, z'', u).
\]
It is easy to check that
\[
\det \left( \frac{\partial (F_1, \ldots, F_h)}{\partial (u_1, \ldots, u_h)} \right) (x, 0, 0) =
= \det \left( \frac{\partial (F_1, \ldots, F_h)}{\partial (y_1, \ldots, y_h)} \right) (x, 0, 0) \delta(x, \overline{y}^*(x, 0))^h = \delta(x, \overline{y}^*(x, 0))^{h+1}.
\]
Since \( \det (M(x, \overline{y}^*(x, t))) = \delta(x, \overline{y}^*(x, t))^{h-1} \) we have
\[
\det \left( \frac{\partial (G_1, \ldots, G_h)}{\partial (u_1, \ldots, u_h)} \right) (x, 0, 0) = 1.
\]
Thus \( \det \left( \frac{\partial (G_1, \ldots, G_h)}{\partial (u_1, \ldots, u_h)} \right) (0, 0, 0) \neq 0 \). Thus the Implicit Function Theorem yields functions \( u_i(x, z'') \in k\{x, z''\}, h < i \leq m, \) such that \( G(x, z'', u(x, z'')) = 0 \). This shows \( F(x, z'', u(x, z'')) = 0 \). Hence we get the result by defining \( y_i(x, z'') := \overline{y}^*_i(x, t) + \delta(x, \overline{y}^*(x, t))u_i(x, z'') \) for \( h < i \leq m \).

**Remark 2.10.** — Let us remark that this result remains true if we replace \( k\{x\} \) by a quotient \( \frac{k[x]}{I} \) as in Corollary 2.7.

**Remark 2.11.** — Let \( I \) be the ideal generated by \( f_1, \ldots, f_r \). The formal solution \( \bar{y}(x) \) of \( f = 0 \) induces a \( k\{x\} \)-morphism \( k\{x, y\} \longrightarrow k[x] \) defined by the substitution of \( \overline{y}(x) \) for \( y \). Then \( I \) is included in the kernel of this morphism thus, by the universal property of the quotient ring, this morphism induces a \( k\{x\} \)-morphism \( \psi : \frac{k\{x, y\}}{I} \longrightarrow k[x] \). On the other hand, any \( k\{x\} \)-morphism \( \psi : \frac{k\{x, y\}}{I} \longrightarrow k[x] \) is clearly defined by substituting for \( y \) a formal power series \( \bar{y}(x) \) such that \( f(x, \bar{y}(x)) = 0 \).

Thus we can reformulate Theorem 2.8 as follows: Let \( \psi : \frac{k\{x, y\}}{I} \longrightarrow k[x] \) be the \( k\{x\} \)-morphism defined by the formal power series solution \( \bar{y}(x) \). Then there exist an analytic \( k\{x\} \)-algebra \( D := k\{x, z\} \) and \( k\{x\} \)-morphisms \( C \longrightarrow D \) (defined via the convergent power series solution \( y(x, z) \) of \( f = 0 \)) and \( D \longrightarrow k[x] \) (defined by substituting \( \bar{z}(x) \) for \( z \)) such that the following diagram commutes:
\[
\begin{array}{ccc}
\frac{k\{x, y\}}{I} & \xrightarrow{\psi} & k[x] \\
\downarrow & & \downarrow \\
k\{x\} & \xrightarrow{\bar{y}} & k[x, z]
\end{array}
\]
We will use and generalize this formulation later (see Theorem 2.17).
2.2. Artin Approximation and Weierstrass Division Theorem. — The proof of Theorem 2.1 uses essentially only two results: the Weierstrass Division Theorem and the Implicit Function Theorem. In particular it is straightforward to check that the proof of Theorem 2.1 remains true if we replace \( k[x, y] \) by \( k(x, y) \), the ring of algebraic power series in \( x \) and \( y \), since this ring satisfies the Weierstrass Division Theorem (cf. \[Laf67\], see Section A) and the Implicit Function Theorem. In \[Ar69\], M. Artin gives a version of Theorem 2.1 in the case of polynomials equations over a field or an excellent discrete valuation ring \( k \), and proves that formal solutions of such equations can be approximated by solutions in the Henselization of the ring of polynomials over \( k \), i.e. in a localization of a finite extension of the ring of polynomials over \( k \). The proof, when \( k \) is an excellent discrete valuation ring, uses Néron \( p \)-desingularization \[Né64\] (see Section 2.3 for a statement of Néron \( p \)-desingularization). This result is very important since it allows to reduce some algebraic problems over complete local rings to local rings which are localization of finitely generated rings over a field or a discrete valuation ring.

For instance, this idea, along with an idea of C. Peskine and L. Szpiro, was used by M. Hochster to reduce problems over complete local rings in characteristic zero to the same problems in positive characteristic. The idea is the following: let us assume that some statement \((T)\) is true in positive characteristic (where you can use the Frobenius map to prove it for instance) and let us assume that there exists an example showing that \((T)\) is not true in characteristic zero. In some cases we can use Artin Approximation Theorem to show the existence of a counterexample to \((T)\) in the Henselization at a prime ideal of a finitely generated algebra over a field of characteristic zero. Since the Henselization is the direct limit of étale extensions, we can show the existence of a counterexample to \((T)\) in a local ring \( A \) which is the localization of a finitely generated algebra over a field of characteristic zero \( k \). If the example involves only a finite number of data in \( A \), then we may lift this counterexample in a ring which is the localization of a finitely generated ring over \( \mathbb{Q} \), and even over \( \mathbb{Z}[\frac{1}{p_1}, ..., \frac{1}{p_s}] \) where the \( p_i \) are prime integers. Finally we may show that this counterexample remains a counterexample to \((T)\) over \( \mathbb{Z}/p_i \mathbb{Z} \) for all but finitely many primes \( p \) by reducing the problem modulo \( p \) (in fact for \( p \neq p_i \) for \( 1 \leq i \leq s \)). This idea was used to prove important results about Intersection Conjectures \[PeSz73\], big Cohen-Macaulay modules \[HR74\], Homological Conjectures \[H75\]. J Denef and L. Lipshitz axiomatized the properties a ring needs to satisfy in order to adapt the proof the main theorem of \[Ar69\] due M. Artin. They called such families of rings Weierstrass Systems. There are two reasons for introducing such rings: the first one is the proof of Theorem 5.16 and the second one is their use in proofs of Strong Artin Approximation results. Independently H. Kurke, G. Pfister, D. Popescu, M. Roczen and T. Mostowski (cf. \[KPPRM78\]) introduced the notion of Weierstrass category which is very similar (see \[KP82\] for a connection between these two notions).

**Definition 2.12.** — \[DeLi80\] Let \( k \) be a field or a discrete valuation ring of maximal ideal \( p \). By a Weierstrass System of local \( k \)-algebras, or a \( W \)-system over \( k \), we mean a family of \( k \)-algebras \( k[[x_1, ..., x_n]] \), \( n \in \mathbb{N} \) such that:
i) For $n = 0$, the $k$-algebra is $k$,

For any $n \geq 1$, $k[x_1, \ldots, x_n]_{(p,x_1,\ldots,x_n)} \subset \mathbb{K}[x_1, \ldots, x_n] \subset \mathbb{K}[x_1, \ldots, x_n]$ and $\mathbb{K}[x_1, \ldots, x_{n+m}] \cap \mathbb{K}[x_1, \ldots, x_n] = \mathbb{K}[x_1, \ldots, x_n]$ for $m \in \mathbb{N}$. For any permutation $\sigma$ of $\{1, \ldots, n\}$ if $f \in \mathbb{K}[x_1, \ldots, x_n]$, then $f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \in \mathbb{K}[x_1, \ldots, x_n]$.

ii) Any element of $\mathbb{K}[x]$, $x = (x_1, \ldots, x_n)$, which is a unit in $\mathbb{K}[x]$, is a unit in $\mathbb{K}[x]$.

iii) If $f \in \mathbb{K}[x]$ and $p$ divides $f$ in $\mathbb{K}[x]$, then $p$ divides $f$ in $\mathbb{K}[x]$.

iv) Let $f \in (p,x)_{(p,x)}$ such that $f \neq 0$. Suppose that $f \notin (p,x_1,\ldots,x_{n-1},x_n^{-1})$. Then for any $g \in \mathbb{K}[x]$ there exist a unique $q \in \mathbb{K}[x]$ and a unique $r \in \mathbb{K}[x_1, \ldots, x_{n-1}]$ with $\deg_{x_n}r < d$ such that $q = rf + r'$. Let $\gamma \in (p,x_1,\ldots,x_{n-1})$ be a Noetherian Henselian regular local ring.

v) If $f(\gamma) = 0$ and $f(\gamma) = 0$, then there exists $g \in \mathbb{K}[y]$ irreducible in $\mathbb{K}[y]$ such that $g(\gamma) = 0$ and such that there does not exist any unit $u(y) \in \mathbb{K}[y]$ with $u(y)g(\gamma) = 0$.

vi) If $\mathbb{K}[x] \rightarrow \mathbb{K}/(p)$ satisfies $\nu$ and $\mathbb{K}[x] \rightarrow \mathbb{K}/(p)$ satisfies $\nu$, then $\mathbb{K}[x] \rightarrow \mathbb{K}/(p)$ satisfies $\nu$.

**Proposition 2.13.** — Let us consider a $W$-system $\mathbb{K}[x]$.

i) For any $n$, $\mathbb{K}[x_1, \ldots, x_n]$ is a Noetherian Henselian regular local ring.

ii) If $f \in \mathbb{K}[x_1, \ldots, x_n]$ and $g := (g_1, \ldots, g_m) \in (p,x)_{(p,x)}$, then $f(x,g(x)) \in \mathbb{K}[x]$.

iii) If $f \in \mathbb{K}[x]$, then $\frac{\partial f}{\partial x} \in \mathbb{K}[x]$.

iv) If $\mathbb{K}[x_1, \ldots, x_n]$ is a family of rings satisfying i)-iv) of Definition 2.12 and if all these rings are excellent, then they satisfy $\nu$ and $\nu$ of Definition 2.12.

**Proof.** — All these assertions are proven in Remark 1.3 [DeLi80], except iv). Thus we prove here iv): let us assume that $\gamma(p) = 0$ and $\gamma \in (p,x)_{(p,x)}$. Let us denote by $I$ the kernel of the $\mathbb{K}[x]$-morphism $\mathbb{K}[x,y] \rightarrow \mathbb{K}[x]$ defined by the substitution of $\gamma$ for $y$ and let us assume that $I \cap \mathbb{K}[x] \neq (0)$. Since $\mathbb{K}[x]$ is excellent, the morphism $\mathbb{K}[x] \rightarrow \mathbb{K}[x]$ is regular. Thus $\text{Frac}(\mathbb{K}[x])$ is a separable extension of $\text{Frac}(\mathbb{K}[x])$, but $\text{Frac}(\mathbb{K}[x])$ is a subfield of $\text{Frac}(\mathbb{K}[x])$, hence $\text{Frac}(\mathbb{K}[x]) \rightarrow \text{Frac}(\mathbb{K}[x])$ is separable. This implies that the field extension $\text{Frac}(k) \rightarrow \text{Frac}(\mathbb{K}[x])$ is a separable field extension. But if for every irreducible $g \in I \cap \mathbb{K}[x]$, there would exist a unit $u(y) \in \mathbb{K}[y]$ with $u(y)g(y) = \sum_{a \in \mathbb{N}} a_0 y^{a_0}$, then the extension $\text{Frac}(k) \rightarrow \text{Frac}(\mathbb{K}[x])$ would be purely inseparable. This proves that Property $\nu$ of Definition 2.12 is satisfied.

The proof that Property $\nu$ of Definition 2.12 is satisfied is identical.

**Example 2.14.** — We give here few examples of Weierstrass systems:

i) If $k$ is a field or a complete discrete valuation ring, the family $\mathbb{K}[x_1, \ldots, x_n]$ is a $W$-system over $k$ (using Proposition 2.13 iv) since complete local rings are excellent rings).
ii) Let $k\langle x_1, \ldots, x_n \rangle$ be the Henselization of the localization of $k[x_1, \ldots, x_n]$ at the maximal ideal $(x_1, \ldots, x_n)$ where $k$ is a field or an excellent discrete valuation ring. Then, for $n \geq 0$, the family $k\langle x_1, \ldots, x_n \rangle$ is a W-system over $k$ (using Proposition 2.13 iv) since the Henselization of an excellent local ring is still excellent - see Proposition C.17).

iii) The family $k\{x_1, \ldots, x_n\}$ (the ring of convergent power series in $n$ variables over a valued field $k$) is a W-system over $k$.

iv) The family of Gevrey power series in $n$ variables over a valued field $k$ is a W-system (see [Br86]).

Then we have the following Approximation result (the case of $k\langle x \rangle$ where $k$ is a field or a discrete valuation ring is proven in [Ar69], the general case is proven in [DeLi80]):

**Theorem 2.15.** — [Ar69, DeLi80] Let $k[[x]]$ be a W-system over $k$, where $k$ is a field or a discrete valuation ring with prime $p$. Let $f \in k[[x,y]]^m$ and $\hat{y} \in (p,x)k[[x]]^m$ satisfy

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

Then, for any $c \in \mathbb{N}$, there exists a convergent power series solution $\tilde{y} \in (p,x)k[[x]]^m$, 

$$f(x, \tilde{y}) = 0 \text{ such that } \tilde{y} - \hat{y} \in (p,x)^c.$$  

Let us mention that Theorem 2.8 extends also for Weierstrass systems (see [Ron10b]).

### 2.3. Néron desingularization and Popescu Theorem

During the 70 and the 80 one of the main goals about Artin Approximation Problem was to find necessary and sufficient conditions on a local ring $A$ for it having the Artin Approximation Property, i.e. such that the set of solutions in $A^m$ of any system of algebraic equations $(S)$ in $m$ variables with coefficients in $A$ is dense for the Krull topology in the set of solutions of $(S)$ in $\hat{A}^m$. Let us recall that the Krull topology on $A$ is the topology induced by the following norm: $|a| := e^{-\ord(a)}$ for all $a \in A \setminus \{0\}$. The problem was to find a way of proving approximation results without using Weierstrass Division Theorem.

**Remark 2.16.** — Let $P(y) \in A[y]$ satisfy $P(0) \in m_A$ and $\frac{\partial P}{\partial y}(0) \notin m_A$. Then, by the Implicit Function Theorem for complete local rings, $P(y)$ has a unique root in $\hat{A}$ equal to 0 modulo $m_A$. Thus if we want being able to approximate roots of $P(y)$ in $\hat{A}$ by roots of $P(y)$ in $A$, a necessary condition is that the root of $P(y)$ constructed by the Implicit Function Theorem is in $A$. Thus it is clear that if a local ring $A$ has the Artin Approximation Property then $A$ is necessarily Henselian.

In fact M. Artin conjectured that a sufficient condition would be that $A$ is an excellent Henselian local ring (Conjecture (1.3) [Ar70]). The idea to prove this conjecture is to generalize Płoski Theorem 2.8 and a theorem of desingularization of A. Néron [Né64]. This generalization is the following (for the definitions see Appendix B):
Theorem 2.17. — [Po85, Po86] Let \( \varphi : A \to B \) be a regular morphism of local Noetherian rings, \( C \) a finitely generated \( A \)-algebra and \( \psi : C \to B \) a morphism of \( A \)-algebras. Then \( \psi \) factors through a finitely generated \( A \)-algebra \( D \) which is smooth over \( A \):

\[
\begin{array}{ccc}
A & \xrightarrow{\varphi} & B \\
\downarrow \psi & & \downarrow \\
C & \xrightarrow{\psi} & D
\end{array}
\]

Historically this theorem has been proven by A. Néron [N64] when \( A \) and \( B \) are discrete valuation rings. Then several authors gave proofs of particular cases (see for instance [Po80, Br83b, ArDe83, ArRo88, or Rot87] - in this last paper the result is proven in the equicharacteristic zero case) until D. Popescu [Po85, Po86] proved the general case. Then, several authors gave simplified proofs or strengthened the result [Og94, Sp99, Sw98, ST]. This result is certainly the most difficult to prove among all the results presented in this paper. We will just give a slight hint of the proof of this result here since there exist very nice presentations of the proof elsewhere (see [Sw98] or [ST] in general, [Qu97] or [Po00] in the equicharacteristic zero case).

Since \( A \to \hat{A} \) is regular if \( A \) is excellent, \( I \) is an ideal of \( A \) and \( \hat{A} := \lim \frac{A}{I^n} \), we get the following result (exactly as Theorem 2.8 implies Theorem 2.1):

Theorem 2.18. — Let \((A, I)\) be an excellent Henselian pair. Let \( f(y) \in A[y]^r \) and \( \hat{y} \in \hat{A}^m \) satisfy \( f(\hat{y}) = 0 \). Then, for any \( c \in \mathbb{N} \), there exists \( \tilde{y} \in A^m \) such that \( \hat{y} - \tilde{y} \in I^c \hat{A} \) and \( f(\tilde{y}) = 0 \).

Proof. — The proof goes as follows: let us denote \( C := \frac{A[y]}{(J)} \) where \( J \) is the ideal generated by \( f_1, \ldots, f_r \). The formal solution \( \hat{y} \in \hat{A} \) defines an \( A \)-morphism \( \hat{\varphi} : C \to \hat{A} \) (see Remark 2.11). By Theorem 2.17 since \( A \to \hat{A} \) is regular (Example B.4), there exists an \( A \)-algebra \( D \) factorizing this morphism. After some technical reductions we may assume that the morphism \( A \to D \) decomposes as \( A \to A[z] \to D \) where \( z = (z_1, \ldots, z_s) \) and \( A[z] \to D \) is standard étale. Let us choose \( \tilde{z} \in \tilde{A}^s \) such that \( \tilde{z} - \tilde{z} \in m_\tilde{A} \tilde{A}^s \) (\( \tilde{z} \) is the image of \( z \) in \( \tilde{A}^s \)). This defines a morphism \( A[z] \to A \). Then \( A \to \frac{D}{(z_1 - \tilde{z}_1, \ldots, z_s - \tilde{z}_s)} \) is standard étale and admits a section in \( \frac{A}{m_A^c} \). Since \( A \) is Henselian, this section lifts to a section in \( A \) by Proposition C.9. This section composed with \( A[z] \to A \) defines a \( A \)-morphism \( D \to A \), and this latter morphism composed with \( C \to D \) yields a morphism \( \tilde{\varphi} : C \to A \) such that \( \tilde{\varphi}(z_i) - \tilde{\varphi}(\tilde{z}_i) \in m_A^c \hat{A} \) for \( 1 \leq i \leq m \).

Remark 2.19. — Let \((A, I)\) be a Henselian pair and let \( J \) be an ideal of \( A \). By applying this result to the Henselian pair \((\frac{A}{I}, \frac{I}{I^2})\) we can prove the following (using the notation of Theorem 2.18): if \( f(\hat{y}) \in J \hat{A} \) then there exists \( \tilde{y} \in A^m \) such that \( f(\tilde{y}) \in J \) and \( \hat{y} - \tilde{y} \in I^c \hat{A} \).
Remark 2.20. — In [Rot90], C. Rotthaus proves the converse of Theorem 2.18 in the local case: if $A$ is a Noetherian local ring that satisfies Theorem 2.18, then $A$ is excellent. In particular Weierstrass systems are excellent local rings. Previously this problem had been studied in [CP81] and [Br83a].

Remark 2.21. — Let $A$ be a Noetherian ring and $I$ be an ideal of $A$. If we assume that $f_1(y), \ldots, f_r(y) \in A[y]$ are linear, then Theorem 2.18 may be proven easily in this case since $A \rightarrow \hat{A}$ is flat (see Example 2). The proof of this flatness result uses the Artin–Rees Lemma.

Example 2.22. — If $A$ is an excellent integral local domain let us denote by $A^h$ its Henselization. Then $A^h$ is the ring of algebraic elements of $\hat{A}$ over $A$. In particular, if $k$ is a field then $k[[x]]$ is the ring of formal power series which are algebraic over $k[x]$. Indeed $A \rightarrow A^h$ is a filtered limit of algebraic extensions, thus $A^h$ is a subring of the ring of algebraic elements of $\hat{A}$ over $A$.

On the other hand if $f \in \hat{A}$ is algebraic over $A$, then $f$ satisfies an equation

$$a_0 f^d + a_1 f^{d-1} + \cdots + a_d = 0$$

where $a_i \in A$ for all $i$. Thus for $c$ large enough there exists $\tilde{f} \in A^h$ such that $\tilde{f}$ satisfies the same polynomial equation and $\tilde{f} - f \in m_A^c$ (by Theorem 2.18 and Theorem C.17).

Since $\bigcap_i m^c_A = (0)$ and a polynomial equation has a finite number of roots, this proves that $\tilde{f} = f$ for $c$ large enough and $f \in A^h$.

Example 2.23. — The strength of this result comes from the fact that it applies to rings that do not satisfy the Weierstrass Preparation Theorem. For example Theorem 2.18 applies to the local ring $B = A[x_1, \ldots, x_n]$ where $A$ is an excellent Henselian local ring (the main example is $A = k[[t]](x)$ where $t$ and $x$ are multivariables). Indeed, this ring is the Henselization of $A[x_1, \ldots, x_n]_{m_A + (x_1, \ldots, x_n)}$. Thus $B$ is an excellent local ring by Example 2.4 and Proposition C.17.

This case was the main motivation of D. Popescu for proving Theorem 2.17 (see also [At70]), since this case implies a nested Artin Approximation result (see Theorem 5.8).

Previous particular cases of this application had been studied before: see [PiP81] for a direct proof that $V[\![x_1, x_2] \!\!]$ satisfies Theorem 2.18 when $V$ is a complete discrete valuation ring, and [BDL83] for the ring $k[[x_1, x_2]][x_3, x_4, x_5]$.

Hint of the proof of Theorem 2.17 — Let $A$ be a Noetherian ring and $C$ be a $A$-algebra of finite type, $C = \frac{A[y_1, \ldots, y_m]}{I f_1, \ldots, f_r}$. We denote by $\Delta_g$ the ideal of $A[y]$ generated by the $h \times h$ minors of the Jacobian matrix $\left(\frac{\partial g_i}{\partial y_j}\right)_{1 \leq i \leq h, 1 \leq j \leq m}$ for $g := (g_1, \ldots, g_h) \in I$. We define the ideal

$$H_{C/A} := \sqrt{\sum_g \Delta_g ((g) : I) C}$$

where the sum runs over all $g := (g_1, \ldots, g_h) \subseteq I$ and $h \in \mathbb{N}$. This ideal is independent of the presentation of $C$ and it defines the singular locus of $C$ over $A$:
Lemma 2.24. — For any \( p \in \text{Spec}(C) \), \( C_\mathfrak{p} \) is smooth over \( A \) if and only if \( \mathcal{H}_{C/A} \not\subset \mathfrak{p} \).

We have the following property:

Lemma 2.25. — Let \( C \) and \( C' \) be two \( A \)-algebras of finite type and let \( A \to C \to C' \) be two morphisms of \( A \)-algebras. Then \( \mathcal{H}_{C'/C} \cap \sqrt{\mathcal{H}_{C/A} C'} = \mathcal{H}_{C'/C} \cap H_{C'/A} \).

The idea of the proof of Theorem 2.17 is the following: if \( \mathcal{H}_{C/A} B \neq B \), then we replace \( C \) by a \( A \)-algebra of finite type \( C' \) such that \( \mathcal{H}_{C/A} B \) is a proper sub-ideal of \( \mathcal{H}_{C'/A} B \).

Using the Noetherian assumption, after a finite number we have \( \mathcal{H}_{C/A} B = B \). Then we use the following proposition:

Proposition 2.26. — Using the notation of Theorem 2.17, let us assume that we have \( \mathcal{H}_{C/A} B = B \). Then \( \psi \) factors as in Theorem 2.17.

Proof of Proposition 2.26. — Let \( (c_1, \ldots, c_s) \) be a system of generators of \( \mathcal{H}_{C/A} \). Then

\[
1 = \sum_{i=1}^{s} b_i \psi(c_i)
\]

for some \( b_i \in B \). Let us define

\[
D := \frac{C[z_1, \ldots, z_s]}{(1 - \sum_{i=1}^{s} c_i z_i)}.
\]

We construct a morphism of \( C \)-algebra \( D \to B \) by sending \( z_i \) onto \( b_i \), \( 1 \leq i \leq s \).

It is easy to check \( D_{c_i} \) is a smooth \( C \)-algebras, thus \( c_i \in H_{D/C} \) by Lemma 2.24 and \( H_{C'/D} \subset H_{D/C} \). By Lemma 2.25 since \( 1 \in H_{C/A} D \), we see that \( 1 \in H_{D/A} \). By Lemma 2.24 this proves that \( D \) is a smooth \( A \)-algebra.

Now to increase the size of \( \mathcal{H}_{C/A} B \) we use the following proposition:

Proposition 2.27. — Using the notation of Theorem 2.17, let \( \mathfrak{p} \) be a minimal prime ideal of \( \mathcal{H}_{C/A} B \). Then there exist a factorization of \( \psi : C \to D \to B \) such that \( D \) is finitely generated over \( A \) and \( \sqrt{\mathcal{H}_{C/A} B} \subset \sqrt{\mathcal{H}_{D/A} B} \not\subset \mathfrak{p} \).

The proof of Proposition 2.27 is done by induction on \( \text{height}(\mathfrak{p}) \). Thus there is two things to prove: first the case \( \text{ht}(\mathfrak{p}) = 0 \) which is equivalent to prove Theorem 2.17 for Artinian rings, then the reduction \( \text{ht}(\mathfrak{p}) = k + 1 \) to the case \( \text{ht}(\mathfrak{p}) = k \). This last case is quite technical, even in the equicharacteristic zero case (i.e. when \( A \) contains \( \mathbb{Q} \), see [Qu97] for a good presentation of this case). In the case \( A \) does not contain \( \mathbb{Q} \) there appear more problems due to the existence of inseparable extensions of residue fields. In this case the André homology is the good tool to handle these problems (see [Sw98]).

3. Strong Artin Approximation

We review here results about the Strong Approximation Property. There is clearly two different cases: the first case is when the base ring is a discrete valuation ring (where life is easy!) and the second case is the general case (where life is less easy).
3.1. Greenberg Theorem: the case of a discrete valuation ring. — Let $V$ be a Henselian discrete valuation ring, $m_V$ its maximal ideal and $K$ be its field of fractions. Let us denote by $\hat{V}$ the $m_V$-adic completion of $V$ and by $\hat{K}$ its field of fractions. If $\text{char}(K) > 0$, let us assume that $K \to \hat{K}$ is a separable field extension (in this case this is equivalent to $V$ being excellent, see Example [B.3 ii) and Example [B.3 iv]).

**Theorem 3.1 (Greenberg Theorem).** — [Gre66] If $f(y) \in V[y]^r$, then there exist $a, b \geq 0$ such that

$\forall c \in \mathbb{N} \quad \forall \tilde{y} \in V^m$ such that $f(\tilde{y}) \in m_V^{ac+b}$

$\exists \tilde{y} \in V^m$ such that $f(\tilde{y}) = 0$ and $\tilde{y} - \tilde{y} \in m_V^c$.

**Sketch of proof.** — We will give the proof in the case $\text{char}(K) = 0$. The result is proved by induction on the height of the ideal generated by $f_1(y), \ldots, f_s(y)$. Let us denote by $I$ this ideal. We will denote by $v$, the $m_V$-adic order on $V$ which is a valuation by assumption.

There exists an integer $e \geq 1$ such that $\sqrt{T^e} \subset I$. Then $f(\tilde{y}) \in m_V^{ce}$ for all $f \in I$ implies that $f(\tilde{y}) \in m_V^c$ for all $f \in \sqrt{T}$ since $V$ is a valuation ring. Moreover if $\sqrt{T} = \mathcal{P}_1 \cap \ldots \cap \mathcal{P}_r$ is prime decomposition of $\sqrt{T}$, then $f(\tilde{y}) \in m_V^{c+1}$ for all $f \in \sqrt{T}$ implies that $f(\tilde{y}) \in m_V^c$ for all $f \in \mathcal{P}_i$ for some $i$. This allows us to assume that $I$ is a prime ideal of $V[y]$.

Let $h$ be the height of $I$. If $h = m + 1$, then $I$ is a maximal ideal of $V[y]$ and thus it contains some non zero element of $V$ denoted by $v$. Then there does not exist $\tilde{y} \in V^m$ such that $f(\tilde{y}) \in m_V^{ce+1}$ for all $f \in I$. Thus the theorem is true for $a = 0$ and $b = v + 1$.

Let us assume that the theorem is proven for ideals of height $h + 1$ and let $I$ be a prime ideal of height $h$. As in the proof of Theorem 2.2, we may assume that $r = h$ and that the determinant of the Jacobian matrix of $f$, denoted by $\delta$, is not in $I$. Let us denote $J := I + (\delta)$. Since $\text{ht}(J) = h + 1$, by the inductive hypothesis, there exist $a, b \geq 0$ such that

$\forall c \in \mathbb{N} \quad \forall \tilde{y} \in V^m$ such that $f(\tilde{y}) \in m_V^{ce+b}$ $\forall f \in J$

$\exists \tilde{y} \in V^m$ such that $f(\tilde{y}) = 0$ $\forall f \in J$ and $\tilde{y}_j - \tilde{y}_j \in m_V^{c+1}$, $1 \leq j \leq m$.

Then let $c \in \mathbb{N}$ and $\tilde{y} \in V^m$ satisfy $f(\tilde{y}) \in m_V^{(2c+1)c+2b}$ for all $f \in I$. If $\delta(\tilde{y}) \in m_V^{ac+b}$, then $f(\tilde{y}) \in m_V^{ac+b}$ for all $f \in J$ and the result is proven by the inductive hypothesis. If $\delta(\tilde{y}) \notin m_V^{ac+b}$, then $f_i(\tilde{y}) \in (\delta(\tilde{y}))^2 m_V^c$ for $1 \leq i \leq r$. Then the result comes from the following result.

**Theorem 3.2 (Tougeron Implicit Function Theorem)**

Let $A$ be a Henselian local ring and $f(y) \in A[y]^r$, $y = (y_1, \ldots, y_m)$, $m \geq r$. Let $\delta(x, y)$ be a $r \times r$ minor of the Jacobian matrix $\frac{\partial(f_1, \ldots, f_r)}{\partial(y_1, \ldots, y_m)}$. Let us assume that there exists $\tilde{y} \in A^m$ such that

$f_i(\tilde{y}) \in (\delta(\tilde{y}))^2 m_A^c$ for all $1 \leq i \leq r$.
and for some \( c \in \mathbb{N} \). Then there exists \( \tilde{y} \in \mathbb{A}^m \) such that
\[
f_i(\tilde{y}) = 0 \quad \text{for all } 1 \leq i \leq r, \quad \text{and} \quad \tilde{y} - \bar{y} \in (\delta(\bar{y}))m^c_A.
\]

**Proof.** — The proof is completely similar to the proof of Theorem 3.1. \( \square \)

In fact we can prove the following result whose proof is identical to the proof of Theorem 3.1.

**Theorem 3.3.** — [Sc83] Let \( V \) be a complete discrete valuation ring and \( f(y, z) \in V[y][z]^\nu \), where \( z := (z_1, \ldots, z_s) \). Then there exist \( a, b \geq 0 \) such that
\[
\forall c \in \mathbb{N}, \quad \exists \tilde{y} \in (mV)^m, \quad \exists \tau \in V^s \text{ such that } f(\tilde{y}, \tau) \in m^{ac+b}_V.
\]

**Remark 3.4.** — M. Greenberg proved this result in order to study \( C_1 \) fields. Previous results about \( C_1 \) fields had been already been studied, in particular by S. Lang in [Lan52] where appeared for the first time a particular case of Artin Approximation Theorem (see Theorem 11 and its corollary in [Lan52]).

**Remark 3.5.** — In the case \( f(y) \) has no solution in \( V \), we can choose \( a = 0 \) and Theorem 3.1 asserts there exists a constant \( b \) such that \( f(y) \) has no solution in \( \frac{V}{m_V} \).

**Remark 3.6.** — The valuation \( \nu \) of \( V \) defines an ultrametric norm on \( K \); we define it as
\[
\frac{V}{z} := e^{\nu(z) - \nu(y)}, \quad \forall y, z \in V \setminus \{0\}.
\]

This norm defines a distance on \( V^m \), for any \( m \in \mathbb{N}^* \), denoted by \( d(.,.) \) and defined by
\[
d(y, z) := \max_{k=1}^m |y_k - z_k|.
\]

Then Theorem 3.1 can be reformulated as a Łojasiewicz Inequality (see [Te12]):
\[
\exists a \geq 1, \quad C > 0 \text{ s.t. } |f(\bar{y})| \geq Cd(f^{-1}(0), \tilde{y})^a \quad \forall \bar{y} \in V^m.
\]

This Łojasiewicz Inequality is well known for algebraic or analytic functions and Theorem 3.1 can be seen as a generalization of this Łojasiewicz Inequality for algebraic or analytic functions defined over \( V \). If \( V = k[l] \) where \( k \) is a field, there is very few results known about the geometry of algebraic varieties defined over \( V \). It is a general problem to extend classical results of differential or analytic geometry over \( \mathbb{R} \) or \( \mathbb{C} \) to this setting. See for instance [HM94], [B-H10] (extension of Rank Theorem), [Reg06] or [FBPP2] (Extension of Curve Selection Lemma), [Hic05] for some results in this direction.

For any \( c \in \mathbb{N} \) let us denote by \( \beta(c) \) the smallest integer such that:

- for all \( \bar{y} \in V^m \) such that \( f(\bar{y}) \in (x)\beta(c) \), there exists \( \tilde{y} \in V^m \) such that \( f(\tilde{y}) = 0 \) and \( \tilde{y} - y \in (x)^c \). Greenberg Theorem asserts that such a function \( \beta : \mathbb{N} \rightarrow \mathbb{N} \) exists and that it is bounded by an affine function. We call this function \( \beta \) the **Greenberg function** of \( f \). We can remark that the Greenberg function is an invariant of the integral closure of the ideal generated by \( f_1, \ldots, f_r \).
Lemma 3.7. — Let us consider \( f(y) \in V[y]^r \) and \( g(y) \in V[y]^q \). Let us denote by \( \beta_f \) and \( \beta_g \) their Greenberg functions. Let \( I \) (resp. \( J \)) be the ideal of \( V[y] \) generated by \( f_1(y), \ldots, f_r(y) \) (resp. \( g_1(y), \ldots, g_q(y) \)). If \( I = J \) then \( \beta_f = \beta_g \). The same is true for Theorem 3.3.

Proof. — Let \( I \) be an ideal of \( V \) and \( \overline{y} \in V^m \). We remark that

\[
 f_1(\overline{y}), \ldots, f_r(\overline{y}) \in I \iff g(\overline{y}) \in I \quad \forall g \in I.
\]

Then by replacing \( I \) by \((0)\) and \( m_\overline{y} \), for all \( c \in \mathbb{N} \), we see that \( \beta_f \) depends only on \( I \).

Now, for any \( c \in \mathbb{N} \), we have:

\[
 g(\overline{y}) \in m_\overline{y}^c \quad \forall g \in I \iff \nu(g(\overline{y})) \geq c \quad \forall g \in I \iff g(\overline{y}) \in m_\overline{y}^c \quad \forall g \in I.
\]

Thus \( \beta_f \) depends only on \( I \).

\[\square\]

In general, it is a difficult problem to compute the Greenberg function of an ideal \( I \). It is even a difficult problem to bound this function in general. If we analyze carefully the proof of Greenberg Theorem, using classical effective results in commutative algebra, we can prove the following result:

Theorem 3.8. — [Ron10a] Let \( k \) be a characteristic zero field and \( V := k[t] \) where \( t \) is a single variable. Then there exists a function

\[
 N^2 \to \mathbb{N}
\]

\[
 (m, d) \mapsto a(m, d)
\]

which is a polynomial function in \( d \) whose degree is exponential in \( m \), such that for any vector \( f(y) \in k[t,y]^r \) of polynomials of total degree \( \leq d \), the Greenberg function of \( f \) is bounded by \( c \mapsto a(m, d)(c + 1) \). Here \( m \) denotes the size of \( y \).

Moreover let us remark that, in the proof of Theorem 3.1, we proved a particular case of the following inequality:

\[
 \beta_f(c) \leq 2\beta_J(c) + c, \quad \forall c \in \mathbb{N}
\]

where \( J \) is the Jacobian Ideal of \( I \) (for a precise definition of the Jacobian Ideal in general and a general proof of this inequality let see [El73]). The coefficient 2 comes from the use of Tougeron Implicit Function Theorem. We can sharpen this bound in the following particular case:

Theorem 3.9. — [Hic93] Let \( k \) be an algebraically closed field of characteristic zero and \( V := k[t] \) where \( t \) is a single variable. Let \( f(y) \in V[y]^r \) be one power series. Let us denote by \( J \) the ideal of \( V[y] \) generated by \( f(y), \frac{\partial f}{\partial y_1}(y), \frac{\partial f}{\partial y_2}(y), \ldots, \frac{\partial f}{\partial y_m}(y) \), and let us denote by \( \beta_f \) the Greenberg function of \( f \) and by \( \beta_J \) the Greenberg function of \( J \). Then

\[
 \beta_f(c) \leq \beta_J(c) + c, \quad \forall c \in \mathbb{N}.
\]
This bound may be used to find sharp bounds of some Greenberg functions (see Remark 3.11).

On the other hand we can describe the behaviour of $\beta$ in the following case:

**Theorem 3.10.** — [De84, DeLo99] Let $V$ be $\mathbb{Z}_p$ or a Henselian discrete valuation ring whose residue field is an algebraically closed field of characteristic zero. Let us denote by $m_V$ the maximal ideal of $V$. Let us denote by $\beta$ the Artin function of $f(y) \in V[y]^r$. Then there exists a finite partition of $\mathbb{N}$ in congruence classes such that on each such congruence class the function $c \mapsto \beta(c)$ is linear for $c$ large enough.

**Hints on the proof in the case the residue field has characteristic zero**

Let us consider the following first order language of three sorts:

1. the field $(K := \text{Frac}(V), +, \times, 0, 1)$
2. the group $(\mathbb{Z}, +, <, \equiv_d (\forall d \in \mathbb{N}^*), 0)$ ($\equiv_d$ is the relation $a \equiv_d b$ if and only if $a - b$ is divisible by $d$ for $a, b \in \mathbb{Z}$)
3. the residue field $(k := \text{Frac}\left(\frac{V}{m_V}\right), +, \times, 0, 1)$

with both following functions:

a) $\nu : K \to \mathbb{Z}^*$

b) $ac : K \to k$ ("angular component")

The function $\nu$ is the valuation of the valuation ring $V$. The function $ac$ may be characterized by axioms, but here let us just give an example: let us assume that $V = \mathbb{k}[t]$. Then $ac$ is defined by $ac(0) = 0$ and $ac(\sum_{n=n_0}^{\infty} a_n t^n) = a_{n_0}$ if $a_{n_0} \neq 0$.

The second sort $(\mathbb{Z}, +, <, \equiv_d, 0)$ admits elimination of quantifiers [Pr29] and the elimination of quantifiers of $(k, +, \times, 0, 1)$ is a classical result of Chevalley. J. Pas proved that the first sort language and the three sorted language admits elimination of quantifiers [Pa89]. This means that any subset of $\mathbb{K}^{n_1} \times \mathbb{Z}^{n_2} \times \mathbb{K}^{n_3}$ defined by a first order formula in this three sorts language (i.e. a logical formula involving $0, 1, +, \times$ (but not $a \times b$ where $a$ and $b$ are integers), $\{,\}, =, <, \wedge, \vee, \neg, \exists, \forall, \nu, ac$, and variables for elements of $K, \mathbb{Z}$ and $k$ may be defined by a formula involving the same symbols except $\forall, \exists$.

Then we see that $\beta$ is defined by the following logical sentence:

$$\forall c \in \mathbb{N} \; \exists \bar{y} \in \mathbb{K}^m (\nu(f(\bar{y})) \geq \beta(c)) \wedge (\nu(\bar{y}) \geq 0) \; \exists \bar{y} \in \mathbb{K}^m (f(\bar{y}) = 0 \wedge \nu(\bar{y} - \bar{y}) \geq c]$$

Applying the latter elimination of quantifiers result we see that $\beta(c)$ may be defined without $\forall$ and $\exists$. Thus $\beta(c)$ is defined by a formula using $+, <, \equiv_d$ (for a finite set of integers $d$). This proves the result.

The case where $V = \mathbb{Z}_p$ needs more work since the residue field of $\mathbb{Z}_p$ is not algebraically closed, but the idea is the same. 

**Remark 3.11.** — When $V = \mathbb{C}\{t\}$, $t$ being a single variable, it is tempting to link together the Greenberg function of a system of equations with coefficients in $V$ and
some geometric invariants of the germ of complex set defined by this system of equations. This has been done in several cases:

i) In [El89], a bound (involving the multiplicity and the Milnor number) of the Greenberg function is given when the system of equations defines a curve in $\mathbb{C}^m$.

ii) Using Theorem $3.9$ [Hic93] gives the following bound of the Greenberg function $\beta$ of a germ of complex hypersurface with an isolated singularity: $\beta(c) \leq \lfloor \lambda c \rfloor + c$ for all $c \in \mathbb{N}$, and this bound is sharp for plane curves. Here $\lambda$ denotes the Łojasiewicz exponent of the germ, i.e.

$$\lambda := \inf \{ \theta \in \mathbb{R} / \exists C > 0 \exists U \text{ neighborhood of } 0 \text{ in } \mathbb{C}^m, |f(z)| + \left| \frac{\partial f}{\partial z_1}(z) \right| + \cdots + \left| \frac{\partial f}{\partial z_m}(z) \right| \geq C|z|^\theta \forall z \in U \}.$$ 

iii) [Hic04] makes the complete computation of the Greenberg function of a branch of plane curve and proves that it is a topological invariant. This computation has been done for several branches in [Sa10]. Some particular cases depending on the Newton polygon of the plane curve singularity are computed in [Wa78].

Finally we mention the following recent result that extends Theorem $3.1$ to non-Noetherian valuation rings and whose proof is based on ultraproducts methods used in [BDLvdD79] to prove Theorem $3.1$ (see $3.3$):

**Theorem 3.12.** — [M-B11] Let $V$ be a Henselian valuation ring and $\nu : V \rightarrow \Gamma$ its associated valuation. Let us denote by $\hat{V}$ its $m_V$-adic completion, $K := \text{Frac}(V)$ and $\hat{K} := \text{Frac}(\hat{V})$. Let us assume that $K \rightarrow \hat{K}$ is a separable extension. Then for any $f(y) \in V[y]^r$ there exist $a \in \mathbb{N}$, $b \in \Gamma^+$ such that

$$\forall c \in \Gamma \forall \bar{y} \in V^m (\nu(f(\bar{y})) \geq ac + b) \implies \exists \tilde{y} \in V^m (f(\tilde{y}) = 0 \land \nu(\tilde{y} - \bar{y}) \geq c).$$

**3.2. Strong Artin Approximation Theorem: the general case.** — In the general case (when $V$ is not a valuation ring), there still exists an approximation function $\beta$. We have the following results:

**Theorem 3.13.** — [Ar68] [BDLvdD79] Let $k$ be a field. For all $n, m, d \in \mathbb{N}$, there exists a function $\beta_{n,m,d} : \mathbb{N} \rightarrow \mathbb{N}$ such that the following holds:

Set $x := (x_1, \ldots, x_n)$ and $y := (y_1, \ldots, y_m)$. Then for all $f(x, y) \in k[x,y]^r$ of total degree $\leq d$, for all $c \in \mathbb{N}$, for all $\bar{y}(x) \in k[x]^m$ such that

$$f(x, \bar{y}(x)) \in (x)^{\beta_{n,m,d}(c)},$$

there exists $\tilde{y}(x) \in k[x]^m$ such that $f(\tilde{y}(x)) = 0$ and $\tilde{y}(x) - \bar{y}(x) \in (x)^c$.

**Remark 3.14.** — By following the proof of M. Artin, D. Lascar proved that there exists a recursive function $\beta$ that satisfies the conclusion of Theorem $3.13$ [Las78]. But the proof of Theorem $3.13$ uses a double induction on the height of the ideal (like in Theorem $3.1$) and on $n$ (like in Theorem $2.1$). In particular, in order to apply the Jacobian Criterion, we need to work with prime ideals, and replace the
original ideal \(I\) generated by \(f_1, \ldots, f_r\) by one of its associated prime and then make a reduction to \(n - 1\) variables. But the bounds of the degree of the generators of such associated prime may be very large compared to the degree of the generators of \(I\). This is essentially the reason why the proof of this theorem does not give much more information about the growth of \(\beta\) than Lascar result.

**Theorem 3.15.** — Let \(A\) be a complete local ring whose maximal ideal is denoted by \(m_A\). Let \(f(y, z) \in A[y][z]^r\), with \(z := (z_1, \ldots, z_s)\). Then there exists a function \(\beta : \mathbb{N} \to \mathbb{N}\) such that the following holds:

For any \(c \in \mathbb{N}\) and any \(\bar{y} \in (m_A, A)^m\) and \(\bar{z} \in A^*\) such that \(f(\bar{y}, \bar{z}) \in m_A^{\beta(c)}\), there exists \(\bar{y} \in (m_A, A)^m\) and \(\bar{z} \in A^*\) such that \(f(\bar{y}, \bar{z}) = 0\) and \(\bar{y} - \bar{y}, \bar{z} - \bar{z} \in m_A^\beta\).

**Example 3.16.** — Set \(f(x_1, x_2, y_1, y_2) := x_1 y_1^2 - (x_1 + x_2) y_2^2\). Set

\[
\sqrt{1 + t} = 1 + \sum_{n \geq 1} a_n t^n \in \mathbb{C}[t]
\]

be the power series such that \(\sqrt{1 + t^2} = 1 + t\). For any \(c \in \mathbb{N}\) set \(y_1^{(c)}(x) := x_1^c\) and \(y_1^{(c)}(x) := x_1^c + \sum_{n=1}^\infty a_n x_1^{-n} x_2^n\). Then

\[
f(x_1, x_2, y_1^{(c)}(x), y_2^{(c)}(x)) \in (x_2^c).
\]

On the other side the equation \(f(x_1, x_2, y_1(x), y_2(x)) = 0\) has no other solution \((y_1(x), y_2(x)) \in k[x, y]^2\) but \((0, 0)\). This proves that Theorem 3.15 is not valid for general Henselian pairs since \((k[x_1, x_2], (x_2))\) is a Henselian pair.

Let us notice that L. Moret-Bailly proved that if a pair \((A, I)\) satisfies Theorem 3.15 then \(A\) has to be an excellent Henselian local ring \([M-B07]\). On the other hand A. More proved that a pair \((A, I)\), where \(A\) is an equicharacteristic excellent regular local ring, satisfies Theorem 3.15 if and only if \(I\) is \(m\)-primary \([M013]\).

It is still an open question to know under which conditions on \(I\) the pair \((A, I)\) satisfies Theorem 3.15 when \(A\) is an excellent Henselian local ring.

**Corollary 3.17.** — Let \(A\) be an excellent Henselian local ring whose maximal ideal is denoted by \(m_A\) and let \(f(y) \in A[y]^r\). Then there exists a function \(\beta : \mathbb{N} \to \mathbb{N}\) such that:

\[
\forall c \in \mathbb{N}, \forall \bar{y} \in A^m \text{ such that } f(\bar{y}) \in m_A^{\beta(c)} \quad \exists \bar{y} \in A^m \text{ such that } f(\bar{y}) = 0 \text{ and } \bar{y} - \bar{y} \in m_A^{\beta}.
\]

**Corollary 3.18.** — Let \(k[x]\) be a W-system over \(k\), where \(k\) is a field or a discrete valuation ring with prime \(p\). Let \(f(x, y) \in k[x, y]^r\). Then there exists a function \(\beta : \mathbb{N} \to \mathbb{N}\) such that for any \(c \in \mathbb{N}\) and any \(\bar{y} \in (p, x)k[x]^m\) such that \(f(x, \bar{y}) \in (p, x)^{\beta(c)}\), there exists \(\bar{y} \in (p, x)k[x]^m\) such that \(f(x, \bar{y}) = 0\) and \(\bar{y} - \bar{y} \in (p, x)^\beta\).

**Proof.** — We first apply Theorem 3.15 then we apply Theorem 2.15

**Remark 3.19.** — As for Theorem 3.1, Corollary 3.17 implies that, if \(f(y)\) has no solution in \(A\), there exists a constant \(c\) such that \(f(y)\) has no solution in \(A/m_A\)
Definition 3.20. — Let $f$ be as in Theorem 3.15 or Corollary 3.17. The least function $\beta$ that satisfies these theorems is called the Artin function of $f$.

Remark 3.21. — As before, the Artin function of $f$ depends only on the integral closure of the ideal $I$ generated by $f_1, \ldots, f_r$ (see Lemma 3.7).

Remark 3.22. — Let $f(y) \in A[y]^r$ and $\overline{y} \in (m_A)^m$ satisfy $f(\overline{y}) \in m_A^c$ and let us assume that $A \rightarrow B := \frac{A[y]}{(f(y))}$ is a smooth morphism. This morphism is local thus it splits as $A \rightarrow C := A[\frac{z}{m_A + (z)}] \rightarrow B$ such that $C \rightarrow B$ is étale (see Definition C.5) and $z := (z_1, \ldots, z_s)$. We remark that $\overline{y}$ defines a morphism of $A$-algebras $\varphi : B \rightarrow \frac{A}{m_A}$. Let us choose any $\tilde{z} \in A^r$ such that $\tilde{z}_i - \tilde{z}_i \in m_A^c$ for all $1 \leq i \leq s$ ($\tilde{z}_i$ denotes the image of $z_i$ in $\frac{A}{m_A}$). Then $A \rightarrow \frac{B}{(z_1 - \tilde{z}_1, \ldots, z_s - \tilde{z}_s)}$ is étale and admits a section in $\frac{A}{m_A}$. By Proposition C.9, this section lifts to a section in $A$. Thus we have a section $B \rightarrow A$ equal to $\varphi$ modulo $m_A$. This proves that $\beta(c) = c$ when $A \rightarrow A[y]/(f(y))$ is smooth.

3.3. Ultraproducts and proofs of Strong Approximation type results. — Historically, M. Artin proved Theorem 3.13 in [Ar69] by slightly modifying the proof of Theorem 2.1, i.e. by an induction on $n$ using the Weierstrass Division Theorem. Then some people tried to prove this kind of result in the same way, but this was not always easy, in particular when the base field was not a characteristic zero field (for example there is a gap in the inseparable case of [PfPo75]). Then four people introduced the use of ultraproducts to give easy proofs of this kind of Strong Approximation type results ([BDLvdD79] and [DeLi80]; see also [Po79] for the general case). The general principle is the following: ultraproducts reduce Strong Artin Approximation Problems to Artin Approximation Problems. We will present here the main ideas.

Let us start with some terminology. A filter $D$ (over $\mathbb{N}$) is a non empty subset of $\mathcal{P}(\mathbb{N})$ that satisfies the following properties:

a) $\emptyset \notin D$,  

b) $\mathcal{E}, \mathcal{F} \in D \implies \mathcal{E} \cap \mathcal{F} \in D$,  

$c) \mathcal{E} \in D, \mathcal{E} \subset \mathcal{F} \implies \mathcal{F} \in D$.

A filter $D$ is principal if $D = \{\mathcal{F} / \mathcal{E} \subset \mathcal{F}\}$ for some subset $\mathcal{E}$ of $\mathbb{N}$. A ultrafilter is a filter which is maximal for the inclusion. It is easy to check that a filter $D$ is an ultrafilter if and only if for any subset $\mathcal{E}$ of $\mathbb{N}$, $D$ contains $E$ or its complement $\mathbb{N} - \mathcal{E}$. In the same way a ultrafilter is non-principal if and only if it contains the filter $E := \{\mathcal{E} \subset \mathbb{N} / \mathbb{N} - \mathcal{E} \text{ is finite}\}$. Zorn Lemma yields the existence of non-principal ultrafilters.

Let $A$ be a Noetherian ring. Let $D$ be a non-principal ultrafilter. We define the ultrapower (or ultraproduct) of $A$ as follows:

$$A^* := \{(a_i)_{i \in \mathbb{N}} \in \prod_A \mid (a_i) \sim (b_i) \text{ iff } \{i / a_i = b_i \} \in D\}.$$
We have a morphism $A \to A^*$ that sends $a$ onto the class of $(a)_{i \in \mathbb{N}}$. We have the following fundamental result:

**Theorem 3.23.** — [CK73] Let $L$ be a first order language, let $A$ be a structure for $L$ and let $D$ be an ultrafilter over $\mathbb{N}$. Then for any $(a_i)_{i \in \mathbb{N}} \in A^*$ and for any first order formula $\varphi(x), \varphi((a_i))$ is true in $A^*$ if and only if $\{i \in \mathbb{N} / \varphi(a_i) \text{ is true in } A\} \in D$.

In particular we can deduce the following properties:

The ultrapower $A^*$ is equipped with a structure of commutative ring. If $A$ is a field then $A^*$ is a field. If $A$ is an algebraically closed field then $A^*$ is an algebraically closed field. If $A^*$ is a local ring with maximal ideal $m_A$ then $A^*$ is a local ring with maximal ideal $m^*_A$ defined by $(a_i)_i \in m^*_A$ if and only if $\{i/a_i \in m_A\} \in D$. If $A$ is a local Henselian ring, then $A^*$ is a local Henselian ring. In fact all these facts are elementary and can be checked directly by hand. Elementary proofs of these results can be found in [BDLvdD79].

Nevertheless if $A$ is Noetherian, then $A^*$ is not Noetherian in general, since Noetherianity is a condition on ideals of $A$ and not on elements of $A$. For example, if $A$ is a Noetherian local ring, then $m^*_\infty := \bigcap_{n>0} m^*_A \neq (0)$ in general. But we have the following lemma:

**Lemma 3.24.** — [Po00] Let $(A, m_A)$ be a Noetherian complete local ring. Let us denote $A_1 := \frac{A}{m^*_A}$. Then $A_1$ is a Noetherian complete local ring of same dimension as $A$ and the composition $A \to A^* \to A_1$ is flat.

In fact, since $A$ is excellent and $m_A A_1$ is the maximal ideal of $A_1$, it is not difficult to prove that $A \to A_1$ is regular. Details can be found in [Po00].

Let us sketch the idea in the case of Theorem 3.17:

**Sketch of the proof of Theorem 3.17** — Let us assume that some system of algebraic equations over an excellent Henselian local ring $A$, denoted by $f = 0$, does not satisfy Theorem 3.17. Using Theorem 2.18 we may assume that $A$ is complete. Thus it means that there exist an integer $c_0 \in \mathbb{N}$ and $\bar{y}^{(c)} \in A^m, \forall c \in \mathbb{N}$, such that $f(\bar{y}^{(c)}) \in m_A^c$ and there does not exist $\bar{y} \in (A^*)^m$ such that $f(\bar{y}) = 0$ and $\bar{y}^{(c)} \in m_A^c$.

Let us denote by $\bar{y}$ the image of $(\bar{y}^{(c)})_c$ in $(A^*)^m$. Since $f(\bar{y}) \in A[\bar{y}]^m$, we may assume that $f(\bar{y}) \in A^*[\bar{y}]$ using the morphism $A \to A^*$. Then $f(\bar{y}) \in m^*_\infty$. Thus $f(\bar{y}) = 0$ in $A_1$. Let us choose $c > c_0$. Since $A \to A_1$ is regular and $A$ is Henselian, following the proof of Theorem 2.18 for any $c \in \mathbb{N}$ there exists $\tilde{y} \in A^m$ such that $f(\tilde{y}) = 0$ and $\tilde{y} - \bar{y} \in m_A^c A_1$. Thus $\bar{y} - \bar{y} \in m_A^c A^*$. Hence the set $\{i \in \mathbb{N} / \bar{y} \bar{y}^{(i)} = m_A^c A^* \}$ is non-empty. This is a contradiction.

We can also prove easily the following proposition with the help of ultraproducts:

**Proposition 3.25.** — [BDLvdD79] Let $f(x, y) \in \mathbb{C}[x, y]^n$. For any $1 \leq i \leq m$ let $J_i$ be a subset of $\{1, \ldots, n\}$.

Let us assume that, for any $c \in \mathbb{N}$, there exist $\bar{y}^{(c)}_i(x) \in \mathbb{C}[x, j \in J_i], 1 \leq i \leq m$, such
that
\[ f(x, \overline{y}^{(c)}(x)) \in (x)^c. \]

Then there exist \( \overline{y}_i(x) \in \mathbb{C}[x, j \in J_i], 1 \leq i \leq m, \) such that \( f(x, \overline{y}(x)) = 0. \)

**Proof.** — Let us denote by \( \overline{y} \in \mathbb{C}[x]^* \) the image of \( (\overline{y}^{(c)})^c. \) Then \( f(x, \overline{y}) = 0 \) modulo \( (x)^{\infty}. \) It is not very difficult to check that \( \frac{C[x]}{(x)^c} \simeq \mathbb{C}[[x]] \) as \( \mathbb{C}[[x]]\)-algebras. Moreover \( \mathbb{C}^* \simeq \mathbb{C} \) as \( \mathbb{k}\)-algebras (where \( \mathbb{k} \) is the subfield of \( \mathbb{C} \) generated by the coefficients of \( f \)), since they are algebraically closed field of same transcendence degree over \( \mathbb{Q} \) and same characteristic. Then the image of \( \overline{y} \) by the isomorphism yields the desired solution in \( \mathbb{C}[x]. \)

Let us remark that the proof of this result remains valid if we replace \( \mathbb{C} \) by any algebraically closed field \( \mathbb{k} \) whose cardinal is strictly greater that the cardinal of \( \mathbb{N}. \) If we replace \( \mathbb{C} \) by \( \mathbb{Q}, \) this result is no more true in general (see Example 5.23).

**Remark 3.26.** — Several authors proved "uniform" Strong Artin approximation results, i.e. they proved the existence of a function \( \beta \) satisfying Theorem 3.15 for a family of \((f_\lambda(y, z))_{\lambda \in \Lambda}\) which satisfy tameness properties that we do not describe here. The main example is Theorem 3.13 that asserts that the Artin functions of polynomials in \( n + m \) variables of degree less than \( d \) are uniformly bounded. There are also two types of proof for these kind of "uniform" Strong Artin approximation results : the ones using ultraproducts (see Theorem 4.2 of [BDLvdD79] which is a generalization of Theorem 3.13 where the base field is not fixed, or Theorems 8.2 and 8.4 of [DeLi80] where uniform Strong Artin approximation results are proven for families of polynomials whose coefficients depend analytically on some parameters) and the ones using the scheme of proof due to Artin (see [ElTo96] where more or less the same results as those of [BDLvdD79] and [DeLi80] are proven).

### 3.4. Effectivity of the behaviour of Artin functions: some examples.

In general the proofs of Strong Artin Approximation results do not give much information about the Artin functions, since ultraproducts methods use a proof by contradiction (see also Remark 3.14). The problem of finding estimates of Artin functions was raised first in [Ar70] and very few general results are known (the only ones in the case of Greenberg Theorem are Theorems 3.9, 3.10 and Remark 3.11 and Remark 3.14 in the general case). We give here a list of examples for which we can give non trivial effective behaviour about their Artin function.

#### 3.4.1. Artin-Rees Lemma.

The following result has been known for long by the specialists and has been communicated to the author by M. Hickel:

**Theorem 3.27.** — [Ron06a] Let \( f(y) \in A[y]^* \) be a vector of linear polynomials with coefficients in a Noetherian ring \( A. \) Let \( I \) be an ideal of \( A. \) Then there exists a constant \( c_0 \geq 0 \) such that:

\[
\forall c \in \mathbb{C} \quad \forall \overline{y} \in A^m \text{ such that } f(\overline{y}) \in I^{c+c_0} \\
\exists \tilde{y} \in A^m \text{ such that } f(\tilde{y}) = 0 \text{ and } \tilde{y} - \overline{y} \in I^c.
\]
This theorem asserts that the function $\beta$ of Theorem \ref{thm:equicharacteristic} is bounded by the function $c \mapsto c + c_0$. Moreover let us remark that this theorem is valid for general Noetherian ring and general ideals $I$ if $A$. This can be compared with the fact that, for linear equations, Theorem \ref{thm:equicharacteristic} is true for any Noetherian ring $A$ without Henselian condition (see Remark \ref{rem:nohenselian}).

Proof. — For convenience, let us assume that there is only one linear polynomial:

$$f(y) = a_1 y_1 + \cdots + a_m y_m.$$ Let us denote by $I$ the ideal of $A$ generated by $a_1, \ldots, a_m$. Artin-Rees Lemma implies that there exists $c_0 > 0$ such that $I \cap I^{c+c_0} \subset I_I$ for any $c \geq 0$. If $\bar{y} \in A^m$ is such that $f(\bar{y}) \in I^{c+c_0}$ then, since $f(\bar{y}) \in I$, there exists $\epsilon \in (I^c)^m$ such that $f(\bar{y}) = f(\epsilon)$. If we define $\tilde{y}_i := \bar{y}_i - \epsilon_i$, for $1 \leq i \leq m$, we have the result. 

We have the following result whose proof is similar:

**Proposition 3.28.** — Let $(A, m_A)$ be a Henselian excellent local ring, $I$ an ideal of $A$ generated by $a_1, \ldots, a_q$ and $f(y) \in A[y]^r$. Set

$$F_i(y, z) := f_i(y) + a_1 z_{i,1} + \cdots + a_q z_{i,q} \in A[y, z], \; 1 \leq i \leq r$$

where the $z_{i,k}$ are new variables and let $F(y, z)$ be the vector whose coordinates are the $F_i(y, z)$ . Let us denote by $\beta$ the Artin function of $f(y)$ seen as a vector of polynomials of $\frac{1}{q}[y]$ and $\gamma$ the Artin function of $F(y, z) \in A[y, z]^r$. Then there exists a constant $c_0$ such that:

$$\beta(c) \leq \gamma(c) \leq \beta(c + c_0), \; \forall c \in \mathbb{N}.$$ 

Proof. — Let $\bar{y} \in A^m$ satisfies $f(\bar{y}) \in m_A^{\beta(c)} \frac{1}{q}$. Then there exists $\bar{z} \in A^{qr}$ such that $F(\bar{y}, \bar{z}) \in m_A^{\beta(c)}$ (we still denote by $\bar{y}$ a lifting of $\bar{y}$ in $A^m$). Thus there exists $\tilde{y} \in A^m$ and $\tilde{z} \in A^{qr}$ such that $F(\tilde{y}, \tilde{z}) = 0$ and $\tilde{y} - \bar{y}, \tilde{z} - \bar{z} \in m_A^{c_0}$. Thus $f(\tilde{y}) = 0$ in $\frac{1}{q}I$. Let $c_0$ be a constant such that $I \cap m_A^{c+c_0} \subset I m_A^{c_0}$ for all $c \in \mathbb{N}$ (such constant exists by the Artin-Rees Lemma). Let $\bar{y} \in A^m, \bar{z} \in A^{qr}$ satisfy $F(\bar{y}, \bar{z}) \in m_A^{\beta(c+c_0)}$. Then $f(\tilde{y}) \in m_A^{\beta(c+c_0)} + I$. Thus there exists $\tilde{y} \in A^m$ such that $f(\tilde{y}) \in I$ and $\tilde{y} - \bar{y} \in m_A^{c+c_0}$. Thus $F(\tilde{y}, \tilde{z}) \in m_A^{c+c_0} \cap I$. Then we conclude by following the proof of Theorem \ref{thm:equicharacteristic}.

**Remark 3.29.** — By Theorem \ref{thm:equicharacteristic} in order to study the behaviour of the Artin function of some ideal we may assume that $A$ is a complete local ring. Let us assume that $A$ is an equicharacteristic local ring. Then $A$ is the quotient of a power series ring over a field by Cohen Structure Theorem \cite{Mat80}. Thus Proposition \ref{prop:equicharacteristic} allows us to reduce the problem to the case $A = k[x_1, \ldots, x_n]$ where $k$ is a field.

### 3.4.2. Izuim Theorem and Diophantine Approximation

Let $(A, m_A)$ be a Noetherian local ring. We denote by $\nu$ the $m_A$-adic order on $A$, i.e.

$$\nu(x) := \max\{n \in \mathbb{N} / x \in m_A^n\} \; \text{ for any } x \neq 0.$$
We always have $\nu(x) + \nu(y) \leq \nu(xy)$ for all $x, y \in A$. But we do not have the equality in general. For instance, if $A := \mathbb{C}[x,y]/(x^2 - y^2)$ then $\nu(x) = \nu(y) = 1$ but $\nu(x^2) = \nu(y^2) = 3$. Nevertheless we have the following theorem:

**Theorem 3.30 (Izumi Theorem).** — [Iz95, Re89] Let $A$ be a local Noetherian ring whose maximal ideal is denoted by $m_A$. Let us assume that $A$ is analytically irreducible, i.e. $\hat{A}$ is irreducible. Then there exist $b \geq 1$, and $d \geq 0$ such that

$$\forall x, y \in A, \quad \nu(xy) \leq b(\nu(x) + \nu(y)) + d.$$  

This result implies easily the following corollary using Corollary 3.28:

**Corollary 3.31.** — [Iz95, Ron06a] Let us consider the polynomial

$$f(y) := y_1y_2 + a_3y_3 + \cdots + a_my_m,$$

with $a_3, \ldots, a_m \in A$ where $(A, m_A)$ is a Noetherian local ring such that $A = (a_3, \ldots, a_m)$ is analytically irreducible. Then there exist $b \geq 1$ and $d \geq 0$ such that the Artin function $\beta$ of Theorem 3.17 satisfies $\beta(c) \leq bc + d$ for all $c \in \mathbb{N}$.

**Proof.** — By Proposition 3.28 we have to prove that the Artin function $\beta$ of $y_1y_2 \in A[y]$ is bounded by an affine function if $A$ is analytically irreducible. Thus let $\bar{y}_1$, $\bar{y}_2 \in A$ satisfy $\bar{y}_1 \bar{y}_2 \in m_A^{bc+d}$ where $b$ and $d$ satisfies Theorem 3.30. This means that

$$2bc + d \leq \nu(\bar{y}_1 \bar{y}_2) \leq b(\nu(\bar{y}_1) + \nu(\bar{y}_2)) + d.$$  

Thus $\nu(\bar{y}_1) \geq c$ or $\nu(\bar{y}_2) \geq c$. In the first case we denote $\tilde{y}_1 = 0$ and $\tilde{y}_2 = \bar{y}_2$ and in the second case we denote $\tilde{y}_1 = \bar{y}_1$ and $\tilde{y}_2 = 0$. Then $\tilde{y}_1 \tilde{y}_2 = 0$ and $\tilde{y}_1 - \bar{y}_1$, $\tilde{y}_2 - \bar{y}_2 \in m_A^c.$

**Hints on the proof of Theorem 3.30 in the complex analytic case**

According to the theory of Rees valuations, there exists discrete valuations $\nu_1, \ldots, \nu_k$ such that $\nu(x) = \min\{\nu_1(x), \ldots, \nu_k(x)\}$ (they are called the Rees valuation of $\nu$). The valuation rings associated to $\nu_1, \ldots, \nu_k$ are the valuation rings associated to the irreducible components of the exceptional divisor of the normalized blowup of $m_A$. Since $\nu_i(xy) = \nu_i(x) + \nu_i(y)$ for any $i$, in order to prove the theorem we have to see that there exists $a \geq 1$ such that $\nu_i(x) \leq a\nu_j(x)$ for any $x \in A$ and any $i$ and $j$. If $A$ is a complex analytic local ring, following S. Izumi proof, we may reduce the problem to the case $\text{dim}(A) = 2$ by using a Bertini type theorem, and then assume that $A$ is normal by using an inequality on the reduced order proved by D. Rees. Then let us consider a resolution of singularities of $\text{Spec}(A)$ (denoted by $\pi$) that factors through the normalized blow-up of $m_A$. In this case, let us denote by $E_1, \ldots, E_s$ the irreducible components of the exceptional divisor of $\pi$. Let us denote $e_{i,j} := E_i \cdot E_j$ for all $1 \leq i, j \leq s$. Let $x$ be an element of $A$. This element defines a germ of analytic hypersurface whose total transform $T_x$ may be written $T_x = S_x + \sum_{j=1}^s m_j E_j$ where $S_x$ is the strict transform of $\{x = 0\}$ and $m_i = \nu_i(x)$, $1 \leq i \leq s$. Then we have

$$0 = T_x \cdot E_i = S_x \cdot E_i + \sum_{j=1}^s m_j e_{i,j}.$$
Since $S_x E_i \geq 0$ for any $i$, the vector $(m_1, ..., m_s)$ is contained in the convex cone $C$ defined by $x_i \geq 0, 1 \leq i \leq s$, and $\sum_{j=1}^{s} e_{i,j} x_j \leq 0, 1 \leq i \leq s$. To prove the theorem, it is enough to prove that $C$ is included in $E_i > 0, 1 \leq i \leq s$. Let assume that it is not the case. Then, after renumbering the $E_i$, we may assume that $(x_1, ..., x_l, 0, ..., 0) \in C$ where $x_1 > 0, 1 \leq i \leq l < s$. Since $e_{i,j} \geq 0$ for all $i \neq j$, $\sum_{j=1}^{s} e_{i,j} x_j = 0$ for $l < i \leq s$ implies that $e_{i,j} = 0$ for all $l < i \leq s$ and $1 \leq j \leq l$. This contradicts the fact that the exceptional divisor of $\pi$ is connected (since $A$ is an integral domain).

Let us mention that Izumi Theorem is the key ingredient to prove the following result:

**Corollary 3.32.** — [Ron06b] [Hic08] [L-108] Let $(A, m_A)$ be a regular Henselian excellent local domain. Let us denote by $K$ and $\hat{K}$ the fraction fields of $A$ and $\hat{A}$ respectively. Let $z \in \hat{K}\backslash K$ be algebraic over $K$. Then

$$\exists a \geq 1, C \geq 0, \forall x \in A \forall y \in A^* \left| z - \frac{x}{y} \right| \geq C|y|^a$$

where $|u| := e^{-\nu(u)}$ and $\nu$ is the usual $m_A$-adic valuation.

This result is equivalent to the following:

**Corollary 3.33.** — [Ron06b] [Hic08] [L-108] Let $(A, m_A)$ be an excellent Henselian local ring and let $f_1(y_1, y_2), ..., f_r(y_1, y_2) \in A[y_1, y_2]$ be homogeneous polynomials. Then the Artin function of $f_1, ..., f_r$ is bounded by an affine function.

3.4.3. Reduction to one quadratic equation and examples. — In general Artin functions are not bounded by affine functions as in Theorem 3.3. Here is such an example:

**Example 3.34.** — [Ron05b] Set $f(y_1, y_2, y_3) := y_1^2 - y_2^2 y_3 \in k[x_1, x_2][y_1, y_2, y_3]$ where $k$ is a field of characteristic zero. Let us denote by $h(T) := \sum_{i=1}^{\infty} a_i T^i \in \mathbb{Q}[T]$ the power series such that $(1 + h(T))^2 = 1 + T$. Let us denote

$$y_1^{(c)} := x_1^{2c+2} \left( 1 + \sum_{i=1}^{c+1} a_i \frac{x_2^i}{x_1^i} \right) = x_1^{2c+2} + \sum_{i=1}^{c+1} a_i x_1^{2(c-i+1)} x_2^i,$$

$$y_2^{(c)} := x_1^{2c+1},$$

$$y_3^{(c)} := x_1^2 + x_2^2.$$

Then in the ring $k(\frac{x_2}{x_1})[x_1]$ we have

$$f(y_1^{(c)}, y_2^{(c)}, y_3^{(c)}) = \left( \frac{y_1^{(c)}}{y_2^{(c)}} \right)^2 - y_3^{(c)} = \left( \frac{y_1^{(c)}}{y_2^{(c)}} \right)^2 - x_1^{2c+2} \left( 1 + \frac{x_2^i}{x_1^i} \right) y_2^{(c)} = \left( \frac{y_1^{(c)}}{y_2^{(c)}} \right) - x_1 \left( 1 + h \left( \frac{x_2}{x_1} \right) \right) \left( \frac{y_1^{(c)}}{y_2^{(c)}} \right) + x_1 \left( 1 + h \left( \frac{x_2}{x_1} \right) \right) y_2^{(c)}.$$
In any case, we have

This allows us to assume that solutions are not too close to the singular locus of $I$. Moreover, the Artin function of $I$ is bounded by the Artin function of $J$.

Lemma 3.36. — [Ron10a, Ron13] Let $k$ be an algebraically closed field of characteristic zero. Let $I$ be an ideal of $k[x_1, x_2][y]$. If $I$ is generated by binomials of $k[y]$ or if $Spec(k[x_1, x_2][y]/I)$ has an isolated singularity, then the Artin function of $I$ is bounded by a function which is doubly exponential, i.e., a function of the form $c \mapsto a^c$ for some constant $a > 1$.

Moreover, the Artin function of $I$ is bounded by an affine function if the approximated solutions are not too close to the singular locus of $I$. We do not know if this doubly exponential bound is sharp since there is no example of Artin function whose growth is greater than a polynomial of degree 2.

In general, in order to investigate bounds on the growth of Artin functions in general, we can reduce the problem as follows, using a trick of [Ron10b]. From now on we assume that $A$ is a complete local ring.

Lemma 3.36. — [Be77b] For any $f(y) \in A[y]^r$ or $A[y]^s$, the Artin function of $f$ is bounded by the Artin function of

$$g(y) := f_1(y)^2 + y_1(f_2(y)^2 + y_1(f_3(y)^2 + \cdots )^2).$$

Proof. — Indeed, if $\beta$ is the Artin function of $g$ and if $f(\tilde{y}) \in m_A^{\beta(c)}$ then $g(\tilde{y}) \in m_A^{\beta(c)}$. Thus there exists $y \in A^n$ such that $g(\tilde{y}) = 0$ and $\tilde{y} - \tilde{y} \in m_A$. But clearly $g(\tilde{y}) = 0$ if and only if $f(\tilde{y}) = 0$. This proves the lemma.

This allows us to assume that $r = 1$ and we define $f(y) := f_1(y)$. If $f(y)$ is not irreducible, then we may write $f = h_1 \cdots h_s$, where $h_i \in A[y]$ is irreducible for $1 \leq i \leq s$, and the Artin function of $f$ is bounded by the sum of the Artin functions of the $h_i$.

Hence we may assume that $f(y)$ is irreducible. We have the following lemma:
Lemma 3.37. — For any \( f(y) \in A[y] \), where \( A \) is a complete local ring, the Artin function of \( f(y) \) is bounded by the Artin function of the polynomial

\[
P(u, x, z) := f(y)u + x_1z_1 + \cdots + x_mz_m \in B[x, z, u]
\]
where \( B := A[y] \).

Proof. — Let us assume that \( f(y) \in \mathfrak{m}_A^\beta(c) \) where \( \beta \) is the Artin function of \( f \). By replacing \( f(y) \) by \( f(y^0 + y) \), where \( y^0 \in A \) is such that \( y^0_i - \overline{y}_i \in \mathfrak{m}_A, 1 \leq i \leq m \), we may assume that \( \overline{y}_i \in \mathfrak{m}_A \) for \( 1 \leq i \leq m \).

Then there exists \( \overline{z}_i(y) \in A[y] \), \( 1 \leq i \leq m \), such that

\[
f(y) + \sum_{i=1}^m (y_i - \overline{y}_i) \overline{z}_i(y) \in (\mathfrak{m}_A + (y))^\beta(c).
\]

Thus there exists \( u(y), f_i(y), z_i(y) \in A[y] \), \( 1 \leq i \leq m \), such that

\[
u(y) - 1, z_i(y) - \overline{z}_i(y), x_i(y) - (y_i - \overline{y}_i) \in (\mathfrak{m}_A + (y))^c, 1 \leq i \leq n
\]

and \( f(y)u(y) + \sum_{i=1}^m x_i(y)z_i(y) = 0 \).

In particular \( u(y) \) is invertible in \( A[y] \) if \( c > 0 \). Let us assume that \( c \geq 2 \). In this case the matrix of the partial derivatives of \( (x_i(y), 1 \leq i \leq m) \) with respect to \( y_1, \ldots, y_m \) has determinant equal to 1 modulo \( \mathfrak{m}_A + (y) \). By the Henselian property there exist \( y_i, c \in \mathfrak{m}_A \) such that \( x_i(y_1, c, \ldots, y_m, c) = 0 \) for \( 1 \leq i \leq m \). Hence, since \( u(y, c) \) is invertible, \( f(y_1, c, \ldots, y_m, c) = 0 \) and \( y_i, c - \overline{y}_i \in \mathfrak{m}_A^c, 1 \leq i \leq m \).

Thus, by Corollary [3.28] in order to study the general growth of Artin functions, it is enough to study the Artin function of the polynomial

\[
y_1y_2 + y_3y_4 + \cdots + y_{2m+1}y_{2m} \in A[y]
\]
where \( A \) is a complete local ring.

4. Examples of Applications

In this part we give some examples of applications of Theorem 2.18 and Corollary 3.17.

Proposition 4.1. — Let \( \hat{A} \) be an excellent Henselian local ring. Then \( \hat{A} \) is reduced (resp. is an integral domain, resp. an integrally closed domain) if and only if \( \hat{A} \) is reduced (resp. is an integral domain, resp. an integrally closed domain).

Proof. — If \( \hat{A} \) is not reduced, then there exists \( \hat{y} \in \hat{A}, \hat{y} \neq 0 \), such that \( \hat{y}^k = 0 \) for some positive integer \( k \). Thus we apply Theorem 2.18 to the polynomial \( y^k \) with \( c \geq \text{ord}(\hat{y}) + 1 \) in order to find \( \overline{y} \in \hat{A} \) such that \( \overline{y}^k = 0 \) and \( \overline{y} \neq 0 \).

In order to prove that \( \hat{A} \) is an integral domain if \( A \) is an integral domain, we apply the same procedure to the polynomial \( y_1y_2 \).

If \( A \) is an integrally closed domain, then \( A \) is an integral domain. Let \( P(z) := \)
Thus Theorem 2.18 for linear equations). We conclude with the help of Proposition 4.2.

Proof. — Let \( \hat{f} \in \hat{A} \) and \( \hat{g} \in \hat{A} \) such that \( \hat{f} \hat{g} = 0 \). By Theorem 2.18 for any \( c \in \mathbb{N} \), there exist \( \hat{a}_c, \hat{g}_c \in A \) such that for some non-negative integers \( \hat{a}_c, \hat{g}_c \), and if there exists a primary decomposition of \( \hat{f} \hat{g} = 0 \), for some integer \( c_0 \), \( \hat{g}_c \neq 0 \). Since \( A \) is an integrally closed domain, then \( \hat{f}_c \in (\hat{g}_c) \) for \( c > c_0 \). Thus \( \hat{f} \in (\hat{g}) + m^c \hat{A} \) for \( c \) large enough. By Nakayama Lemma this implies that \( \hat{f} \in Q \hat{A} \) and \( \hat{A} \) is integrally closed.

\( \square \)

**Proposition 4.2.** — Let \( A \) be an excellent Henselian local ring. Let \( Q \) be a primary ideal of \( A \). Then \( Q \hat{A} \) is a primary ideal of \( \hat{A} \).

Proof. — Let \( \hat{f} \in \hat{A} \) and \( \hat{g} \in \hat{A} \) satisfy \( \hat{f} \hat{g} = 0 \). By Theorem 2.18 for any \( c \in \mathbb{N} \), there exist \( \hat{f}_c, \hat{g}_c \in A \) such that \( \hat{f}_c \hat{g}_c \in Q \hat{f} \). For some \( c \) large enough, \( \hat{g}_c \notin Q \hat{A} \). Since \( A \) is a primary ideal, this proves that \( \hat{f}_c \in Q \hat{A} \) for \( c \) large enough, hence \( \hat{f} \in Q \hat{A} \).

\( \square \)

**Corollary 4.3.** — Let \( A \) be an excellent Henselian local ring. Let \( I = Q_1 \cap \cdots \cap Q_s \) be a primary decomposition of \( I \) in \( A \). Then \( Q_1 \hat{A} \cap \cdots \cap Q_s \hat{A} \) is a primary decomposition of \( I \hat{A} \).

Proof. — Since \( I = \bigcap_{i=1}^s Q_i \), then \( I \hat{A} = \bigcap_{i=1}^s (Q_i \hat{A}) \) by faithful flatness (or by Theorem 2.18 for linear equations). We conclude with the help of Proposition 4.2.

\( \square \)

**Corollary 4.4.** — Let \( A \) be an excellent Henselian local integrally closed domain. If \( \hat{f} \in \hat{A} \) and if there exists \( \hat{g} \in \hat{A} \) such that \( \hat{f} \hat{g} = 0 \), then there exists a unit \( \hat{u} \in \hat{A} \) such that \( \hat{u} \hat{f} \in \hat{A} \).

Proof. — Let \( (\hat{f} \hat{g}) A = Q_1 \cap \cdots \cap Q_s \) be a primary decomposition of the principal ideal of \( A \) generated by \( \hat{f} \hat{g} \). Since \( A \) is an integrally closed domain, it is a Krull ring and \( Q_i = p_i^{(n_i)} \) for some prime ideal \( p_i \), \( 1 \leq i \leq s \), where \( p_i^{(n_i)} \) denote the \( n \)-th symbolic power of \( p \) (see [Mat80] p.88). In fact \( n_i := \nu_{p_i}(\hat{f} \hat{g}) \), where \( \nu_{p_i} \) is the \( p_i \)-adic valuation of the valuation ring \( A_{p_i} \). By Corollary 4.3, \( p_i^{(n_i)} \hat{A} \cap \cdots \cap p_s^{(n_s)} \hat{A} \) is a primary decomposition of \( (\hat{f} \hat{g}) \hat{A} \). Since \( \nu_{p_i} \) are valuations, then

\[
\hat{f} \hat{A} = \left( p_1^{(k_1)} \hat{A} \right) \cap \cdots \cap \left( p_s^{(k_s)} \hat{A} \right) = \left( p_1^{(k_1)} \cap \cdots \cap p_s^{(k_s)} \right) \hat{A}
\]

for some non-negative integers \( k_1, \ldots, k_s \). Let \( h_1, \ldots, h_r \in A \) be generators of the ideal \( p_1^{(k_1)} \cap \cdots \cap p_s^{(k_s)} \). Then \( \hat{f} = \sum_{i=1}^r \hat{a}_i h_i \) and \( h_i = \hat{b}_i \hat{f} \) for some \( \hat{a}_i, \hat{b}_i \in A \), \( 1 \leq i \leq r \).

Thus \( \sum_{i=1}^r \hat{a}_i \hat{b}_i = 1 \), since \( \hat{A} \) is an integral domain. Thus one of the \( \hat{b}_i \) is invertible and we choose \( \hat{u} \) to be this invertible \( \hat{b}_i \).

\( \square \)
Corollary 4.5. — [To72] Let be an excellent Henselian local ring. For \( f(y) \in A[y] \) let \( I \) be the ideal of \( A[y] \) generated by \( f_1(y), \ldots, f_r(y) \). Let us assume that \( ht(I) = m \). Let \( \tilde{y} \in \hat{A}^m \) satisfy \( f(\tilde{y}) = 0 \). Then \( \tilde{y} \in A^m \).

Proof. — Set \( p := (y_1 - \tilde{y}_1, \ldots, y_m - \tilde{y}_m) \). It is a prime ideal of \( \hat{A} \) and \( ht(p) = m \). Of course \( I \hat{A} \subset p \) and \( ht(I \hat{A}) = m \) by Corollary 4.3. Thus \( p \) is of the form \( p' \hat{A} \) where \( p' \) is minimal prime of \( I \). Then \( \tilde{y} \in \hat{A}^m \) is the only common zero of all the elements of \( p' \). By Theorem 2.18 \( \tilde{y} \) can be approximated by a common zero of all the elements of \( p' \) which is in \( A^m \). Thus \( \tilde{y} \in A^m \).

Proposition 4.6. — [KPPRM78, Po86] Let \( A \) be an excellent Henselian local domain. Then \( A \) is a unique factorization domain if and only if \( \hat{A} \) is a unique factorization domain.

Proof. — If \( \hat{A} \) is a unique factorization domain, then any irreducible element of \( \hat{A} \) is prime. Thus any irreducible element of \( A \) is prime. Since \( A \) is a Noetherian integral domain, it is a unique factorization domain.

Let us assume that \( \hat{A} \) is a Noetherian integral domain but not a unique factorization domain. Thus there exists an irreducible element \( \hat{x}_1 \in \hat{A} \) that is not prime. This equivalent to

\[
\exists \hat{x}_2, \hat{x}_3, \hat{x}_4 \in \hat{A} \text{ such that } \hat{x}_1 \hat{x}_2 - \hat{x}_3 \hat{x}_4 = 0
\]

\[
\hat{A}\hat{x}_2, \hat{x}_3 \in \hat{A} \text{ such that } \hat{x}_1 \hat{x}_2 - \hat{x}_3 = 0 \text{ and } \hat{x}_2 \hat{x}_3 - \hat{x}_4 = 0
\]

and \( \hat{A}\hat{y}_1, \hat{y}_2 \in m_A \hat{A} \) such that \( \hat{y}_1 \hat{y}_2 - \hat{x}_1 = 0 \).

Let us denote by \( \beta \) the Artin function of

\[
(\hat{y}, z) := (\hat{x}_1 \hat{z}_1 - \hat{x}_3, \hat{x}_2 \hat{z}_2 - \hat{x}_4, y_1 y_2 - \hat{x}_1) \in \hat{A}[y][z].
\]

Since \( f(y, z) \) has no solution in \( (m_A \hat{A})^2 \times \hat{A}^2 \), by Remark 3.19 \( \beta \) is a constant, and \( f(y, z) \) has no solution in \( (m_A \hat{A})^2 \times \hat{A}^2 \) modulo \( m^3_A \).

On the other hand by Theorem 2.18 applied to \( x_1 x_2 - x_3 x_4 \), there exists \( \hat{x}_1 \in A \), \( 1 \leq i \leq 4 \), such that \( \hat{x}_1 \hat{x}_2 - \hat{x}_3 \hat{x}_4 = 0 \) and \( \hat{x}_1 - \hat{x}_i \in m^{3+1}_A \), \( 1 \leq i \leq 4 \). Hence

\[
g(y, z) := (\hat{x}_1 \hat{z}_1 - \hat{x}_3, \hat{x}_2 \hat{z}_2 - \hat{x}_4, y_1 y_2 - \hat{x}_1) \in \hat{A}[y][z]
\]

has no solution in \( (m_A \hat{A})^2 \times \hat{A}^2 \) modulo \( m^3_A \), hence has no solution in \( (m_A \hat{A})^2 \times \hat{A}^2 \). This means that \( \hat{x}_1 \) is an irreducible element of \( A \) but it is not prime. Hence \( A \) is not a unique factorization domain.

5. Approximation with constraints

We will now discuss the problem of the Artin Approximation with constraints that is the following:

Problem 1 (Artin Approximation with constraints):
Let $A$ be an excellent Henselian local subring of $\mathbb{k}[x_1, \ldots, x_n]$ and $f(y) \in A[y]^r$. Let us assume that we have a formal solution $\hat{y} \in \hat{A}^m$ of $f = 0$ and assume moreover that

$$\hat{y}_i(x) \in \hat{A} \bigcap x_j, j \in J_i$$

for some subset $J_i \subset \{1, \ldots, n\}$, $1 \leq i \leq m$. Is it possible to approximate $\hat{y}(x)$ by a solution $\tilde{y}(x) \in A^m$ of $f = 0$ such that $\tilde{y}_i(x) \in A \bigcap x_j, j \in J_i$, $1 \leq i \leq m$?

Another problem is the following:

**Problem 2** (Strong Artin Approximation with constraints):

Let us consider $f(y) \in \mathbb{k}[x][y]^r$ and $J_i \subset \{1, \ldots, n\}$, $1 \leq i \leq m$. Does there exist a function $\beta : \mathbb{N} \to \mathbb{N}$ such that:

for all $c \in \mathbb{N}$ and all $y_i(x) \in \mathbb{k}[x_j, j \in J_i]$, $1 \leq i \leq m$, such that

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

there exist $\tilde{y}_i(x) \in \mathbb{k}[x_j, j \in J_i]$ such that $f(\tilde{y}(x)) = 0$ and $\tilde{y}_i(x) - y_i(x) \in (x)^c$, $1 \leq i \leq m$?

If such function $\beta$ exists, the smallest function satisfying this property is called the Artin function of the system $f = 0$.

Let us remark that we have already given a positive answer to a similar weaker problem (see Proposition 3.25). The answer will be no in general for both problems and yes for some particular cases. We present here the positive and negative results concerning these problems. We will see that some systems yield a positive answer to Problem 2 but a negative answer to Problem 1.

### 5.1. Examples

First of all we give here a list of examples that show that there is no hope, in general, to have a positive answer to Problem 1 without any more specific hypothesis, even if $A$ is the ring of algebraic or convergent power series. These examples are constructed by looking at the Artin Approximation Problem for equations involving differentials (Examples 5.3 and 5.6) and operators on germs of functions (Examples 5.4 and 5.5). To construct these examples, the following lemma will be used repeatedly:

**Lemma 5.1**. Let $(A, \mathfrak{m}_A)$ be a Noetherian local ring and let $B$ be a Noetherian local subring of $A[y]$ such that $\hat{B} = \hat{A}[y]$. For any $P(y) \in B$ and $\hat{y} \in (\mathfrak{m}_A A)^m$, $P(\hat{y}) = 0$ if and only if there exists $\hat{h}(y) \in B^m$ such that

$$P(y) + \sum_{i=1}^m (y_i - \hat{y}_i)\hat{h}(y) = 0.$$
Proof of Lemma 5.1 — If $P(\hat{y}) = 0$ then, by Taylor expansion, we have:

$$P(y) - P(\hat{y}) = \sum_{\alpha \in \mathbb{N}^m \setminus \{0\}} \frac{1}{\alpha_1! \ldots \alpha_m!} (y_1 - \hat{y}_1)^{\alpha_1} \ldots (y_m - \hat{y}_m)^{\alpha_m} \frac{\partial^\alpha P(\hat{y})}{\partial y^\alpha}.$$ 

Thus there exists $\hat{h}(y) \in A[y]^m$ such that

$$P(y) + \sum_{i=1}^m (y_i - \hat{y}_i)\hat{h}(y) = 0.$$ 

Since $B \rightarrow \hat{B} = \hat{A}[x]$ is faithfully flat and we may assume that $\hat{h}(y) \in B$ (See Example 3).

On the other hand if $P(y) + \sum_{i=1}^m (y_i - \hat{y}_i)\hat{h}(y) = 0$, by substitution of $y_i$ by $\hat{y}_i$, we get $P(\hat{y}) = 0$.

Example 5.2. — Let us consider $P(x, y, z) \in \mathbb{k}[x, y, z]$ where $x$, $y$ and $z$ are single variables and $\hat{y} \in (x, \mathbb{k}[x])$. Then $P(x, \hat{y}, \frac{\partial \hat{y}}{\partial x}) = 0$ if and only if $P(x, \hat{y}, \hat{z}) = 0$ and $\hat{z} - \frac{\partial \hat{y}}{\partial x} = 0$.

Moreover $\hat{z} - \frac{\partial \hat{y}}{\partial x} = 0$ if and only if $\hat{z} - \left( \frac{\hat{y}(x+t) - \hat{y}(x)}{t} \right) \in (t)\mathbb{k}[x, t]$. By Lemma 5.1 this is equivalent to: there exist $\hat{h}(x, t, u)$, $\hat{k}(x, t, u) \in \mathbb{k}[x, t, u]$ such that

$$t\hat{z}(x) - \hat{y}(u) - \hat{g}(x) + t^2\hat{h}(x, t, u) + (u - x - t)\hat{k}(x, t, u) = 0.$$ 

Finally we see that

$$P \left( x, \hat{y}(x), \frac{\partial \hat{y}}{\partial x}(x) \right) = 0 \iff \exists \hat{z}(x) \in \mathbb{k}[x], \hat{h}(x, t, u), \hat{k}(x, t, u), \hat{h}(x, t, u) \in \mathbb{k}[x, t, u], \hat{g}(u) \in \mathbb{k}[u] \text{ s.t.}$$

\[\begin{cases} 
P(x, \hat{y}(x), \hat{z}(x)) = 0 \\
(\hat{z}(x) - \hat{y}(u) - \hat{g}(x) + t^2\hat{h}(x, t, u) + (u - x - t)\hat{k}(x, t, u) = 0 \\
\hat{g}(u) - \hat{y}(x) + (u - y)\hat{h}(x, t, u) = 0 
\end{cases}\]

Lemma 5.1 and Example 5.2 allow us to transform any system of equations involving partial differentials and compositions of power series into a system of algebraic equations whose solutions depend only on some of the $x_i$. Of course there exists plenty of examples of such systems of equations with algebraic or analytic coefficients that do not have algebraic or analytic solutions. These kinds of examples will provide counterexamples to Problem 1 as follows:

Example 5.3. — Let us consider the following differential equation: $y' = y$. The solutions of this equation are the convergent power series $ce^x \in \mathbb{C}[x]$ where $c$ is a complex number.

On the other hand, Example 5.2 $\hat{y}(x)$ is convergent power series solution of this
equation if and only if there exists \( \tilde{g}_1(x_1) \in \mathbb{C}\{x_1\}, \tilde{g}_2(x_2) \in \mathbb{C}\{x_2\} \) and \( \tilde{h}(x_1, x_2, x_3), \tilde{k}(x_1, x_2, x_3) \in \mathbb{C}\{x_1, x_2, x_3\} \) such that \((\text{with } \tilde{g}_1 := \tilde{g})\):
\[
\begin{align*}
\tilde{g}_2(x_2) - \tilde{g}_1(x_1) &= x_3\tilde{g}_1(x_1) + x_2^2\tilde{h}(x_1, x_2, x_3) + (x_2 - x_1 - x_3)\tilde{k}(x_1, x_2, x_3) \\
\tilde{g}_2(x_2) - \tilde{g}_1(x_1) &= (x_2 - x_1)\tilde{h}(x_1, x_2, x_3)
\end{align*}
\]
Thus the former system of equations has a convergent solution
\[
(\tilde{g}_1, \tilde{g}_2, \tilde{h}, \tilde{k}) \in \mathbb{C}\{x_1\} \times \mathbb{C}\{x_2\} \times \mathbb{C}\{x_1, x_2, x_3\},
\]
but no algebraic solution in \( \mathbb{C}(x_1) \times \mathbb{C}(x_2) \times \mathbb{C}(x_1, x_2, x_3) \).

**Example 5.4 (Kashiwara-Gabber Example).** \(\text{[Hir77 p. 75]}\) Let us perform the division of \( xy \) by
\[
g := (x - y^2)(y - x^2) = xy - x^3 - y^3 + x^2y^2
\]
as formal power series in \( \mathbb{C}\{x, y\} \) with respect to the monomial \( xy \) (see Example 5.3 in the introduction). The remainder of this division can be written \( r(x) + s(y) \) where \( r(x) \in (x)\mathbb{C}\{x\} \) and \( s(y) \in (y)\mathbb{C}\{y\} \) since this remainder has no monomial divisible by \( xy \). By symmetry, we get \( r(x) = s(x) \), and by substituting \( y \) by \( x^2 \) we get the following equation:
\[
r(x^2) + r(x) - x^3 = 0.
\]
This relation yields the expansion
\[
r(x) = \sum_{i=0}^{\infty} (-1)^i x^{3^i}
\]
and shows that the remainder of the division is not algebraic. This proves that the equation
\[
xy - gQ(x, y) - R(x) - S(y) = 0
\]
has a convergent solution \((\tilde{q}(x, y), \tilde{r}(x), \tilde{s}(y)) \in \mathbb{C}\{x, y\} \times \mathbb{C}\{x\} \times \mathbb{C}\{y\}\) but has no algebraic solution \((q(x, y), r(x), s(y)) \in \mathbb{C}(x, y) \times \mathbb{C}(x) \times \mathbb{C}(y)\).

**Example 5.5 (Becker Example).** \(\text{[Be77b]}\) By direct computation we show that there exists a unique power series \( f(x) \in \mathbb{C}[x] \) such that \( f(x + x^2) = 2f(x) - x \) and that this power series is not convergent. But, by Lemma 5.1 we have:
\[
f(x + x^2) - 2f(x) + x = 0
\]
\(\Leftrightarrow \exists g(y) \in \mathbb{C}[y], h(x, y), k(x, y) \in \mathbb{C}[x, y] \) s.t.
\[
\left\{
\begin{align*}
F_1 &:= g(y) - 2f(x) + x + (y - x - x^2)h(x, y) = 0 \\
F_2 &:= g(y) - f(x) + (x - y)k(x, y) = 0
\end{align*}
\right.
\]
Then this system of equations has solutions in \( \mathbb{C}(x) \times \mathbb{C}(y) \times \mathbb{C}(x, y)^2 \) but no solution in \( \mathbb{C}(x) \times \mathbb{C}(y) \times \mathbb{C}(x, y)^2, \) even no solution in \( \mathbb{C}\{x\} \times \mathbb{C}\{y\} \times \mathbb{C}\{x, y\}\).
**Example 5.6.** — Set \( \hat{g}(x) := \sum_{i \geq 0} \frac{d}{dx} x^{i+1} \in \mathbb{C}[x] \). This power series is divergent and we have shown in Example 12 that it is the only solution of the equation

\[ x^2 y' - y + x = 0 \] (Euler Equation).

By Example 5.2, \( \hat{g}(x) \) is a solution of this differential equation if and only if there exist \( \hat{y}_2(x_2) \in \mathbb{C}[x_2] \) and \( \hat{k}(x_1, x_2, x_3), \hat{h}(x_1, x_2, x_3), \hat{l}(x_1, x_2, x_3) \in \mathbb{C}[x_1, x_2, x_3] \) such that \( (x := (x_1, x_2, x_3)) \):

\[
\begin{align*}
x_1^2(\hat{y}_2(x_2) - \hat{y}_1(x_1)) - x_3\hat{y}_1(x_1) + x_3x_1 + x_3\hat{k}(x) + (x_1 + x_3 - x_2)\hat{h}(x) &= 0 \\
\hat{y}_2(x_2) - \hat{y}_1(x_1) - (x_2 - x_1)\hat{l}(x) &= 0
\end{align*}
\]

with \( \hat{y}_1(x_1) := \hat{g}(x_1) \). Thus this system has no solution in

\[ \mathbb{C}\{x_1\} \times \mathbb{C}\{x_2\} \times \mathbb{C}\{x_1, x_2, x_3\}^3 \]

but it has solutions in

\[ \mathbb{C}[x_1] \times \mathbb{C}[x_2] \times \mathbb{C}[x_1, x_2, x_3]^3. \]

**Remark 5.7.** — By replacing \( f_1(y), \ldots, f_r(y) \) by

\[ g(y) := f_1(y)^2 + y_1(f_2(y)^2 + y_1(f_3(y)^2 + \ldots)^2)^2 \]

in these examples as in the proof of Lemma 3.36, we can construct the same kind of examples involving only one equation. Indeed \( f_1 = f_2 = \cdots = f_r = 0 \) if and only if \( g = 0 \).

### 5.2. Nested Approximation in the algebraic case.

All the examples of Section 5.1 involve components that depends on separate variables. Indeed, Example 5.2 shows that equations involving partial derivatives yield algebraic equations whose solutions have components with separate variables.

In the case the variables are nested (i.e. \( y_i = y_i(x_1, \ldots, x_{s(i)}) \) for some integer \( i \), which is equivalent to say that \( J_i \) contains or is contained in \( J_j \) for any \( i \) and \( j \) with notations of Problems 1 and 2) it is not possible to construct a counterexample as we did in Section 5.1 from differential equations or equations as in Example 5.5. We will see, in the nested case, that the algebraic case is completely different from the analytic case. First of all in the algebraic case, we have a nested Artin approximation result as follows:

**Theorem 5.8.** — [KPPRM78, Po86] Let \( (A, m_A) \) be an excellent Henselian local ring and \( f(x, y) \in A(x, y)^r \). Let \( \tilde{g}(x) \) be a solution of \( f = 0 \) in \( (m_A + (x))\hat{A}[x]^m \). Let us assume that \( \tilde{y}_i \in \hat{A}[x_1, \ldots, x_s], 1 \leq i \leq m, \) for integers \( s_i, 1 \leq s_i \leq n. \) Then for any \( c \in \mathbb{N} \) there exists a solution \( \tilde{y}(x) \in A(x)^m \) such that for all \( i, \tilde{y}_i(x) \in A(x_1, \ldots, x_{s_i}) \) and \( \tilde{y}(x) - \tilde{g}(x) \in (m_A + (x))^c. \)

This result has a lot of applications and is one of the most important about Artin Approximation. The proof we present here uses the formalism of codes for algebraic power series and is a bit different from the classical one. The key point is the fact
that $A(x)$ satisfies Theorem 2.18 for any excellent Henselian local ring $A$ (see Remark 2.23).

Proof of Theorem 5.8 — Lemma 5.9. — Let $A$ be a complete normal local domain, $u := (u_1, ..., u_n)$, $v := (v_1, ..., v_m)$. Then

$$A[u][v] = \{ f \in A[u, v] / \exists s \in \mathbb{N}, g \in A(v, z_1, ..., z_s), \hat{z}_i \in (m_A + (u))A[u], f = g(v, \hat{z}_1, ..., \hat{z}_s) \}.$$  

Proof of Lemma 5.7. — Let us denote

$$B := \{ f \in A[u, v] / \exists s \in \mathbb{N}, g \in A(v, z_1, ..., z_s), \hat{z}_i \in (m_A + (u))A[u], f = g(v, \hat{z}_1, ..., \hat{z}_s) \}.$$  

Clearly $B$ is a subring of $A[u][v]$.

If $f \in A[u][v]$ we can write $f = f^0 + f^1$ where $f^0 \in A$ and $f^1 \in (m_A + (v))A[v]$. There exist $F_1, ..., F_r \in A[u][v][x_1, ..., x_r]$ such that $\frac{\partial (F_1, ..., F_r)}{\partial (x_1, ..., x_r)}$ is non-zero modulo $m_A + (u, v, x)$ and such that the unique $(f_1, ..., f_r) \in (m_A + (u, v))A[u][v]^r$ with $F(f_1, ..., f_r) = 0$ (by the Implicit Function Theorem) is such that $f^1 = f_1$ (cf. Proposition 5.10). Let us write

$$F_i := \sum_{\alpha, \beta} F_{i, \alpha, \beta} v^\alpha X^\beta, \quad 1 \leq i \leq r$$

with $F_{i, \alpha, \beta} \in A[u]$ for all $i, \alpha, \beta$. We can write $F_{i, \alpha, \beta} = F^0_{i, \alpha, \beta} + z_{i, \alpha, \beta}$ where $F^0_{i, \alpha, \beta} \in A$ and $z_{i, \alpha, \beta} \in (m_A + (u))A[u]$. Let us denote

$$G_i := \sum_{\alpha, \beta} (F^0_{i, \alpha, \beta} + z_{i, \alpha, \beta}) v^\alpha X^\beta, \quad 1 \leq i \leq r$$

where $z_{i, \alpha, \beta}$ are new variables. Let us denote by $z$ the vector whose coordinates are the variables $z_{i, \alpha, \beta}$. Then $\frac{\partial (G_1, ..., G_r)}{\partial (x_1, ..., x_r)} = \frac{\partial (F_1, ..., F_r)}{\partial (x_1, ..., x_r)}$ modulo $m_A + (u, v, z, X)$. Hence, by the Implicit Function Theorem, there exists $h := (h_1, ..., h_r) \in (m_A + (v, z))A(v, z)^r$ such that $G(h) = 0$. Moreover $f^1 = f_1 = h_1(v, z)$, thus we have $f = g(v, z)$ where $g(v, z) := f^0 + h_1(v, z)$. This proves the lemma. 

Then we can prove Theorem 5.8 by induction on $n$. First of all, since $A = \frac{B}{I}$ where $B$ is a complete regular local ring (by Cohen Structure Theorem), by using the same trick as in the proof of Corollary 2.7 we may replace $A$ by $B$ and assume that $A$ is a complete regular local ring. Let us assume that Theorem 5.8 is true for $n - 1$. We denote $x' := (x_1, ..., x_{n-1})$. We will denote by $y_1, ..., y_k$ the unknowns depending only on $x'$ and by $y_{k+1}, ..., y_m$ the unknowns depending on $x_n$. Let us consider the following system of equations

$$(2) \quad f(x', x_n, y_1(x'), ..., y_k(x'), y_{k+1}(x', x_n), y_m(x', x_n)) = 0.$$ 

By Theorem 2.18 and Remark 2.23 we may assume that $\tilde{y}_{k+1}, ..., \tilde{y}_m \in \mathbb{k}[x'][x]$. Thus by Lemma 5.9 we can write $\tilde{y}_i = \sum_{j \in \mathbb{N}} h_{i, j}(z) x_{i, j}^n$ with $\sum_{j \in \mathbb{N}} h_{i, j}(z) x_{i, j}^n \in \mathbb{k}(z, x_n)$ and
\( \tilde{z} = (\tilde{z}_1, \ldots, \tilde{z}_s) \in (x')^k k[x']^s \). We can write
\[
f \left( x', x_n, y_1, \ldots, y_k, \sum_j h_{k+1,j}(z)x_n^j, \ldots, \sum_j h_{m,j}(z)x_n^j \right) = \sum_j G_j(x', y_1, \ldots, y_k, z)x_n^j
\]
where \( G_j(x', y_1, \ldots, y_k, z) \in k[x', y_1, \ldots, y_k, z] \) for all \( j \in \mathbb{N} \). Thus \( \tilde{y}_1, \ldots, \tilde{y}_k, \tilde{z}_1, \ldots, \tilde{z}_s \in k[x'] \) is a solution of the equations \( G_j = 0 \) for all \( j \in \mathbb{N} \). Since \( k(t, y_1, \ldots, y_k, z) \) is Noetherian, this system of equations is equivalent to a finite system \( G_j = 0 \) with \( j \in E \) where \( E \) is a finite subset of \( \mathbb{N} \). Thus by the induction hypothesis applied to the system \( G_j(x', y_1, \ldots, y_k, z) = 0 \), \( j \in E \), there exist \( \tilde{y}_1, \ldots, \tilde{y}_k, \tilde{z}_1, \ldots, \tilde{z}_s \in k(x') \), with nested conditions, such that \( \tilde{y}_1 - \tilde{y}_i, \tilde{z}_1 - \tilde{z}_i \in (x')^r \), for \( 1 \leq i \leq k \) and \( 1 \leq l \leq s \), and \( G_j(x', \tilde{y}_1, \ldots, \tilde{y}_k, \tilde{z}) = 0 \) for all \( j \in E \), thus \( G_j(x', \tilde{y}_1, \ldots, \tilde{y}_k, \tilde{z}) = 0 \) for all \( j \in \mathbb{N} \).

Set \( \tilde{y}_n = \sum_{j \in \mathbb{N}} h_{i,j}(\tilde{z})x_n^j \) for \( k < j \leq m \). Then \( \tilde{y}_1, \ldots, \tilde{y}_m \) satisfy the conclusion of the theorem.

\[\Box\]

Proposition 5.10. — [ArMa65] [AMR92] Let \( A \) be a complete regular local ring and \( v := (v_1, \ldots, v_n) \). If \( f \in (m_A + (v))A(v) \) then there exists an integer \( r \in \mathbb{N} \) and \( F_1, \ldots, F_r \in A[v][X_1, \ldots, X_r] \) such that \( \partial(F_1, \ldots, F_r)/\partial(x_1, \ldots, x_r) \) is non-zero modulo \( m_A + (v, X) \) and such that the unique \( (f_1, \ldots, f_r) \in (m_A + (v))A(v)^r \) with \( F(f_1, \ldots, f_r) = 0 \) (according to the Implicit Function Theorem) is such that \( f = f_1 \).

Proof. — Let \( P(v, X_1) \in A[v][X_1] \) be an irreducible polynomial such that \( P(v, f) = 0 \). Set \( R := A[v][X_1]/(P(v, X_1)) \) and let \( \overline{R} \) be its normalization. Let \( \varphi : R \rightarrow A(v) \) be the \( A[v] \)-morphism defined by \( \varphi(X_1) = f \). Since \( A(v) \) is the Henselization of a regular local ring (by Example 2.22) it is also a regular local ring thus a normal local ring hence, by the universal property of the normalization, the morphism \( \varphi \) factors through \( R \rightarrow \overline{R} \). Let \( \overline{\varphi} : \overline{R} \rightarrow A(v) \) be the extension of \( \varphi \) to \( \overline{R} \).

Since \( R \) is finitely generated over a local complete domain \( A \), then \( \overline{R} \) is module-finite over \( R \). Hence \( \overline{R} = A[v_1, v_2, \ldots, v_n][F_1, \ldots, F_r]/(F_1, \ldots, F_r) \). Let \( f_i := \overline{\varphi}(X_i) \), for \( 2 \leq i \leq r \). By replacing \( X_1 \) by \( X_1 + a_i \) for some \( a_i \in A \) we may assume that \( f_1 \in m_A + (v) \). Let us denote \( B := \overline{R}_{m_A + (v)}[X_1, \ldots, X_r] \). Thus \( \overline{\varphi} \) induces a \( A[v] \)-morphism \( B \rightarrow A(v) \) whose image contains \( A[v] \). Thus by the universal property of the Henselization it induces a surjective \( A[v] \)-morphism \( B^h \rightarrow A(v) \) where \( B^h \) denotes the Henselization of \( B \). Moreover \( A[v]_{m_A + (v)} \rightarrow B \) induces a morphism between \( A(v) \) and \( B^h \) which is finite since \( A[v] \rightarrow \overline{R} \) is finite. Since \( B \) is an integrally closed local domain then its completion is a local domain [Za48], hence \( B^h \) is a local domain. If \( b \in B^h \) is in the kernel of \( B^h \rightarrow A(v) \), since \( b \) is finite over \( A(v) \), then \( b \) would satisfy \( b^k = 0 \) for some positive integer \( k \). But \( B^h \) being a domain, then \( b \) has to be zero. Thus \( B^h \rightarrow A(v) \) is injective hence \( B^h \) and \( A(v) \) are isomorphic. Moreover we have \( B^h \simeq B \otimes_{A[v]_{m_A + (v)}} A(v) \). Using the definition of an étale morphism, since \( A[v]_{m_A + (v)} \rightarrow A(v) \) is faithfully flat, it is an exercice to check that \( A[v]_{m_A + (v)} \rightarrow B \) is étale. Thus \( s = r \) and \( \partial(F_1, \ldots, F_r)/\partial(x_1, \ldots, x_r) \) is non-zero modulo \( m_A + (v, X) \) and the unique solution of \( F = 0 \) in \( (m_A + (v))A(v)^r \) is \( (f, f_2, \ldots, f_r) \). \[\Box\]
Using ultraproducts methods we can deduce the following Strong Approximation result:

**Corollary 5.11.** [BDLvdD79] Let \( k \) be a field and \( f(x, y) \in k(x, y)^r \). There exists \( \beta : \mathbb{N} \rightarrow \mathbb{N} \) satisfying the following:

Let \( c \in \mathbb{N} \) and \( \bar{y}(x) \in ((k[x])^m)^n \) satisfy \( f(x, \bar{y}(x)) \in (x)^{\beta(c)} \). Let us assume that \( \bar{y}_i(x) \in k[[x_1, \ldots, x_n]] \), \( 1 \leq i \leq m \), for integers \( s_i, 1 \leq s_i \leq n \).

Then there exists a solution \( \tilde{y}(x) \in ((k[x])^m)^n \) such that \( \bar{y}_i(x) \in k[x_1, \ldots, x_n] \) and \( \bar{y}(x) - \tilde{y}(x) \in (x)^{\beta} \).

### 5.3. Nested Approximation in the analytic case.

In the analytic case, Theorem 5.8 is no more valid, as shown in the following example:

**Example 5.12 (Gabrielov Example).** [Ga71] Let \( \varphi : \mathbb{C}\{x_1, x_2, x_3\} \rightarrow \mathbb{C}\{y_1, y_2\} \) be the morphism of analytic \( \mathbb{C} \)-algebras defined by

\[
\varphi(x_1) = y_1, \quad \varphi(x_2) = y_1y_2, \quad \varphi(x_3) = y_1e^{y_2}.
\]

Let \( f \in \text{Ker}(\bar{\varphi}) \) be written as \( f = \sum_{d=0}^{+\infty} f_d \) where \( f_d \) is a homogeneous polynomial of degree \( d \) for all \( d \in \mathbb{N} \). Then \( 0 = \bar{\varphi}(f) = \sum_0 y_1^d f_d(1, y_2, y_2e^{y_2}) \). Thus \( f_d = 0 \) for all \( d \in \mathbb{N} \) since \( 1, y_2 \) and \( y_2e^{y_2} \) are algebraically independent over \( \mathbb{C} \). Hence \( \text{Ker}(\bar{\varphi}) = \{0\} \) and \( \text{Ker}(\varphi) = \{0\} \). This remark is due to W. S. Osgood [Os16].

- We may remark that "\( \varphi(x_3 - x_2e^{x_2}) = 0^n \). But \( x_3 - x_2e^{x_2} \) is not an element of \( \mathbb{C}\{x_1, x_2, x_3\} \).

Let us denote

\[
f_n := \left( x_3 - x_2 \sum_{i=0}^{\infty} \frac{1}{i!} \frac{x_2^i}{x_1^i} \right) x_1^n \in \mathbb{C}\{x_1, x_2, x_3\}, \quad \forall n \in \mathbb{N}.
\]

Then

\[
\varphi(f_n) = y_1^{n+1}y_2 \sum_{i=n+1}^{+\infty} \frac{y_2^i}{i!} \quad \forall n \in \mathbb{N}.
\]

Then we see that \( (n+1)!\varphi(f_n) \) is a convergent power series whose coefficients have module less than 1. Moreover if the coefficient of \( y_k^iy_j^l \) in the Taylor expansion of \( \varphi(f_n) \) is non zero then \( k = n + 1 \). Thus \( h := \sum_n (n+1)!\varphi(f_n) \) is a convergent power series since each of its coefficients has module less than 1. But \( \bar{\varphi} \) being injective, the unique element whose image is \( h \) is necessarily \( \bar{\varphi}(\sum_n (n+1)!f_n) \).

But

\[
\bar{\varphi}(\sum_n (n+1)!f_n) = \left( \sum_n (n+1)!x_1^n \right) x_3 + \tilde{f}(x_1, x_2)
\]

and \( \sum_n (n+1)!x_1^n \) is a divergent power series and \( \bar{\varphi}(\tilde{g}(x)) = h(y) \in \mathbb{C}\{y\} \).

Hence \( \varphi(\mathbb{C}\{x\}) \subseteq \bar{\varphi}(\mathbb{C}[x]) \cap \mathbb{C}\{y\} \).
• By Lemma [5.1] \( \hat{\phi}(\hat{g}(x)) = h(y) \) is equivalent to say that there exist \( \hat{k}_1(x, y), \hat{k}_2(x, y), \hat{k}_3(x, y) \in \mathbb{C}[x, y] \) such that

\[
(3) \quad \hat{g}(x) + (x_1 - y_1)\hat{k}_1(x, y) + (x_2 - y_1y_2)\hat{k}_2(x, y) + (x_3 - y_1y_2)\hat{k}_3(x, y) - h(y) = 0.
\]

Since \( \hat{g}(x) \) is the unique element whose image under \( \hat{\phi} \) equals \( h(y) \), Equation (3) has no convergent solution \( g(x) \in \mathbb{C}[x] \), \( k_1(x, y) \), \( k_2(x, y) \), \( k_3(x, y) \in \mathbb{C}[x, y] \). Thus Theorem [5.8] is not true in the analytic setting.

Let us denote \( \hat{g}_1(x_1, x_2) := \sum_n (n + 1)!x_1^n \) and \( \hat{g}_2(x_1, x_2) := \hat{f}(x_1, x_2) \). By replacing \( y_1 \) by \( x_1 \), \( y_2 \) by \( y \) and \( x_3 \) by \( x_1 \) in Equation (3) we see that the equation

\[
(4) \quad \hat{g}_1(x_1, x_2)x_1e^y + \hat{g}_2(x_1, x_2) + (x_2 - x_1y)\hat{k}(x, y) - h(x_1, y) = 0.
\]

has a nested formal solution but no nested convergent solution.

Nevertheless there are, at least, three positive results about the nested approximation problem in the analytic category. They are the followings.

5.3.1. Grauert Theorem. — The first one is due to H. Grauert who proved it in order to construct analytic deformations of a complex analytic germ in the case it has an isolated singularity. The approximation result of H. Grauert may be reformulated as: "a system of complex analytic equations, considered as a formal nested system, admits an Artin function (as in Problem 2) which is the Identity function, then it has nested analytic solutions". We present here the result.

Set \( x := (x_1, ..., x_n) \), \( t := (t_1, ..., t_l) \), \( y = (y_1, ..., y_m) \) and \( z := (z_1, ..., z_p) \). Let \( f := (f_1, ..., f_r) \) be in \( \mathbb{C}[t, x, y, z]^r \). Let \( I \) be an ideal of \( \mathbb{C}[t] \).

**Theorem 5.13.** — [Gra72] Let \( d_0 \in \mathbb{N} \) and \( \overline{\mathfrak{g}}(t), \overline{\mathfrak{z}}(t, x) \in \mathbb{C}[t][m] \times \mathbb{C}[x][t]^p \) satisfy

\[
f(t, x, \overline{\mathfrak{g}}(t), \overline{\mathfrak{z}}(t, x)) \in I + (t)^{d_0}.
\]

Let us assume that for any \( d \geq d_0 \) and for any \( (y^{(d)}(t), z^{(d)}(t, x)) \in \mathbb{C}[t][m] \times \mathbb{C}[x][t]^p \) such that \( \overline{\mathfrak{g}}(t) - y^{(d)}(t) \in (t)^{d_0} \) et \( \overline{\mathfrak{z}}(t, x) - z^{(d)}(t, x) \in (t)^{d_0} \), and such that

\[
f(t, x, y^{(d)}(t), z^{(d)}(t, x)) \in I + (t)^d,
\]

there exists \( (\varepsilon(t), \eta(t, x)) \in \mathbb{C}[t][m] \times \mathbb{C}[x][t]^p \) homogeneous in \( t \) of degree \( d \) such that

\[
f(t, x, y^{(d)}(t) + \varepsilon(t), z^{(d)}(t, x) + \eta(t, x)) \in I + (t)^{d+1}.
\]

Then there exists \( \tilde{\mathfrak{g}}(t), \tilde{\mathfrak{z}}(t, x) \in \mathbb{C}[t][m] \times \mathbb{C}[x][t]^p \) such that

\[
f(t, x, \tilde{\mathfrak{g}}(t), \tilde{\mathfrak{z}}(t, x)) \in I \quad \text{and} \quad \tilde{\mathfrak{g}}(t) - \overline{\mathfrak{g}}(t), \tilde{\mathfrak{z}}(t, x) - \overline{\mathfrak{z}}(t, x) \in (t)^{d_0}.
\]

The main ingredient of the proof is a result of Functional Analysis called "voisinages privilégiés" and proven by H. Cartan ([Ca44, Théorème α]). We do not give details here, but the reader may consult [LJP100].
5.3.2. Gabrielov Theorem. — The second positive result about the nested approximation problem in the analytic category is due to A. Gabrielov. Before giving his result, let us explain the context.

Let \( \varphi : A \to B \) be a morphism of analytic algebras where \( A := \mathbb{C}[x_1, \ldots, x_n] / I \) and \( B := \mathbb{C}[y_1, \ldots, y_m] / J \) are analytic algebras. Let us denote \( \varphi_i := \varphi(x_i) \) for \( 1 \leq i \leq n \). Let us denote by \( \hat{\varphi} : \hat{A} \to \hat{B} \) the morphism induced by \( \varphi \). A. Grothendieck [Gro60] and S. S. Abhyankar [Ar71] raised the following question: Does \( \ker(\hat{\varphi}) = \ker(\varphi) \hat{A} \)? Without loss of generality, we may assume that \( A \) and \( B \) are regular, i.e., \( A = \mathbb{C}[x_1, \ldots, x_n] \) and \( B = \mathbb{C}[y_1, \ldots, y_m] \).

In this case, an element of \( \ker(\varphi) \) (resp. of \( \ker(\hat{\varphi}) \)) is called an analytic (resp. formal) relation between \( \varphi_1(y), \ldots, \varphi_m(y) \). Hence the previous question is equivalent to the following: is any formal relation \( \hat{S} \) between \( \varphi_1(y), \ldots, \varphi_n(y) \) a linear combination of analytic relations?

This question is also equivalent to the following: may any formal relation between \( \varphi_1(y), \ldots, \varphi_n(y) \) be approximated by analytic relations for the \( (x) \)-adic topology? In this form the problem is the "dual" problem to the Artin Approximation Problem. In fact this problem is also a nested approximation problem. Indeed let \( \hat{S} \) be a formal relation between \( \varphi_1(y), \ldots, \varphi_n(y) \). This means that \( \hat{S}(\varphi_1(y), \ldots, \varphi_n(y)) = 0 \). By Lemma 5.2, this is equivalent to the existence of \( \hat{h}_1(x, y), \ldots, \hat{h}_n(x, y) \in \mathbb{C}[x, y] \) such that

\[
\hat{S}(x_1, \ldots, x_n) = \sum_{i=1}^n (x_i - \varphi_i(y))\hat{h}_i(x, y) = 0.
\]

If this equation has an analytic nested solution \( S(x) \in \mathbb{C}\{x\} \), \( h_1(x, y), \ldots, h_n(x, y) \in \mathbb{C}[x, y] \), it gives an analytic relation between \( \varphi_1(y), \ldots, \varphi_n(y) \).

Example 5.14. — [Ga71] Let us consider now the morphism

\[
\psi : \mathbb{C}\{x_1, x_2, x_3, x_4\} \to \mathbb{C}\{y_1, y_2\}
\]

defined by

\[
\psi(x_1) = y_1, \quad \psi(x_2) = y_1y_2, \quad \psi(x_3) = y_1y_2e^{y_2}, \quad \psi(x_4) = h(y_1, y_2).
\]

Then \( x_4 - \hat{g}(x_1, x_2, x_3) \in \ker(\hat{\psi}) \). On the other hand, the morphism induced by \( \hat{\psi} \) on \( \mathbb{C}\{x_1, \ldots, x_4\} / (x_4 - \hat{g}(x_1, x_2, x_3)) \) is isomorphic to \( \hat{\varphi} \) (where \( \varphi \) is the morphism of Example 5.12) that is injective. Thus we have \( \ker(\hat{\psi}) = (x_4 - \hat{g}(x_1, x_2, x_3)) \).

Since \( \ker(\varphi) \) is a prime ideal of \( \mathbb{C}\{x\} \), \( \ker(\psi) \mathbb{C}[x] \) is a prime ideal of \( \mathbb{C}[x] \) included in \( \ker(\hat{\psi}) \) by Proposition 5.1. Let us assume that \( \ker(\psi) \neq (0) \), then \( \ker(\psi) \mathbb{C}[x] = \ker(\hat{\psi}) \) since \( \text{ht}(\ker(\hat{\psi})) = 1 \). Thus \( \ker(\hat{\psi}) \) is generated by one convergent power series denoted by \( f \in \mathbb{C}[x_1, \ldots, x_4] \) (in unique factorization domains, prime ideals of height one are principal ideals). Since \( \ker(\varphi) = (x_4 - \hat{g}(x_1, x_2, x_3)) \), there exists \( u(x) \in \mathbb{C}[x] \), \( u(0) \neq 0 \), such that \( f = u(x)(x_4 - \hat{g}(x_1, x_2, x_3)) \). By applying Weierstrass Preparation Theorem to \( f \) with respect to \( x_4 \) we see that \( u(x) \) and \( x_4 - \hat{g}(x_1, x_2, x_3) \) must be convergent, which is impossible since \( \hat{g} \) is a divergent power series. Hence \( \ker(\psi) = (0) \) but \( \ker(\psi) \neq (0) \).
Nevertheless A. Gabrielov proved the following theorem:

**Theorem 5.15.** — [Ga73] Let \( \varphi : A \to B \) be a morphism of complex analytic algebras. Let us assume that the generic rank of the Jacobian matrix is equal to \( \dim(\tilde{\ker}(\varphi)) \). Then \( \ker(\tilde{\varphi}) = \ker(\varphi)\tilde{\tilde{A}} \).

**Sketch of the proof.** — We give a sketch of the proof given by J.-Cl. Tougeron [To90]. As before we may assume that \( \tilde{A} = \mathbb{C}\{x_1, \ldots, x_n\} \) and \( \tilde{B} = \mathbb{C}\{y_1, \ldots, y_m\} \). Let us assume that \( \tilde{\ker}(\varphi)\tilde{\tilde{A}} \not\subseteq \ker(\tilde{\varphi}) \). Using a Bertini type theorem we may assume that \( n = 3 \), \( \varphi \) is injective and \( \dim(\tilde{\ker}(\varphi)) = 2 \) (in particular \( \ker(\tilde{\varphi}) \) is a principal ideal). Moreover, in this case we may assume that \( m = 2 \). After a linear change of coordinates we may assume that \( \ker(\tilde{\varphi}) \) is generated by an irreducible Weierstrass polynomial of degree \( d \) in \( x_3 \). Using change of coordinates and quadratic transforms on \( \mathbb{C}\{y_1, y_2\} \) and using changes of coordinates of \( \mathbb{C}\{x\} \) involving only \( x_1 \) and \( x_2 \), we may assume that \( \varphi_1 = y_1 \) and \( \varphi_2 = y_1 y_2 \). Let us denote \( f(y) := \varphi_3(y) \). Then we have

\[
    f(y)^d + \hat{a}_1(y_1, y_1 y_2) f(y)^{d-1} + \cdots + \hat{a}_d(y_1, y_1 y_2) = 0
\]

for some \( \hat{a}_i(x) \in \mathbb{C}[x_1, x_2] \), \( 1 \leq i \leq d \). Then we want to prove that the \( \hat{a}_i \) may be chosen convergent in order to get a contradiction. Let us denote

\[
    P(Z) := Z^d + \hat{a}_1(x_1, x_2) Z^{d-1} + \cdots + \hat{a}_d(x_1, x_2) \in \mathbb{C}[[x]] [Z].
\]

Since \( \ker(\tilde{\varphi}) \) is prime we may assume that \( P(Z) \) is irreducible. J.-Cl. Tougeron studies the algebraic closure \( \tilde{\mathbb{K}} \) of the field \( \mathbb{C}\{(x_1, x_2)\} \). Let consider the following valuation ring

\[
    V := \left\{ f | g, f, g \in \mathbb{C}[x_1, x_2], g \neq 0, \text{ord}(f) \geq \text{ord}(g) \right\},
\]

let \( \tilde{V} \) be its completion and \( \tilde{\mathbb{K}} \) the fraction field of \( \tilde{V} \). J.-Cl. Tougeron proves that the algebraic extension \( \mathbb{K} \to \tilde{\mathbb{K}} \) splits into \( \mathbb{K} \to \mathbb{K}_1 \to \tilde{\mathbb{K}} \) where \( \mathbb{K}_1 \) is a subsfield of the following field

\[
    \mathbb{L} := \left\{ A \in \tilde{\mathbb{K}} / \exists \delta, a_i \in \mathbb{K}[x] \text{ is homogeneous} \forall i, \text{ord} \left( \frac{a_i}{\delta^{m(i)}} \right) = i, \exists a, b \text{ such that } m(i) \leq ai + b \forall i \text{ and } A = \sum_{i=0}^{\infty} \frac{a_i}{\delta^{m(i)}} \right\}.
\]

Moreover the algebraic extension \( \mathbb{K}_1 \to \tilde{\mathbb{K}} \) is the extension of \( \mathbb{K}_1 \) generated by all the roots of polynomials of the form \( Z^q + g_1(x) Z^{q-1} + \cdots + g_q(x) \) where \( g_i \in \mathbb{C}(x) \) are homogeneous rational fractions of degree \( e_i \), \( 1 \leq i \leq q, e \in \mathbb{Q} \). A root of such polynomial is called a homogeneous element of degree \( e \). For example, square roots of \( x_1 \) or of \( x_1 + x_2 \) are homogeneous elements of degree 2. We have \( \tilde{\mathbb{K}} \cap \mathbb{L} = \mathbb{K}_1 \).

In the same way he proves that the algebraic closure \( \mathbb{K}^{an} \) of \( \mathbb{K}^{an} \), the fraction field of \( \mathbb{C}\{(x_1, x_2)\} \) can be factorized as \( \mathbb{K}^{an} \to \mathbb{K}_1^{an} \to \tilde{\mathbb{K}}^{an} \) with \( \mathbb{K}_1^{an} \subset \mathbb{L}^{an} \) where

\[
    \mathbb{L}^{an} := \left\{ A \in \tilde{\mathbb{K}} / \exists \delta, a_i \in \mathbb{K}[x] \text{ is homogeneous} \forall i, \text{ord} \left( \frac{a_i}{\delta^{m(i)}} \right) = i, A = \sum_{i=0}^{\infty} \frac{a_i}{\delta^{m(i)}} \right\}.
\]
\[ \exists a, b \text{ such that } m(i) \leq ai + b \forall i \quad \text{and} \quad \exists r > 0 \text{ such that } \sum_i |a_i|^r < \infty \]

and \[ ||a(x)|| := \max_{|i| \leq 1} |a(z_1, z_2)| \] for a homogeneous polynomial \[ a \in \mathbb{K}[x]. \]

Clearly, \( \xi := f(x_1, \frac{x_2}{x_1}) \) is an element of \( \mathbb{K} \) since it is a root of \( P(Z) \). Moreover \( \xi \) may be written \( \xi = \sum_{i=1}^d \xi_i \gamma^i \) where \( \gamma \) is a homogenous element and \( \xi_i \in \mathbb{L}^{an} \cap \mathbb{K} \) for any \( i \), i.e. \( \xi \in \mathbb{L}^{an}[\gamma] \). Thus the problem is to show that \( \xi_i \in \mathbb{K}_1^{an} \) for any \( i \), i.e. \( \mathbb{L}^{an} \cap \mathbb{K} = \mathbb{K}_1^{an} \).

Then the idea is to resolve, by a sequence of blowing-ups, the singularities of the discriminant locus of \( P(Z) \) which is a germ of plane curve. Let us call \( \pi \) this resolution map. Then the discriminant of \( \pi^*(P)(Z) \) is normal crossing and \( \pi^*(P)(Z) \) defines a germ of hypersurface along the exceptional divisor of \( \pi \), denoted by \( E \). Let \( p \) be a point of \( E \). At this point \( \pi^*(P)(Z) \) may factor as a product of polynomials and \( \xi \) is a root of one of these factors denoted by \( Q_1(Z) \) and this root is a germ of an analytic function at \( p \). Then the other roots of \( Q_1(Z) \) are also in \( \mathbb{L}^{an}[\gamma] \) according to the Abhyankar-Jung Theorem, for some homogeneous element \( \gamma' \). Thus the coefficients of \( Q_1(Z) \) are in \( \mathbb{L}^{an} \) and are analytic at \( p \).

Then the idea is to use the special form of the elements of \( \mathbb{L}^{an} \) to prove that the coefficients of \( Q_1(Z) \) may be extended as analytic functions along the exceptional divisor \( E \) (the main ingredient in this part is the Maximum Principle). We can repeat the latter procedure in another point \( p' \): we take the roots of \( Q_1(Z) \) at \( p' \) and using Abhyankar-Jung Theorem we construct new roots of \( \pi^*(P)(Z) \) at \( p' \) and the coefficients of \( Q_2(Z) := \prod (Z - \sigma_i) \), where \( \sigma_i \) runs over all these roots, are in \( \mathbb{L}^{an} \) and are analytic at \( p' \). Then we extend the coefficients of \( Q_2(Z) \) everywhere along \( E \). Since \( \pi^*(P)(Z) \) has exactly \( d \) roots, this process stops after a finite number of steps. The polynomial \( Q(Z) := \prod (Z - \sigma_k) \), where the \( \sigma_k \) are the roots of \( \pi^*(P)(Z) \) that we have constructed, is a polynomial whose coefficients are analytic everywhere and it divides \( \pi^*(P)(Z) \). Thus, by Grauert Direct Image Theorem, there exists \( R(Z) \in \mathbb{C}(x)[Z] \) such that \( \pi^*(R)(Z) = Q(Z) \). Thus \( R(Z) \) divides \( P(Z) \), but since \( P(Z) \) is irreducible, then \( P(Z) = R(Z) \in \mathbb{C}(x)[Z] \) and the result is proven.

5.3.3. One variable Nested Approximation. — In the example of A. Gabrielov we can remark that the nested part of the solutions depends on two variables \( x_1 \) and \( x_2 \). In the case they depend only on one variable the nested approximation property is true. This is the following theorem:

**Theorem 5.16.** (cf. Theorem 5.1 [DeLi80]) Let \( k \) be a field and \( k[[x]] \) be a \( W \)-system over \( k \). Let \( t \) be one variable, \( x = (x_1, \ldots, x_n) \), \( y = (y_1, \ldots, y_m, y_{m+1}, \ldots, y_{m+k}) \), \( f \in k[[t, x, y]] \). Let \( \tilde{y}_1, \ldots, \tilde{y}_m \in (t)k[[t]] \) and \( \tilde{y}_{m+1}, \ldots, \tilde{y}_{m+k} \in (t, x)k[[t, x]] \) satisfy \( f(t, x, \tilde{y}) = 0 \). Then, for any \( c \in \mathbb{N} \), there exists \( \tilde{y}_1, \ldots, \tilde{y}_m \in (t)k[[t]], \; y_{m+1}, \ldots, \; y_{m+k} \in (t, x)k[[t, x]] \) such that \( f(t, x, \tilde{y}) = 0 \) and \( \tilde{y} = \tilde{y} \in (t, x)^c \).

**Example 5.17.** — The main example is the case where \( k \) is a valued field and \( k[[x]] \) is the ring of convergent power series over \( k \).
Proof. — The proof is very similar to the second proof of Theorem 5.8. Set \( u := (u_1, ..., u_j) \), \( j \in \mathbb{N} \) and Set
\[
\mathbb{k}[t][[u]] := \left\{ f(z_1(t), ..., z_s(t), u) \in \mathbb{k}[t, u] \right. \]
\[
\left. f(z_1, ..., z_s, u) \in \mathbb{k}[z, u] \right\}
\]
The rings \( \mathbb{k}[t][[u]] \) form a \( \mathbb{W} \)-system over \( \mathbb{k}[t] \) (cf. Lemma 52, [DeLi80] but it is straightforward to check it since \( \mathbb{k}[x] \) is a \( \mathbb{W} \)-system over \( \mathbb{k} \) - in particular, if \( \text{char}(\mathbb{k}) > 0 \), vi) of Definition 2.12 is satisfied since v) of Definition 2.12 is satisfied for \( \mathbb{k}[x] \)). By Theorem 2.15 applied to
\[
f(t, \hat{y}_1, ..., \hat{y}_m, y_{m+1}, ..., y_{m+k}) = 0
\]
there exist \( \gamma_{m+1}, ..., \gamma_{m+k} \in \mathbb{k}[[t]] \) such that \( f(t, \hat{y}_1, ..., \hat{y}_m, \gamma_{m+1}, ..., \gamma_{m+k}) = 0 \) and \( \hat{y}_i - \hat{y}_i \in \mathbb{k}(t, x)^c \) for \( m < i \leq m + k \).

Let us write \( \gamma_i = \sum_{\alpha \in \mathbb{N}^n} h_{i, \alpha}(\hat{x})x^\alpha \) with \( \sum_{\alpha \in \mathbb{N}^n} h_{i, \alpha}(\hat{x})x^\alpha \in \mathbb{k}[[x, t]] \) and \( \hat{x} = (\hat{z}_1, ..., \hat{z}_s) \in \mathbb{k}[t] \). We can write
\[
f(t, x, y_1, ..., y_m, \sum_{\alpha} h_{m+1, \alpha}(x)z^\alpha, ..., \sum_{\alpha} h_{m+k, \alpha}(z)z^\alpha) = \sum_{\alpha} G_{\alpha}(t, y_1, ..., y_m, z)z^\alpha
\]
where \( G_{\alpha}(t, y_1, ..., y_m, z) \in \mathbb{k}[t, y_1, ..., y_m, z] \) for all \( \alpha \in \mathbb{N}^n \). Thus \( \hat{y}_1, ..., \hat{y}_m, \hat{z}_1, ..., \hat{z}_s \) is a solution of the equations \( G_{\alpha} = 0 \) for all \( \alpha \in \mathbb{N}^n \). Since \( \mathbb{k}[[t, y_1, ..., y_m, z]] \) is Noetherian, this system of equations is equivalent to a finite system \( \mathbb{G}_{\alpha} = 0 \) with \( \alpha \in \mathbb{E} \) where \( \mathbb{E} \) is a finite subset of \( \mathbb{N}^n \). Thus by Theorem 2.15 applied to the system
\[
G_{\alpha}(t, y_1, ..., y_m, z) = 0, \alpha \in \mathbb{E},
\]
there exist \( \gamma_1, ..., \gamma_m, \gamma_{m+1}, ..., \gamma_{m+k} \in \mathbb{k}[t] \) such that \( \gamma_i - \gamma_i \in \mathbb{k}(t, x)^c \) for \( 1 \leq i \leq m \) and \( 1 \leq j \leq s \), and \( G_{\alpha}(t, \gamma_1, ..., \gamma_m, \gamma_z) = 0 \) for all \( \alpha \in \mathbb{E} \), thus \( G_{\alpha}(t, \gamma_1, ..., \gamma_m, \gamma_z) = 0 \) for all \( \alpha \in \mathbb{N}^n \).

Set \( \gamma_i = \sum_{\alpha \in \mathbb{N}^n} h_{i, \alpha}(\hat{x})x^\alpha \) for \( m < i \leq m + k \). Then \( \gamma_1, ..., \gamma_{m+k} \) satisfy the conclusion of the theorem.

\[ \square \]

Remark 5.18. — The proof of this theorem uses in an essential way the Weierstrass Division Property (in order to show that \( \mathbb{k}[t][[u]] \) is a Noetherian local ring, which is the main condition to use Theorem 2.15. The Henselian and excellent conditions may be proven quite easily). On the other hand the Weierstrass Division Property (at least in dimension 2) is necessary to obtain this theorem. Indeed if \( \mathbb{k}[x] \) is a family of rings satisfying Theorem 5.16 and \( f(t, y) \in \mathbb{k}[t, y] \) is \( y \)-regular of order \( d \) and \( g(t, y) \in \mathbb{k}[t, y] \) is another series, we can write formally in unique way
\[
g(t, y) = \tilde{g}(t, y)f(t, y) + \tilde{r}_0(t) + \tilde{r}_1(t)y + \cdots + \tilde{r}_{d-1}(t)y^{d-1}
\]
where \( \tilde{g}(t, y) \in \mathbb{k}[t, y] \) and \( \tilde{r}_i(t) \in \mathbb{k}[t] \) for all \( i \). Thus by Theorem 5.16 \( \tilde{g}(t, y) \in \mathbb{k}[t, y] \) and \( \tilde{r}_i(t) \in \mathbb{k}[t] \) for all \( i \). This means that \( \mathbb{k}[x, t] \) satisfies the Weierstrass Division Theorem.

For example let \( C_n \) be the ring of germs of \( \mathbb{k} \)-valued Denjoy-Carleman functions defined at the origin of \( \mathbb{R}^n \), where \( \mathbb{k} = \mathbb{R} \) or \( \mathbb{C} \) (see [Th08] for definitions and
properties of these rings. It is still an open problem to know if $C_n$ is Noetherian or not for $n \geq 2$ ($C_1$ is a discrete valuation ring, thus it is Noetherian). These rings have similar properties to the Weierstrass systems (stability by partial derivates, stability by division by coordinates, ...), except that there is no Weierstrass Division Theorem. For instance, there exists $f \in C_1$ and $\hat{g} \in k[[x]] \setminus C_1$ such that $f(x) = \hat{g}(x^2)$. This implies that

$$f(x) = (x^2 - y)\hat{h}(x,y) + \hat{g}(y)$$

where $\hat{h}(x,y) \in k[x,y]$ but Equation (5) has no nested solution in $C_1 \times C_2$.

On the other hand, if the rings $C_n$ were Noetherian, since their completions are regular local rings, they would be regular. Then using Example [B.4 iii] we see that they would be excellent (see also [ElKh11]). Thus these rings would satisfy Theorem 2.18 but they do not satisfy Theorem 5.16 since Equation (5) has no solutions in $C_1 \times C_2$.

5.4. Other examples of approximation with constraints. — We present here some examples of positive or negative answers to Problems 1 and 2 in several contexts.

Example 5.19. — [Mi78b] P. Milman proved the following theorem:

**Theorem 5.20.** — Let $f \in \mathbb{C}\{x,y,u,v\}^r$ where $x := (x_1, ..., x_n)$, $y := (y_1, ..., y_m)$, $u := (u_1, ..., u_m)$, $v := (v_1, ..., v_m)$. Then the set of convergent solutions of the following system:

$$\begin{cases}
\frac{\partial u_k}{\partial x_j}(x,y) - \frac{\partial v_k}{\partial y_j}(x,y) = 0 \\
\frac{\partial v_k}{\partial x_j}(x,y) + \frac{\partial u_k}{\partial y_j}(x,y) = 0
\end{cases}$$

(6)

is dense (for the $(x,y)$-adic topology) in the set of formal solutions of this system.

Hints on the proof. — Let $(\hat{u}(x,y), \hat{v}(x,y)) \in \mathbb{C}[x,y]^{2m}$ be a solution of (6). Let us denote $z := x + iy$ and $w := u + iv$. In this case the Cauchy-Riemann equations of (6) are equivalent to $\hat{u}(z) := \hat{u}(x,y) + \hat{v}(x,y) \in \mathbb{C}[z]$ (or in $\mathbb{C}(x)$). Let $\varphi : \mathbb{C}\{z,\bar{z},w,\bar{w}\} \to \mathbb{C}\{z,\bar{z}\}$ and $\psi : \mathbb{C}\{z,w\} \to \mathbb{C}\{z\}$ be the morphisms defined by

$$\varphi(h(z,\bar{z},w,\bar{w})) := h(z,\bar{z},\hat{u}(z),\hat{v}(z)) \quad \text{and} \quad \psi(h(z,w)) := h(z,\hat{w}(z)).$$

Milman proved that

$$\text{Ker}(\varphi) = \text{Ker}(\psi) \cdot \mathbb{C}\{z,\bar{z},w,\bar{w}\} + \text{Ker}(\psi) \cdot \mathbb{C}\{z,\bar{z},w,\bar{w}\}.$$ 

Since $\text{Ker}(\psi)$ (as an ideal of $\mathbb{C}\{z,w\}$) satisfies Theorem 2.1, the result follows.

This proof does not give the existence of an Artin function for this kind of system, since the proof consists in reducing Theorem 5.20 to Theorem 2.1, but this reduction depends on the formal solution of (6). Nevertheless in [Hic-Ro11], it is proven that such a system admits an Artin function using ultraproducts methods. The survey [Mir13] is a good introduction for application of Artin Approximation in CR geometry.
Example 5.21. — [BM79]
Let $G$ be a reductive algebraic group. Suppose that $G$ acts linearly on $\mathbb{C}^n$ and $\mathbb{C}^m$. We say that $y(x) \in \mathbb{C}[x]^m$ is equivariant if $y(\sigma x) = \sigma y(x)$ for all $\sigma \in G$. E. Bierstone and P. Milman proved that, in Theorem 2.1, the constraint for the solutions of being equivariant may be preserved for convergent solutions:

Theorem 5.22. — [BM79] Let $f(x, y) \in \mathbb{C}[x, y]^r$. Then the set of equivariant convergent solutions of $f = 0$ is dense in the set of equivariant formal solutions of $f = 0$ for the $(x)$-adic topology.

This result remains true if we replace $\mathbb{C}$ (resp. $\mathbb{C}(x)$ and $\mathbb{C}(x, y)$) by any field of characteristic zero $k$ (resp. $k(x)$ and $k(x, y)$).

Using ultraproducts methods we may probably prove that Problem 2 has a positive answer in this case.

Example 5.23. — [BDLvdD79] Let $k$ be a field. Let us consider the following differential equation:

$$ a^2 x_1 \frac{\partial f}{\partial x_1}(x_1, x_2) - x_2 \frac{\partial f}{\partial x_2}(x_1, x_2) = \sum_{i,j \geq 1} x_i^1 x_j^j \left( = \left( \frac{x_1}{1 - x_1} \left( \frac{x_2}{1 - x_2} \right) \right) \right). $$

For $a \in k$, $a \neq 0$, this equation has only the following solutions

$$ f(x_1, x_2) := b + \sum_{i,j \geq 1} \frac{x_i^1 x_j^j}{a^2 i - j}, \quad b \in k. $$

Let us consider the following system of equations:

$$ y_0^2 x_1 y_5(x_1, x_2) - x_2 y_7(x_1, x_2) = \sum_{i,j \geq 1} x_i^1 x_j^j \left( \right) $$

$$ y_1(x_1, x_2) = y_2(x_3, x_4, x_5) + (x_1 - x_3) z_1(x) + (x_2 - x_4) z_2(x) $$

$$ y_2(x_3, x_4, x_5) = y_1(x_1, x_2) + x_5 y_4(x_1, x_2) + $$

$$ x_5^2 y_5 + (x_3 - x_1 - x_5) z_3(x) + (x_4 - x_2) z_4(x) $$

$$ y_3(x_3, x_4, x_5) = y_1(x_1, x_2) + x_5 y_6(x_1, x_2) + $$

$$ x_5^2 y_7 + (x_3 - x_1) z_5(x) + (x_4 - x_2) z_6(x) $$

$$ y_8(x_1, x_2) = y_9(x_3, x_4, x_5) \text{ i.e. } y_9 \in k \text{ and } y_8 y_9 = 1. $$

It is straightforward, by Lemma 5.1 and Example 5.2, to check that $(a, f(x_1, x_2))$ is a solution of (7) if and only if (8) has a solution with $y_1 = f$ and $y_8 = a$. Moreover, if $y_1, ..., y_10, z_1, ..., z_6$ is a solution of Equation (8), then $(y_8, y_1)$ is a solution of (7). Thus (7) has no solution in $\mathbb{Q}[x]$. But clearly, (7) has solutions in $\mathbb{Q}(x)$ for any $c \in \mathbb{N}$ and the same is true for (8). This shows that Proposition 3.25 is not valid if the base field is not $\mathbb{C}$.

Example 5.24. — [BDLvdD79] Let us assume that $k = \mathbb{C}$ and consider the latter example. The system of equations (8) does not admit an Artin function. Indeed, for any $c \in \mathbb{N}$, there is $a_c \in \mathbb{Q}$, such that (8) has a solution modulo $(x)^c$ with $y_8 = a_c$. 


But there is no solution in \( \mathbb{C}[x] \) with \( y_k = a_c \) modulo \( (x) \), otherwise \( y_k = a_c \) which is not possible.

Thus systems of equations with constraints do not satisfy Problem 2 in general.

**Example 5.25.** — [Ron08] Let \( \varphi : \mathbb{C}\{x\} \to \mathbb{C}\{y\} \) be a morphism of complex analytic algebras and let us denote \( \varphi_i(y) := \varphi(x_i) \). Let us denote by \( \hat{\varphi} : \mathbb{C}[x] \to \mathbb{C}[y] \) the morphism induced on the completions. According to a lemma of Chevalley (Lemma 7 of [Ch43]), there exists a function \( \beta : \mathbb{N} \to \mathbb{N} \) such that \( \varphi^{-1}(y) \beta(c) \subset \text{Ker}(\hat{\varphi}) \) for all \( c \in \mathbb{N} \). It is called the Chevalley function of \( \varphi \). Using Lemma 5.1 we check easily that this function \( \beta \) satisfies the following statement (in fact both statements are equivalent [Ron08]): Let \( \tilde{f}(x) \in \mathbb{C}[x] \) and \( \tilde{h}_i(x, y) \in \mathbb{C}[x, y] \), \( 1 \leq i \leq n \), satisfy

\[
\tilde{f}(x) + \sum_{i=1}^{n}(x_i - \varphi_i(y))\tilde{h}_i(x, y) \in (x, y)^\beta(c).
\]

Then there exists \( \tilde{g}(y) \in \mathbb{C}[x], \tilde{h}_i(x, y) \in \mathbb{C}[x, y], 1 \leq i \leq n \), such that

\[
(9) \quad \tilde{g}(y) + \sum_{i=1}^{n}(x_i - \varphi_i(y))\tilde{h}_i(x, y) = 0.
\]

and \( \tilde{f}(x) - \tilde{f}(x) \in (x, y)^c \), \( \tilde{h}_i(x, y) - \tilde{h}_i(x, y) \in (x, y)^c, 1 \leq i \leq n \).

In particular Problem 2 has a positive answer for Equation (9), but not Problem 1 (see Example 5.12). In fact, the conditions of Theorem 5.15 are equivalent to the fact that \( \beta \) is bounded by an affine function [Lz86].

The following example is given in [Ron08] and is inspired by Example 5.12. Let \( \alpha : \mathbb{N} \to \mathbb{N} \) be an increasing function. Let \( (n_i)_i \) be a sequence of integers such that \( n_{i+1} > \alpha(n_i + 1) \) for all \( i \) and such that the convergent power series \( \xi(Y) := \sum_{i \geq 1} Y^{n_i} \) is not algebraic over \( \mathbb{C}(Y) \). Then we define the morphism \( \varphi : \mathbb{C}\{x_1, x_2, x_3\} \to \mathbb{C}\{y_1, y_2\} \) in the following way:

\[
(\varphi(x_1), \varphi(x_2), \varphi(x_3)) = (y_1, y_1y_2, y_1\xi(y_2)).
\]

It is easy to prove that \( \hat{\varphi} \) is injective following Example 5.12. For any integer \( i \) we define:

\[
\tilde{f}_i := x_1^{n_i} x_3 - \left(x_2^{n_1} x_1^{n_1 - n_2} + \cdots + x_2^{n_2} x_1^{n_2 - n_1} x_3 + x_2^{n_2}\right).
\]

Then

\[
\varphi(\tilde{f}_i) = y_1^{n_1}\xi(y_2) - y_1^{n_1} \sum_{k=1}^{i} y_2^{n_k} \in (y)^{n_1 + n_i + 1} \subset (y)^{\alpha(n_i + 1)}
\]

but \( \tilde{f}_i \notin (x)^{n_1 + 1} \) for any \( i \). Thus the Chevalley function of \( \varphi \) satisfies \( \beta(n_1 + 1) > \alpha(n_i + 1) \) for all \( i \in \mathbb{N} \). Hence \( \lim \sup \frac{\beta(c)}{\alpha(c)} \geq 1 \). In particular if the growth of \( \alpha \) is too big, then \( \beta \) is not recursive.
Appendix A

Weierstrass Preparation Theorem

In this part set \(x := (x_1, \ldots, x_n)\) and \(x' := (x_1, \ldots, x_{n-1})\). Moreover \(k\) will denote a local ring of maximal ideal \(m\) (if \(k\) is a field, \(m = (0)\)). A local subring of \(k[[x]]\) will be a subring of \(k[x]\) which is a local ring and whose maximal ideal is generated by \((m + (x)) \cap A\).

**Definition A.1.** — If \(f \in k[[x]]\) we say that \(f\) is regular of order \(d\) with respect to \(x_n\) if \(f = ux_n^d\) modulo \(m + (x')\) where \(u\) is invertible in \(\frac{k[[x]]}{m + (x')} \cong \frac{k}{m}[x_n]\).

**Definition A.2.** — Let \(A\) be a local subring of \(k[[x]]\). We say that \(A\) has the Weierstrass Division Property if for any \(f, g \in A\) such that \(f\) is regular of order \(d\) with respect to \(x_n\), there exist \(q \in A\) and \(r \in (A \cap k[x'])[x_n]\) such that \(\deg_{x_n}(r) < d\) and \(g = qf + r\).

**Definition A.3.** — Let \(A\) be a local subring of \(k[[x]]\). We say that \(A\) satisfies the Weierstrass Preparation Theorem if for any \(f \in A\) which is regular with respect to \(x_n\), there exist an integer \(d\), a unit \(u \in A\) and \(a_1(x'), \ldots, a_d(x') \in A \cap (x')k[[x']]\) such that

\[
f = u \left( x_n^d + a_1(x')x_n^{d-1} + \cdots + a_d(x') \right).
\]

In this case \(f\) is necessarily regular of order \(d\) with respect to \(x_n\).

**Remark A.4.** — Clearly, if they exist, \(q\) and \(r\) are unique in Definition A.2. The same is true for \(u\) and the \(a_i(x')\) in Definition A.3.

**Lemma A.5.** — If a local subring \(A\) of \(k[[x]]\) has the Weierstrass Division Property then it satisfies the Weierstrass Preparation Theorem.

**Proof.** — If \(A\) has the Weierstrass Division Property and if \(f \in A\) is regular of order \(d\) with respect to \(x_n\), then we can write \(x_n^d = qf + r\) where \(r \in (A \cap k[x'])[x_n]\) such that \(\deg_{x_n}(r) < d\). Thus \(qf = x_n^d - r\). Since \(f\) is regular of order \(d\) with respect to \(x_n\), then \(q\) is invertible in \(k[[x]]\) and \(r \in (m + (x'))\). Thus \(q \notin (m + (x))\) and \(q\) is invertible in \(A\). Hence \(f = q^{-1}(x_n^d - r)\).

In fact the converse implication is true under some mild conditions:

**Lemma A.6.** — Let \(A_n\) be a subring of \(k[x_1, \ldots, x_n]\) for all \(n \in \mathbb{N}\) such that

i) \(A_{n+m} \cap k[x_1, \ldots, x_n] = A_n\) for all \(n\) and \(m\),

ii) if \(f \in A_n\) is written \(f = \sum_{k \in \mathbb{N}} f_k x_n^k\) with \(f_k \in k[[x']]\) for all \(k\), then \(f_k \in A_{n-1}\) for all \(k\).

iii) \(A_n\) is stable by permutation of the \(x_i\).

Then \(A_n\) has the Weierstrass Division Property if \(A_n\) and \(A_{n+1}\) have the Weierstrass Preparation Property.
Proof. — Let \( f(x) \in A_n \) be \( x_n \)-regular of order \( d \). By the Weierstrass Preparation Property for \( A_n \) we may write
\[
    f = u \left( x_n^d + a_1(x')x_n^{d-1} + \cdots + a_d(x') \right) = uP
\]
where \( u \) is a unit in \( A_n \) and \( P \in A_{n-1}[x_n] \). Now let \( g(x) \in A_n \) and set \( h := P + x_{n+1}g \). Then \( h \) is also \( x_n \)-regular of order \( d \), thus by the Weierstrass Preparation Property for \( A_{n+1} \) we may write \( h = vQ \) where \( v \) is a unit and \( Q \) a polynomial of degree \( d \) in \( x_n \). Let us write
\[
    v = \sum_{k \in \mathbb{N}} v_k x_{n+1}^k \quad \text{and} \quad Q = \sum_{k \in \mathbb{N}} Q_k x_{n+1}^k
\]
where \( v_k \in A_n \) and \( Q_k \in A_{n-1}[x_n] \) for all \( k \) (\( \deg x_n(Q_k) \leq d \)). Thus we deduce from this that
\[
    v_0Q_0 = P \quad \text{and} \quad v_1Q_0 + v_0Q_1 = g.
\]
By unicity of the decomposition in Weierstrass Preparation Theorem, the first equality implies that \( v_0 = 1 \) and \( Q_0 = P \). Thus the second yields \( g = v_1P + Q_1 \), i.e.
\[
    g = v_1u^{-1}f + Q_1
\]
and \( Q_1 \in A_{n-1}[x_n] \) is a polynomial in \( x_n \) of degree \( \leq d \).

\[ \square \]

Theorem A.7. — The following rings have the Weierstrass Division Property:

i) The ring \( A = k[[x]] \) where \( k \) is complete local ring (\([Bo65]\)).

ii) The ring \( A = k(x) \) of algebraic power series where \( k \) is a field or a Noetherian Henselian local ring of characteristic zero which is analytically normal (\([Laf65]\) and \([Laf67]\)).

iii) The ring \( A = k\{x\} \) of convergent power series over a valued field \( k \) (\([Na62]\)).

Appendix B

Regular morphisms and excellent rings

Definition B.1. — Let \( \varphi : A \rightarrow B \) be a morphism of Noetherian rings. We say that \( \varphi \) is regular if it is flat and if for any prime ideal \( \mathcal{P} \) of \( A \), the \( \kappa(\mathcal{P}) \)-algebra \( B \otimes_A \kappa(\mathcal{P}) \) is geometrically regular (where \( \kappa(\mathcal{P}) := \frac{A_{\mathcal{P}}}{\mathcal{P}_{\mathcal{P}}} \) is the residue field of \( A_{\mathcal{P}} \)). This means that \( B \otimes_A \kappa \) is a regular Noetherian ring for any finite field extension of \( \kappa(\mathcal{P}) \).

Example B.2. —

i) If \( A \) and \( B \) are fields, then \( A \rightarrow B \) is regular if and only if \( B \) is a separable field extension of \( A \).

ii) If \( A \) is excellent, for any ideal \( I \) of \( A \), the morphism \( A \rightarrow \hat{A} \) is regular where \( \hat{A} := \varprojlim \frac{A}{I^n} \) is the \( I \)-adic completion of \( A \) (cf. \([GrDi65]\) 7.8.3).

iii) If \( V \) is a discrete valuation ring, then the completion morphism \( V \rightarrow \hat{V} \) is regular if and only if \( \text{Frac}(V) \rightarrow \text{Frac}(\hat{V}) \) is separable. Indeed, \( V \rightarrow \hat{V} \) is always flat and this morphism induces an isomorphism on the residue fields.
iv) Let $X$ be compact Nash manifold, let $\mathcal{N}(X)$ be the ring of Nash functions on $X$ and let $\mathcal{O}(X)$ be the ring of real analytic functions on $X$. Then the natural inclusion $\mathcal{N}(X) \to \mathcal{O}(X)$ is regular (cf. [CRS95]).

v) Let $L \subset \mathbb{C}^n$ be a compact polynomial polyhedron and $B$ the ring of holomorphic function germs at $L$. Then the morphism of constants $\mathbb{C} \to B$ is regular (cf. [Le95]). This example and the previous one allow to use Theorem 2.17 to show global approximation results in complex geometry or real geometry.

In the case of Artin Approximation, we will be mostly interested in the morphism $A \to \hat{A}$. Thus we need to know what is an excellent ring.

**Definition B.3.** — A Noetherian ring $A$ is excellent if the following conditions hold:

i) $A$ is universally catenary.

ii) For any $p \in \text{Spec}(A)$, the formal fibre of $A_p$ is geometrically regular.

iii) For any $p \in \text{Spec}(A)$ and for any finite separable extension $\mathbb{K}$, there exists a finitely generated sub-$\frac{A}{p}$-algebra $B$ of $\mathbb{K}$, containing $\frac{A}{p}$, and such that $\text{Frac}(B) = \mathbb{K}$ and the set of regular points of $\text{Spec}(B)$ contains a non-empty open set.

This definition may be a bit obscure at first sight. Thus we give here the main examples of excellent rings:

**Example B.4.** —

i) Local complete rings (thus any field) are excellent. Dedekind rings of characteristic zero are excellent. Any ring which is essentially of finite type over an excellent ring is excellent. ([GrDi65] 7-8-3).

ii) If $k$ is a complete valued field, then $k\{x_1, ..., x_n\}$ is excellent [Ki69].

iii) We have the following result: Let $A$ be a regular ring containing a field of characteristic zero denoted by $k$. Suppose that for any maximal ideal $m$, the field extension $k \to \frac{k}{m}$ is algebraic and $\text{ht}(m) = n$. Suppose moreover that there exist $D_1, ..., D_n \in \text{Der}_k(A)$ and $x_1, ..., x_n \in A$ such that $D_i(x_j) = \delta_{i,j}$. Then $A$ is excellent (cf. Theorem 102 [Mat80]).

iv) A Noetherian local ring $A$ is excellent if and only if it is universally catenary and $A \to \hat{A}$ is regular ([GrDi65] 7-8-3 i)). In particular, if $A$ is a quotient of a local regular ring, then $A$ is excellent if and only if $A \to \hat{A}$ is regular (cf. [GrDi65] 5-6-4).

**Example B.5.** — [Na62] [Mat80] Let $k$ be a field of characteristic $p > 0$ such that $[k:k^p] = \infty$ (for instance let us take $k = \mathbb{F}_p(t_1, ..., t_n,...)$). Let $V := k^p[[x]]$ where $x$ is a single variable, i.e. $V$ is the ring of power series $\sum_{i=0}^{\infty} a_i x^i$ such that $[k^p(a_0, a_1, ...)] : k^p < \infty$. Then $V$ is a discrete valuation ring whose completion is $k[[x]]$ and it is a Henselian ring.

We have $\hat{V} \subset V$, thus $[\text{Frac}(\hat{V}) : \text{Frac}(V)]$ is purely inseparable. Hence $V \to \hat{V}$ is not regular by Example B.2, and $V$ is not excellent by Example B.4 (v).

On the other hand, let $f$ be the power series $\sum_{i=0}^{\infty} a_i x^i$, $a_i \in k$ such that $[k^p(a_0, a_1, ...)] :$
\[ \mathbb{k}^p = \infty. \] Then \( f \in \mathbb{V} \) but \( f \notin V \), and \( f^p \in V \). Thus \( f \) is the only root of the polynomial \( y^p - f^p \). This shows that the polynomial \( y^p - f^p \in V[y] \) does not satisfy Theorem 2.17.

## Appendix C

### Étale morphisms and Henselian rings

The material presented here is very classical and has first been studied by G. Azumaya and M. Nagata. We will give a quick review of the definitions and properties that we need for the understanding of the rest of the chapter. Nevertheless, the reader may consult [Na62], [GrDi65], [Ra70] or [Iv73].

**Example C.1** — In classical algebraic geometry, the Zariski topology has too few open sets. For instance, there is no Implicit Function Theorem. Let \( X \) be the zero set of the polynomial \( y^2 - x^2(x + 1) \) in \( \mathbb{C}^2 \). On an affine open neighborhood of 0, denoted by \( U \), \( X \cap U \) is equal to \( X \) minus a finite number of points, thus \( X \cap U \) is irreducible since \( X \) is irreducible. In the analytic topology, we can find an open neighborhood of 0, denoted by \( U \), such that \( X \cap U \) is reducible, for instance take \( U = \{(x, y) \in \mathbb{C}^2 / |x|^2 + |y|^2 < 1/2\} \). This comes from the fact that \( x^2(1 + x) \) is the square of an analytic function defined on \( U \cap (\mathbb{C} \times \{0\}) \). Let \( z(x) \) be such an analytic function, \( z(x)^2 = x^2(1 + x) \).

In fact we can obtain \( z(x) \) from the Implicit Function Theorem. We see that \( z(x) \) is a root of the polynomial \( Q(x, z) := z^2 - x^2(1 + x) \). We have \( Q(0, 0) = \frac{\partial Q}{\partial t}(0, 0) = 0 \), thus we can not use directly the Implicit Function Theorem to obtain \( z(x) \) from its minimal polynomial.

Nevertheless let us take \( P(x, t) := (t + 1)^2 - (1 + x) = t^2 + 2t - x \). Then \( P(0, 0) = 0 \) and \( \frac{\partial P}{\partial t}(0, 0) = 2 \neq 0 \). Thus, from the Implicit function Theorem, there exists \( t(x) \) analytic on a neighborhood of 0 such that \( t(0) = 0 \) and \( P(x, t(x)) = 0 \). If we denote \( z(x) := x(1 + t(x)) \), we have \( z^2(x) = x^2(1 + x) \). In fact \( z(x) \in B := \mathbb{C}^{[x, t(x), t]}(P(x, t)) \). The morphism \( \mathbb{C}[x] \rightarrow B \) is an example of étale morphism.

**Definition C.2** — Let \( \varphi : A \rightarrow B \) be a ring morphism. We say that \( \varphi \) is smooth (resp. étale) if for any \( A \)-algebra \( C \) along with an ideal \( I \) such that \( I^2 = (0) \) and any morphism of \( A \)-algebras \( \psi : B \rightarrow C \) there exists a morphism \( \sigma : B \rightarrow C \) (resp. a unique morphism) such that the following diagram commutes:

\[
\begin{array}{ccc}
\mathbb{A} & \xrightarrow{\varphi} & \mathbb{B} \\
\downarrow{\psi} & & \downarrow{\sigma} \\
\mathbb{C} & \xrightarrow{\psi} & \mathbb{C} \\
\end{array}
\]

**Example C.3** — Let \( k := \mathbb{R} \) or \( \mathbb{C} \) and let us assume that \( A = \mathbb{k}[x_1, \ldots, x_n]_J \) and \( B = \mathbb{A}[y_1, \ldots, y_m]_K \) for some ideals \( J \) and \( K \). Let \( X \) be the zero locus of \( J \) in \( \mathbb{k}^n \) and \( Y \) be the zero locus of \( K \) in \( \mathbb{k}^{n+m} \). The morphism \( \varphi : A \rightarrow B \) defines a regular map
\( \Phi : Y \to X \). Let \( C := k[\tau_j] \) and \( I := (t) \). Let \( f_1(x), \ldots, f_r(x) \) be generators of \( J \).

A morphism \( A \to C \) is given by elements \( a_j, b_i \in \mathbb{k} \) such that \( f_j(a_1 + b_1 t, \ldots, a_n + b_n t) \in (t)^2 \) for \( 1 \leq j \leq r \). We have

\[
f_j(a_1 + b_1 t, \ldots, a_n + b_n t) = f_j(a_1, \ldots, a_n) + \left( \sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(a_1, \ldots, a_n)b_i \right) t \text{ mod. } (t)^2.
\]

Thus a morphism \( A \to C \) is given by a point \( x := (a_1, \ldots, a_n) \in X \) (i.e. such that \( f_j(a_1, \ldots, a_n) = 0 \) for all \( j \)) and a tangent vector \( u := (b_1, \ldots, b_n) \) to \( X \) at \( x \) (i.e. such that \( \sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(a_1, \ldots, a_n)b_i = 0 \) for all \( j \)). In the same way a \( A \)-morphism \( B \to \mathbb{C} \) is given by a point \( y \in Y \). Moreover the first diagram commutes if and only if \( \Phi(y) = x \).

Then \( \varphi \) is smooth if for any \( x \in X \), any \( y \in Y \) and any tangent vector \( u \) to \( X \) at \( x \) such that \( \Phi(y) = x \), there exists a tangent vector \( v \) to \( Y \) at \( y \) such that \( D_y(\Phi)(v) = u \). And \( \varphi \) is étale if and only if \( v \) is unique. This shows that smooth morphisms correspond to submersions and étale morphisms to local diffeomorphisms.

**Example C.4.** — Let \( \varphi : A \to B_p \) be the canonical morphism where \( B := A[x]/(P(x)) \) and \( p \) is a prime ideal of \( B \) such that \( \frac{\partial P}{\partial x}(x) \notin p \). If we have such a commutative diagram

\[
\begin{array}{ccc}
A & \xrightarrow{\varphi} & B_p \\
\downarrow & & \downarrow \\
C & \xrightarrow{\psi} & \mathbb{C} \\
\end{array}
\]

then the morphism \( B_p \to \mathbb{C} \) is given by an element \( c \in C \) such that \( P(c) \in I \).

Looking for a lifting of \( \psi \) is equivalent to finding \( \in I \) such that \( P(c + \varepsilon) = 0 \). We have

\[
P(c + \varepsilon) = P(c) + \frac{\partial P}{\partial x}(c)\varepsilon
\]

since \( I^2 = (0) \). Since \( \frac{\partial P}{\partial x} \) is invertible in \( B_p \), \( \frac{\partial P}{\partial x}(c) \) is invertible in \( \mathbb{C} \), i.e. there exists \( a \in C \) such that \( a\frac{\partial P}{\partial x}(c) = 1 \) mod. \( I \). Moreover \( a \) is unique modulo \( I \). For any \( \eta \in I \) let us set \( \varepsilon := -P(c)(a + \eta) \). Since \( P(c) \in I \), \( \varepsilon \) does not depend on \( \eta \) and the lifting of \( \psi \) is unique. This proves that \( \varphi \) is étale. Compare this example with Example C.1

**Definition C.5.** — Étale morphisms of Example C.4 are called standard étale morphisms. We can prove that if \( A \) and \( B \) are local rings then any étale morphism is standard ([IV, III. 2]).

**Example C.6 (Jacobian Criterion).** — We can generalize the former example as follows. If \( \mathbb{k} \) is a field and \( \varphi : \mathbb{k} \to B := \mathbb{k}[x_1, \ldots, x_n]/(g_1, \ldots, g_r) \) where \( \mathfrak{m} := (x_1 - c_1, \ldots, x_n - c_n) \) then \( \varphi \) is smooth if and only if the jacobian matrix \( \left( \frac{\partial g_i}{\partial x_j}(c) \right) \) has rank equal to the
height of \((g_1, \ldots, g_r)\). This is equivalent to say that \(V(I)\) has a non-singular point at the origin. Let us recall that the fibers of submersions are always smooth.

**Definition C.7.** — Let \(A\) be a local ring. An étale neighbourhood of \(A\) is an étale local morphism \(A \rightarrow B\) inducing an isomorphism on the residue fields. If \(A\) is a local ring, the étale neighbourhoods of \(A\) form a filtered inductive limit and the limit of this system is called the Henselization of \(A\) ([Lü73] III. 6. or [Ra69] VIII) and denoted by \(A^h\).

We say that \(A\) is Henselian if \(A = A^h\). The morphism \(A \rightarrow A^h\) is universal among all the morphisms \(A \rightarrow B\) inducing an isomorphisms on the residue fields and where \(B\) is Henselian.

**Proposition C.8.** — If \(A\) is a Noetherian local ring, then its Henselization \(A^h\) is a Noetherian local ring and \(A \rightarrow A^h\) is faithfully flat. If \(\varphi : A^h \rightarrow B\) is an étale neighbourhood of \(A^h\), then there is a section \(\sigma : B \rightarrow A\), i.e. \(\sigma \circ \varphi = \text{id}_{A^h}\).

**Proposition C.9.** — Let \(A\) be a Henselian local ring and let \(\varphi : A \rightarrow B\) be an étale neighbourhood that admits a section in \(\frac{A}{\mathfrak{m}_A}\) for \(c \geq 1\), i.e. a morphism of \(A\)-algebra \(\sigma : B \rightarrow \frac{A}{\mathfrak{m}_A^c}\). Then there exists a section \(\tilde{s} : B \rightarrow A\) such that \(\tilde{s} = s \mod \mathfrak{m}_A^c\).

**Proof.** — Since \(A\) is Henselian and \(\varphi\) is étale then \(A\) is isomorphic to the Henselization of \(B\). Moreover \(\frac{A}{\mathfrak{m}_A}\) is Henselian. The result comes from the universal property of the Henselization. \(\blacksquare\)

**Definition C.10.** — Let \(A\) be a Henselian local ring and \(x := (x_1, \ldots, x_n)\). Then the Henselization of \(A[x]/\mathfrak{m}_A+(x)\) is denoted by \(A(x)\).

**Remark C.11.** — Let \(P(y) \in A[y]\) and \(a \in A\) satisfy \(P(a) \in \mathfrak{m}_A\) and \(\frac{\partial P}{\partial y}(a) \notin \mathfrak{m}_A\). If \(A\) is Henselian, then \(A \rightarrow A[y]/(P(y)+\mathfrak{m}_A+(y-a))\) is an étale neighborhood of \(A\), thus it admits a section. This means that there exists \(\tilde{y} \in \mathfrak{m}_A\) such that \(P(a + \tilde{y}) = 0\). If \(A\) is a local ring, then any étale neighborhood of \(A\) is of the previous form. Thus, by Proposition C.8 we have the following proposition:

**Proposition C.12.** — Let \(A\) be a local ring. Then \(A\) is Henselian if and only if for any \(P(y) \in A[y]\) and \(a \in A\) such that \(P(a) \in \mathfrak{m}_A\) and \(\frac{\partial P}{\partial y}(a) \notin \mathfrak{m}_A\) there exists \(\tilde{y} \in \mathfrak{m}_A\) such that \(P(a + \tilde{y}) = 0\).

We can generalize this proposition as follows:

**Theorem C.13 (Implicit Function Theorem).** — Set \(y = (y_1, \ldots, y_m)\) and let \(f(y) \in A[y]^r\) with \(r \leq m\). Let \(J\) be the ideal of \(A[y]\) generated by the \(r \times r\) minors of the Jacobian matrix of \(f(y)\). If \(A\) is Henselian, \(f(0) = 0\) and \(J \notin \mathfrak{m}_A\), then there exists \(\tilde{y} \in \mathfrak{m}_A^2\) such that \(f(\tilde{y}) = 0\).

**Example C.14.** — The ring of germs of \(C^\infty\) function at the origin of \(\mathbb{R}^n\) is a Henselian local ring but it is not Noetherian. The ring of germ of analytic functions at the origin of \(C^n\) is a Noetherian Henselian local ring; it is the ring of convergent power series.
Example C.15. — If $A = k[[x_1, ..., x_n]]$ for some Weierstrass system over $k$, then $A$ is a Henselian local ring by Proposition C.12. Indeed, let $P(y) \in A[y]$ satisfies $P(0) = 0$ and $\frac{\partial P}{\partial y}(0) \not\in (p, x)$. Thus $P(y)$ contains a nonzero term of the form $cy$, $c \in k^*$. Then we have $y = P(y)Q(y) + R$ where $R \in m_A$. Clearly $Q(y)$ is a unit, thus $P(R) = 0$.

Proposition C.16 (Hensel Lemma). — Let $(A, mA)$ be a local ring. Then $A$ is Henselian if and only if for any monic polynomial $P(y) \in A[y]$ such that $P(y) = f(y)g(y) \mod mA$ for some $f(y), g(y) \in A[y]$ which are coprime modulo $mA$, there exists $f(y), g(y) \in A[y]$ such that $P(y) = f(y)\bar{g}(y)$ and $f(y) - f(y), \bar{g}(y) - g(y) \in mA[y]$.

Proof. — Let us prove the sufficiency of the condition. Let $P(y) \in A[y]$ and $a \in A$ satisfy $P(a) \in mA$ and $\frac{\partial P}{\partial y}(a) \notin mA$. This means that $P(X) = (X - a)Q(X)$ where $X - a$ and $Q(X)$ are coprime modulo $m$. Then this factorization lifts to $A[X]$, this means $\bar{y} \in mA$ such that $P(a + \bar{y}) = 0$. This proves that $A$ is Henselian.

To prove that the condition is necessary, let $P(y) \in A[y]$ be a monic polynomial, $P(y) = y^2 + a_1y^{d-1} + \cdots + a_d$. Let $k := \frac{A}{mA}$ be the residue field of $A$, and for any $a \in A$, let us write $\bar{a}$ for the image of $a$ in $k$. Let us assume that $\bar{P}(y) = f(y)g(y)$ mod $mA$ for some $f(y), g(y) \in k[y]$ which are coprime in $k[y]$. Let us write $f(y) = \bar{f}(y)y^\delta + b_1y^{\delta-1} + \cdots + b_\delta$, $g(y) = \bar{g}(y)y^\delta + c_1y^{\delta-1} + \cdots + c_\delta$ where $b = (b_1, \cdots, b_\delta) \in k^{d_1}$, $c = (c_1, \cdots, c_\delta) \in k^{d_2}$. The product of polynomials $\bar{P} = fg$ defines a map $\Phi : k^{d_1} \times k^{d_2} \to k^d$, that is polynomial in $b$ and $c$ with integer coefficients, and $\Phi(b, c) = \bar{a} := (\bar{a}_1, ..., \bar{a}_d)$. The determinant of the Jacobian matrix $\frac{\partial a}{\partial (b, c)}$ is the resultant of $f(y)$ and $g(y)$, and hence is nonzero at $(b, c)$. Using the Implicit Function Theorem C.13 there exist $\tilde{b} \in A^{d_1}$, $\tilde{c} \in A^{d_2}$ such that $P(y) = P_1(y)P_2(y)$ where $P_1(y) = y^\delta + \tilde{b}_1y^{\delta-1} + \cdots + \tilde{b}_\delta$, and $P_2(y) = y^\delta + \tilde{c}_1y^{\delta-1} + \cdots + \tilde{c}_\delta$.

Proposition C.17. — ([GrDi67] 18-7-6) If $A$ is an excellent local ring, then its Henselization $A^h$ is also an excellent local ring.

References


M. Raynaud, Grothendieck et la théorie des schémas, unpublished notes.


