Inventiones math. 51, 189–203 (1979)

Inventiones mathematicae

Ultraproducts and Approximation in Local Rings I

Joseph Becker[1]*, J. Denef[2]**, L. Lipshitz[3], and L. van den Dries[4]***

[1] Purdue University, West Lafayette, Indiana 47907
[2] Princeton University, Princeton, New Jersey 08540
[3] Institute for Advanced Study, Princeton, New Jersey 08540, USA
[4] University of Utrecht, Utrecht, Netherlands

In this paper we give new and quite elementary proofs of the strong approximation theorems of M. Greenberg (Theorem 2.1), M. Artin (Theorem 3.2) and D. Popescu (Theorem 3.3). Our proofs use the ultraproduct construction. In § 1 we give a brief outline of this construction and derive the properties we shall use.

Our method of proof gives a quite general way of deriving a strong approximation theorem from the corresponding approximation theorem. (By an "approximation theorem" we mean a theorem of the type of Theorem 3.1 and by a "strong approximation theorem" we mean a theorem of the type of Theorem 3.2 or 3.3.) For example if we have an approximation theorem with some of the Y's constrained to depend only on some of the X's then we can immediately derive the strong approximation theorem with the same constraints. In § 4 we give a pair of such theorems with the constraints that some of the Y's depend only on X_1.

In § 5 we present some counterexamples: Let k be the algebraic closure of the field with p elements. There is a system of polynomials $f(X_1, X_2, Y_1, \ldots, Y_7)$ over k such that the system $f(X_1, X_2, Y_1, \ldots, Y_7)=0$; $Y_1, Y_6 \in k[\![X_2]\!]$; $Y_2, Y_3, Y_6 \in k[\![X_1]\!]$ has a solution mod $(X_1, X_2)^j$, for every $j \in \mathbb{N}$ but has no solution in $k[\![X_1, X_2]\!]$. We also give a similar counterexample over the rationals.

In § 6 we discuss effectivity.

In Ultraproducts and Approximation in Local Rings II we shall extend the methods of this paper to give proofs of some further strong approximation theorems, e.g. the Theorem of Pfister and Popescu [17] as well as some new results.

The usefulness of the ultraproduct construction (or equivalently an appropriate form of the Gödel Completeness and Compactness Theorems) in analyzing algebraic questions about valued fields is due to Ax and Kochen [3] (cf. also

* Supported in part by the N.S.F.
** Supported by the Belgian Nationaal Fonds voor Wetenschappelijk Onderzoek
*** The proofs in Sects. 2, 3 and 6 of this paper were found by L. van den Dries and the other three authors independently. Van den Dries would like to thank W. Baur for a helpful communication in connection with the proof of Theorem 2.1 below

0020-9910/79/0051/0189/$03.00

Eršov [9]). In [3] the similarity between the fields $\mathbb{Q}_p$ of p-adic numbers and the fields $F_p((t))$ of formal power series over the p-element field is made precise by the theorem (under the continuum hypothesis) that if D is a nonprincipal ultrafilter on the set of primes (see the definitions in § 1 below) then the ultraproducts $(\prod_p \mathbb{Q}_p)/D$ and $(\prod_p F_p((t)))/D$ are isomorphic. (The proof of this theorem is entirely algebraic.) This has the immediate consequence that for any first order statement φ about valued fields, φ is true in all but a finite number of the $\mathbb{Q}_p$'s if and only if φ is true in all but a finite number of the fields $F_p((t))$.

Our use of the ultraproduct construction is somewhat similar.[1] Classically the additional difficulty in proving strong approximation theorems, over and above that of proving approximation theorems, seems to be that one does not have a point at which to apply the algebraic and geometric methods – one only has an "approximate point". The ultraproduct construction eliminates this difficulty by providing an appropriate point and hence allows one to deduce the strong approximation theorems from corresponding approximation theorems.

As a simple illustration of these ideas consider the following well-known result. (This example does not quite fit the pattern given above but illustrates the role played by the ultraproduct construction.) Let $f(X, Y) \in \mathbb{C}[X, Y]$ ($X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_N)$) and suppose that for every $j \in \mathbb{N}$ there is a $\bar{y}_j \in \mathbb{C}[X]$ such that $f(X, \bar{y}_j) \equiv 0 \bmod (X)^j$. Then there is a $y \in \mathbb{C}[\![X]\!]$ such that $f(X, y) = 0$. One can give a very simple proof of this as follows. Let * denote the ultrapower with respect to a nonprincipal ultrafilter on $\mathbb{N}$, (see § 1). Then from the assumption that for every $j \in \mathbb{N}$ there exists a $\bar{y}_j$ such that $f(X, \bar{y}_j) \equiv 0 \bmod (X)^j$ we know that in $\mathbb{C}[X]^*$ there is a $\bar{y}$ such that $f(X, \bar{y}) \equiv 0 \bmod (X)^\infty$, where $(X)^\infty = \bigcap_{j \in \mathbb{N}} (X)^j \subset \mathbb{C}[X]^*$. But $\mathbb{C}[X]^*/(X)^\infty \cong \mathbb{C}^*[\![X]\!]$ and hence there is a $\bar{y} \in \mathbb{C}^*[\![X]\!]$ such that $f(X, \bar{y}) = 0$. Also $\mathbb{C}^* \cong \mathbb{C}$ (over the subfield generated by the coefficients of f, since they are both algebraically closed and of transcendence degree $2^{\aleph_0}$). Hence the required solution in $\mathbb{C}[\![X]\!]$ exists. Notice that the same proof works if we have constraints that some of the Y_i's depend only on some of the X_j's.

A. Robinson in [19] gave a proof of a special case of Greenberg's Theorem, using the results of Ax and Kochen [3]. Motivation has also come from the paper of Nerode [16].

§ 1. Ultraproducts. The reader who is familiar with the ultraproduct construction can skip this section. For the reader who is unfamiliar with this construction we shall give a brief survey of the facts we shall use. More information about ultraproducts can be found in § 1 of [3] or any standard model theory text such as [7].

In the cases we shall consider the index set will always be $\mathbb{N}$, the natural numbers. A *filter* on $\mathbb{N}$ is a nonempty family D of subsets of $\mathbb{N}$ satisfying (i) $\phi \notin D$, (ii) if $s, t \in D$ then $s \cap t \in D$ and (iii) if $s \in D$ and $s \subseteq t \subseteq \mathbb{N}$ then $t \in D$. A filter D on $\mathbb{N}$ is *principal* if there is an $r \in D$ such that $D = \{t \mid r \subseteq t \subseteq \mathbb{N}\}$, otherwise it is

[1] It is well-known that the results of Ax and Kochen and Eršov don't generalise to local rings of dimension greater than one, see for example [8]

nonprincipal. An *ultrafilter* D on $\mathbb{N}$ is a filter on $\mathbb{N}$ which is maximal with respect to inclusion in the class of all filters on $\mathbb{N}$. It is an exercise to show that a filter D on $\mathbb{N}$ is an ultrafilter if and only if for every $r \subseteq \mathbb{N}$ exactly one of r or $\mathbb{N}-r$ belongs to D. It is also easy to show that an ultrafilter D is nonprincipal if and only if it contains the Fréchet filter F of cofinite subsets of $\mathbb{N}$, i.e. $D \supseteq F = \{s \subseteq \mathbb{N} \mid \mathbb{N}-s \text{ is finite}\}$. The existence of nonprincipal ultrafilters follows by applying Zorn's Lemma to the class of filters on $\mathbb{N}$ which contain F.

Let A_i, $i \in \mathbb{N}$ be a family of algebraic structures (e.g. rings, valuation rings, local rings, algebras $k_i[X]$ where k_i is a field). Let D be an ultrafilter on $\mathbb{N}$. We form the ultraproduct $A^* = (\prod_{i \in \mathbb{N}} A_i)/D$ (or the ultrapower $A^{\mathbb{N}}/D$ if all the $A_i = A$) as follows. On the cartesian product $\prod_{i \in \mathbb{N}} A_i$ (an element a of the cartesian product is a sequence $i \mapsto a(i)$) define an equivalence relation $\sim$ by $a \sim a'$ if and only if $\{i \mid a(i) = a'(i)\} \in D$. We denote the equivalence class of a by $[a]$. We define $A^* = (\prod_{i \in \mathbb{N}} A_i)/D$ as the set of equivalence classes $[a]$ with the operations and relations on A^* defined componentwise e.g. $[a]+[b]=[a+b]$ and $[a]\cdot[b] = [a \cdot b]$, where $(a+b)(i) = a(i)+b(i)$ and $(a \cdot b)(i) = a(i) \cdot b(i)$.

The zero of A^* is the equivalence class of the constant sequence $0(i) = 0$ and the unity of A^* is the equivalence class of the constant sequence $1(i) = 1$. If each A_i has a total order relation $\leqq$ (e.g. $A_i = \mathbb{Z}$), then A^* has a total order relation $\leqq$ defined by $[a] \leqq [a'] \leftrightarrow \{i \mid a(i) \leqq a'(i)\} \in D$. If each A_i has a valuation ord_i: $A_i - \{0\} \to \mathbb{Z}$ then A^* has a valuation ord*: $A^* - \{0\} \to \mathbb{Z}^* = \mathbb{Z}^{\mathbb{N}}/D$ defined by $\mathrm{ord}^*([a]) = [\mathrm{ord}_i(a(i))]$. It is easy to see that these operations are well defined, using the fact that D is an ultrafilter. If all the $A_i = A$ then there is a natural embedding μ: $A \to A^*$ defined by $\mu(a) = [a(i)]$ where $a(i) = a$ for all i. In such cases we shall identify A with its image in A^*.

The fundamental property of ultraproducts is the following

Theorem (Łoś, cf. [7, Theorem 4.1.9, p. 170]). *Let L be a first order language, let A_i, $i \in I$ be structures for L and let D be an ultrafilter on I. Then for any $[a] \in (\prod_{i \in I} A_i)/D = A^*$ and any first order formula $\varphi(x)$, $\varphi([a])$ is true in A^* if and only if $\{i \mid \varphi(a(i))$ is true in $A_i\} \in D$.*

The proof of this is quite straightforward (cf. [7, Theorem 4.1.9, p. 170]) but we shall not give it. We shall indicate below the proofs in a few special cases which we shall use later.

(i) If k is a field and D is an ultrafilter on $\mathbb{N}$ then it is easy to see that $k^* = k^{\mathbb{N}}/D$ is a field. For example if $[a] \in k^*$ and $[a] \neq 0$ then there is an $s \in D$ such that $a(i) \neq 0$ for all $i \in s$. Define $b(i) = a(i)^{-1}$ if $i \in s$ and $b(i) = 0$ if $i \notin s$. Then for $i \in s$ $a(i) \cdot b(i) = 1$ and hence $[a] \cdot [b] = 1$.

Let R be a discrete valuation ring with valuation ord: $R - \{0\} \to \mathbb{Z}$, and prime p. Let $R^* = R^{\mathbb{N}}/D$ for some nonprincipal ultrafilter D. Let $\mathbb{Z}^* = \mathbb{Z}^{\mathbb{N}}/D$, $k = R/(p)$ and $k^* = k^{\mathbb{N}}/D$. Then R^* is a valuation ring with valuation ord* (defined as above) into $\mathbb{Z}^*$, with prime p, (we identify R with $\mu(R) \subset R^*$ defined as above) and residue class field k^* (i.e. $R^*/(p) = k^*$). If R is Henselian then so is R^*. $\mathbb{Z}$ is a convex subgroup of $\mathbb{Z}^*$ (under the embedding μ) i.e., if $[a] \in \mathbb{Z}^*$, m, $n \in \mathbb{Z}$ and

$m \leqq [a] \leqq n$ then $[a] \in \mathbb{Z}$ (more precisely $\mu(\mathbb{Z})$). Moreover $\mathbb{Z}$ is properly contained in $\mathbb{Z}^*$, since D is nonprincipal. All of this follows directly from the definitions. We shall prove some of these assertions:

One can see that R^* is a valuation ring as follows. Suppose $\operatorname{ord}^*([a]) \leqq \operatorname{ord}^*([b])$. Then, for some $s \in D$, $\operatorname{ord}(a(i)) \leqq \operatorname{ord}(b(i))$ for all $i \in s$. Hence for $i \in s$ there exists a $c_i \in R$ such that $a(i)\, c_i = b(i)$. Let $c(i) = c_i$ for $i \in s$ and $c(i) = 0$ for $i \in \mathbb{N} - s$. Then $a(i)\, c(i) = b(i)$ for $i \in s$, and $[a][c] = [b]$, since $s \in D$.

Suppose that R is Henselian (cf. [15, § 30, p. 103]). Let

$$f(x) = x^n + [a_{n-1}]\, x^{n-1} + \ldots + [a_0],$$

$$g_0(x) = x^k + [b_{k-1}]\, x^{k-1} + \ldots + [b_0] \quad \text{and}$$

$$h_0(x) = x^{n-k} + [c_{n-k-1}]\, x^{n-k-1} + \ldots + [c_0] \in R^*[x]$$

satisfy (1) $f(x) - g_0(x)\, h_0(x) \in p\, R^*[x]$ and (2) g_0 modulo p and h_0 modulo p have no common root (i.e. $g_0\, R^*[x] + h_0\, R^*[x] + p\, R^*[x] = R^*[x]$). Then, there is an $s \in D$ such that for all $i \in s$ we have $f_i(x) - g_{0i}(x)\, h_{0i}(x) \in p\, R[x]$ and $g_{0i}(x) \bmod p$ and $h_{0i}(x) \bmod p$ have no common root (where $f_i(x) = x^n + a_{n-1}(i)\, x^{n-1} + \ldots + a_0(i)$ etc.). Hence (since R is Henselian) for $i \in s$ there exist polynomials

$$g_i(x) = x^k + b'_{k-1}(i)\, x^{k-1} + \ldots + b'_0(i) \quad \text{and}$$

$$h_i(x) = c'_{n-k}(i)\, x^{n-k} + \ldots + c'_0(i) \in R[x]$$

such that $g_i(x) \equiv g_{0i}(x)$ and $h_i(x) \equiv h_{0i}(x) \bmod p$ and $f_i(x) = g_i(x)\, h_i(x)$. Let

$$g(x) = x^k + [b'_{k-1}]\, x^{k-1} + \ldots + [b'_0] \quad \text{and}$$

$$h(x) = [c'_{n-k}]\, x^{n-k} + \ldots + [c'_0] \in R^*[X].$$

Then $g(x) \equiv g_0(x)$, $h(x) \equiv h_0(x) \bmod p$ (in $R^*[x]$) and $f(x) = g(x)\, h(x)$ (since these statements are true for all $i \in s$). Hence R^* is Henselian.

To see that $\mathbb{Z}$ is a convex subgroup of $\mathbb{Z}^*$, suppose that $m \leqq [a] \leqq n$ for m, $n \in \mathbb{Z}$ and $[a] \in \mathbb{Z}^*$. Then, for some $s \in D$, $m \leqq a(i) \leqq n$ for all $i \in s$. Let $s_j = \{i \in s \mid a(i) = j\}$ for $m \leqq j \leqq n$. Then $s = \bigcup_{j=m}^{n} s_j$ and hence, for some j_0 ($m \leqq j_0 \leqq n$), $s_{j_0} \in D$. (For suppose not. Then $\mathbb{N} - s_j \in D$ for $j = m, \ldots, n$ and hence $\bigcap_{j=m}^{n} (\mathbb{N} - s_j) = \mathbb{N} - s \in D$). Thus $[a] = j_0 \in \mathbb{Z}$. Moreover let $[\gamma] \in \mathbb{Z}^*$ be the equivalence class of the sequence $\gamma(i) = i$. Then $[\gamma] \neq z$, for all $z \in \mathbb{Z}$, since D is nonprincipal.

(ii) For each $i \in \mathbb{N}$ let A_i be a quasi-local ring, i.e. a ring with only one maximal ideal $\mathcal{M}_i$. Let D be a nonprincipal ultrafilter on $\mathbb{N}$ and let $A^* = (\prod_{i \in \mathbb{N}} A_i)/D$. Then A^* is again a quasi-local ring with maximal ideal $\mathcal{M}$, defined by

$$[a] \in \mathcal{M} \leftrightarrow \{i \mid a(i) \in \mathcal{M}_i\} \in D.$$

If the residue class field of A_i is k_i, then the residue class field of A^* is $k^* = (\prod_{i \in \mathbb{N}} k_i)/D$. If the ideal $\mathcal{M}_i$ is generated by $x_0(i), \ldots, x_n(i)$ for every i, then $\mathcal{M}$ is

generated by $x_0, \ldots, x_n$ where x_j denotes the equivalence class of the sequence $x_j(i)$.

As above, if A is Henselian, then A^* is Henselian too.

(iii) Ultraproducts have certain saturation properties (cf. [7, § 6.1, p. 305]) of which we shall use only the following special case. Let A_i, $\mathcal{M}_i, D, A^*$ and $\mathcal{M}$ be as in (ii) above. Suppose $a_j \in A^*$ and $a_{j+1} \equiv a_j \bmod \mathcal{M}^j$ for all $j \in \mathbb{N}$. Then there is a $z \in A^*$ satisfying $z \equiv a_j \bmod \mathcal{M}^j$ for all $j \in \mathbb{N}$.

Proof. Let $a_j(i)$ be a representative of a_j. For each j there is an $s_j \in D$ such that $a_j(i) \equiv a_l(i) \bmod \mathcal{M}_i^l$ for $l = 0, \ldots, j$ and $i \in s_j$. For any $i \in \mathbb{N}$ let $j(i)$ be the largest integer in

$$\{j \mid j \leqq i \text{ and } i \in s_j\} \cup \{0\},$$

and set $z(i) = a_{j(i)}(i)$. Then $z(i) \equiv a_l(i) \bmod \mathcal{M}_i^l$ for $l = 0, 1, \ldots, j(i)$ and $i \in s_{j(i)}$. For all j and all i satisfying $j \leqq i$ and $i \in s_j$ we have $j \leqq j(i)$ and $i \in s_{j(i)}$, and hence $z(i) \equiv a_j(i) \bmod \mathcal{M}_i^j$. Thus, since D is nonprincipal, $z = [z(i)] \equiv a_j \bmod \mathcal{M}^j$, for all j.

A similar argument, combined with Łoś' theorem, proves that A^* is ω_1-saturated, i.e. a countable set Φ of first order formulas is satisfiable in A^*, provided that every finite subset of Φ is satisfiable in A^*. (We say that a set of first order formulas $\varphi_j(z)$ is satisfiable in A^* if there exists a $z \in A^*$ satisfying $\varphi_j(z)$ for all j.) We shall however not need this.

Remark. For our purposes it will be sufficient to use ultraproducts with respect to any nonprincipal ultrafilter. For other uses it can be necessary to choose the ultrafilter more carefully or to use the more abstract notion of enlargements (due to A. Robinson). For a discussion of enlargements see Chapter 1 of [14].

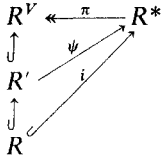
§ 2. In this section we shall give a new proof of a theorem of Greenberg (Theorem 2.1). Recall that a discrete valuation ring R is called *excellent* if the fraction field of the completion of R is separable over the fraction field of R. (Cf. [12 Chapter 10, § 6].)

Theorem 2.1 (Greenberg [10, Theorem 1]). *Let R be an excellent Henselian discrete valuation ring, with prime p. Let $Y = (Y_1, \ldots, Y_n)$ and $f = (f_1, \ldots, f_m) \in R[Y]$. Then there is an integer $N \in \mathbb{N}$, depending on f, such that for any nonzero $\alpha \in \mathbb{N}$ and any $\bar{y} \in R$ satisfying $f(\bar{y}) \equiv 0 \bmod p^{N\alpha}$, there is a $y \in R$ satisfying $f(y) = 0$ and $y \equiv \bar{y} \bmod p^{\alpha}$.*

Let D be a nonprincipal ultrafilter on $\mathbb{N}$ and let $R^* = R^{\mathbb{N}}/D$. Then (cf. § 1(i)) R^* is a Henselian valuation ring with valuation ord* into $\mathbb{Z}^* = \mathbb{Z}^{\mathbb{N}}/D$. Let $k = R/(p)$, then $R^*/(p) = k^* = k^{\mathbb{N}}/D$. Let V be any convex subgroup of $\mathbb{Z}^*$ containing $\mathbb{Z}$ (i.e. $0 \leqq \alpha \leqq \beta$, $\alpha \in \mathbb{Z}^*$, $\beta \in V$ imply that $\alpha \in V$). Set $I_V = \{x \in R^* \mid \text{ord}^*(x) \notin V\}$. Since V is convex I_V is a prime ideal of R^*. Set $R^V = R^*/I_V$. It is easy to see that R^V is a valuation ring with value group V and residue class field k^*. We denote the natural projection of R^* onto R^V by π. The natural embedding i of R in R^* induces an embedding of R into R^V, and we consider R as a subring of R^V. The algebraic part of our proof of Greenberg's Theorem is contained in Lemmas 2.2 and 2.3. We shall find it convenient to use Lemma 2.2, in the case $V = \mathbb{Z}$, in §§ 3 and 4. Lemma 2.2 is an easy logical consequence of Greenberg's Theorem 2.1

(see Remark (ii) at the end of § 2), and hence for our uses in §§ 3 and 4 we could have omitted the proofs below of 2.2 and 2.3.

Lemma 2.2. *Let R, R^*, V, R^V, π, i be as above. Let R' be any subring of R^V containing R and finitely generated over R. Then R' can be lifted into R^*, i.e. there is a homomorphism $\psi: R' \to R^*$ such that the following diagram commutes*



Proof. In Lemma 2.3 we shall show that the fraction field of R^V is separable over the fraction field of R. Hence we have $R' = R[y_1, \ldots, y_r, w_1, \ldots, w_s]$ with the y_i, $w_j \in R^V$, $y_1, \ldots, y_r$ algebraically independent over R and $w_1, \ldots, w_s$ separably algebraic over $R[y_1, \ldots, y_r]$. It is obvious that $R[y_1, \ldots, y_r]$ can be lifted to R^* – just send y_i to any element of $\pi^{-1}(y_i)$. Hence let R_0 be any subring of R^V containing R, let ψ_0 be a lifting of R_0 into R^* and let $w \in R^V$ be separably algebraic over R_0. It is sufficient to prove that ψ_0 can be extended to a lifting of $R_0[w]$ into R^*. Let f be an irreducible polynomial for w over R_0. Thus $f(w) = 0$ and $f'(w) \neq 0$ (where f' denotes the derivative of f). Choose $\bar{w} \in R^*$ such that $\pi(\bar{w}) = w$. Let g be the polynomial obtained by applying ψ_0 to the coefficients of f. Hence $\pi(g(\bar{w})) = 0$ and $\pi(g'(\bar{w})) \neq 0$, i.e. $\operatorname{ord}^* g(\bar{w}) \notin V$ and $\operatorname{ord}^*(g'(\bar{w})) \in V$. Since V is a convex subgroup of $\mathbb{Z}^*$ we obtain $\operatorname{ord}^*(g(\bar{w})/(g'(\bar{w}))^2) > 0$. Hence by the Hensel-Rychlik lemma for $g(\bar{w})$ (which follows from Hensel's lemma applied to the polynomial $G(h) = \dfrac{1}{g(\bar{w})} g\left(\bar{w} + \dfrac{g(\bar{w})}{g'(\bar{w})} h\right) \equiv 1 + h \bmod p$) there is a $\bar{\bar{w}} \in R^*$ such that $g(\bar{\bar{w}}) = 0$ and $\bar{\bar{w}} \equiv \bar{w} \bmod \dfrac{g(\bar{w})}{g'(\bar{w})}$. This congruence implies that $\pi(\bar{\bar{w}}) = w$. Hence by mapping w to $\bar{\bar{w}}$, ψ_0 extends to a lifting of $R_0[w]$ into R^*.

Remark. From the above proof it is easy to see that R^V can be lifted into R^* if the characteristic of R is zero.

Lemma 2.3. *Assume the hypotheses of Lemma* 2.2. *Then* $\operatorname{Frac}(R^V)$ *is separable over* $\operatorname{Frac}(R)$ *(*Frac *denotes the fraction field).*

Proof. Suppose that the characteristic of R is $p \neq 0$. Let $\bar{R}$ be the completion of R. The natural embedding of $R \hookrightarrow \bar{R}$ induces embeddings $R^* \hookrightarrow \bar{R}^*$ $(= \bar{R}^{\mathbb{N}}/D)$ and $R^V \hookrightarrow \bar{R}^V$. Since $\operatorname{Frac}(\bar{R})$ is separable over $\operatorname{Frac}(R)$ it is sufficient to prove that $\operatorname{Frac}(\bar{R}^V)$ is separable over $\operatorname{Frac}(\bar{R})$. Thus we may suppose that R is complete. Let $K = \operatorname{Frac}(R)$, $L = \operatorname{Frac}(R^V)$. Let K_1 be any finite field extension of K contained in $K^{1/p}$. We fix an extension of the valuation on L to LK_1. Let k_1 be the residue class field of K_1 and let Γ be the order group of K_1. We have (cf. [12, Ch. 12, § 4, Prop. 13]) that

$$[LK_1 : L] \geqq e(LK_1/L) f(LK_1/L)$$

where e is the ramification index and f the residue class degree. But

$$e(LK_1/L) \geqq [V+\Gamma : V] \geqq [\mathbb{Z}^*+\Gamma : \mathbb{Z}^*] = [\Gamma : \mathbb{Z}] = e(K_1/K) \tag{1}$$

$$f(LK_1/L) \geqq [k^* k_1 : k^*] = [k_1 : k] = f(K_1/K). \tag{2}$$

In (1) only $[\mathbb{Z}^*+\Gamma : \mathbb{Z}^*] = [\Gamma : \mathbb{Z}]$ requires proof. Let $\gamma \in \Gamma \cap \mathbb{Z}^*$. Set $e = e(K_1/K)$, then $e\gamma \in \mathbb{Z}$. Since $\mathbb{Z}$ is a convex subgroup of $\mathbb{Z}^*$ (cf. § 1(i)), we deduce $\gamma \in \mathbb{Z}$. This proves (1). In (2) only the equality $[k^* k_1 : k^*] = [k_1 : k]$ required proof. Notice that $k^* k_1 \subset k_1^* (= k_1^{\mathbb{N}}/D)$. Let $\alpha_1, \ldots, \alpha_r \in k_1$ be linearly independent over k. Suppose $\sum_j a_j \alpha_j = 0$ in k_1^*, with $a_j = [a_j(i)] \in k^*$. Thus for some $s \in D$, we have $\sum_j a_j(i) \alpha_j = 0$ in k_1, for all $i \in s$. Hence $a_j(i) = 0$ for all $i \in s$. Thus $a_j = 0$, and the α_j are linearly independent over k^*. This proves (2).

Since R is complete we have (cf. [12, Ch. 12, § 6, Prop. 18]) that $[K_1 : K] = e(K_1/K) f(K_1/K)$. Hence $(LK_1 : L) \geqq [K_1 : K]$. Thus L is linearly disjoint from $K^{1/p}$ over K and hence L is separable over K (cf. [12, Ch. 10, § 6]).

Proof of Theorem 2.1. Suppose the theorem is false. Thus $\forall N \in \mathbb{N}$, $\exists \alpha \in \mathbb{N}$, $\exists x \in R$ [$\alpha \neq 0$ and $\operatorname{ord}(f(x)) \geqq N\alpha$ and $\neg\, \exists y \in R (f(y) = 0$ and $\operatorname{ord}(y-x) \geqq \alpha)$], where $\neg$ denotes "not". This gives sequences $\alpha_N \in \mathbb{N}$, $x_N \in R$ for $N = 0, 1, 2, \ldots$ which determine elements $\alpha \in \mathbb{N}^*$, $x \in R^*$ ($\alpha = [(\alpha_N)]$, $x = [(x_N)]$) satisfying $\alpha \neq 0$ and

$$\operatorname{ord}^*(f(x)) \geqq m\alpha \quad \text{for all } m \in \mathbb{N} \tag{3}$$

$$\neg\, \exists y \in R^* (f(y) = 0 \quad \text{and} \quad \operatorname{ord}^*(y-x) \geqq \alpha). \tag{4}$$

Inequality (3) follows from the definitions, since D contains every cofinite subset of $\mathbb{N}$. Also (4) is easy to see directly from the definition of R^*, since if there were a $y \in R^*$ satisfying $f(y) = 0$ and $\operatorname{ord}^*(y-x) \geqq \alpha$ then for infinitely many $i \in \mathbb{N}$ we would have $f(y(i)) = 0$ and $\operatorname{ord}(y(i) - x_i) \geqq \alpha_i$ (with $y = [y(i)]$). But there is no such $y(i)$ by the definition of α_i, x_i. Set $V = \{\beta \in \mathbb{Z}^* \mid -m\alpha \leqq \beta \leqq m\alpha$ for some $m \in \mathbb{N}\}$. Then V is a convex subgroup of $\mathbb{Z}^*$. Let $\pi\colon R^* \twoheadrightarrow R^V$ be as above. From (3) it follows that $f(\pi(x)) = \pi f(x) = 0$. Set $R' = R[\pi(x)] \subset R^V$. By Lemma 2.2 there is a lifting $\psi\colon R' \to R^*$. Set $y = \psi(\pi(x))$. Then $f(y) = 0$ and $\pi(y) = \pi(x)$ and hence $\operatorname{ord}^*(y - x) > \alpha$. This contradicts (4) and completes the proof.

Remarks. (i) Using similar methods one easily obtains a new proof of the result of Greenberg [11], (cf. van den Dries [20, p. 146]). Effectivity can be proved by the methods of Sect. 6. For an explicit algorithm see Birch and McCann [6].

(ii) Lemma 2.2 is an easy logical consequence of Greenberg's theorem 2.1.

Proof. Let $R' = R[\pi(\bar{y})] \subset R^V$, with $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_n) \in R^*$. Let $f = (f_1, \ldots, f_m) \in R[Y]$ generate the ideal of all $f \in R[Y]$ vanishing on $\pi(\bar{y})$. To prove Lemma 2.2 it suffices to find a $y \in R^*$ such that $f(y) = 0$ and $\pi(y) = \pi(\bar{y})$. Let $N \in \mathbb{N}$ be as in Greenberg's theorem 2.1. There is a positive α in $\mathbb{Z}^*$ such that $\alpha \notin V$ and $\operatorname{ord}^* f(\bar{y}) \geqq N\alpha$. Indeed, since $f(\pi(\bar{y})) = 0$, we have $\operatorname{ord}^* f(\bar{y}) \notin V$, and moreover we can write $\operatorname{ord}^* f(\bar{y}) = N\alpha + r$, with $0 \leqq r < N$. Thus by Greenberg's theorem 2.1 and the definition of ultraproducts (or Łoś' Theorem), there is a $y \in R^*$ satisfying $f(y) = 0$ and $\operatorname{ord}^*(y - \bar{y}) \geqq \alpha$. And since $\alpha \notin V$, this implies $\pi(y) = \pi(\bar{y})$.

§ 3. Artin [2] has proved the following theorems (3.1 and 3.2). Related theorems are proved in [1].

Theorem 3.1 ([2, Thm. 1.10]). *Let R be a field or an excellent discrete valuation ring with prime p. Let A be the henselization of $R[X_1, \ldots, X_n]$ at the maximal ideal $\mathcal{M} = (p, X_1, \ldots, X_n)$ (cf. [15, § 43, p. 180]). Given an arbitrary system of polynomial equations $f(Y) = 0$, $Y = (Y_1, \ldots, Y_N)$ with coefficients from A, a solution $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_N)$ in the $\mathcal{M}$-adic completion $\hat{A}$ of A and an integer c, there exists a solution $y = (y_1, \ldots, y_N) \in A$ with $y_i \equiv \bar{y}_i \bmod \mathcal{M}^c$ for $i = 1, \ldots, N$.*

Theorem 3.2 ([2, Thm. 6.1]). *There is an integer valued function $\beta = \beta(n, N, d, \alpha)$ defined for non-negative values of n, N, d, α with the following property: Let k be any field, let $f = (f_1, \ldots, f_m)$ be polynomials in $k[X, Y]$ where $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_N)$ with $\deg(f) \leqq d$. Suppose given polynomials $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_N) \in k[X]$ such that $f(X, \bar{y}) \equiv 0 \bmod (X)^\beta$. Then there are elements $y = (y_1, \ldots, y_N) \in k[X]^\sim$ (the Henselization of $k[X]$ at the maximal ideal (X)) solving the equations $f(X, Y) = 0$ and such that $y \equiv \bar{y} \bmod (X)^\alpha$.*

D. Popescu has extended the methods of Artin to prove the following theorem. (See also [17].)

Theorem 3.3 ([18, Thm. 1.4]). *Let R be an excellent Henselian discrete valuation ring with prime p and let $f = (f_1, \ldots, f_m)$ be polynomials in $R[X, Y]$ where $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_N)$. There is an integer valued function $v(\alpha)$, depending on f, such that, given polynomials $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_N) \in R[X]$ satisfying $f(X, \bar{y}) \equiv 0 \bmod (p, X)^{v(\alpha)}$, there are elements $y = (y_1, \ldots, y_N) \in R[X]^\sim$ (the Henselization of $R[X]$ at the maximal ideal (p, X)) solving the equations $f(X, Y) = 0$ and such that $y \equiv \bar{y} \bmod (p, X)^\alpha$.*

Theorem 3.2 is proved in [2, § 6] by a careful analysis of the proof of Theorem 3.1. The proof of Theorem 3.3 given in [18] is very intricate and entails a substantial extension of the Néron p-desingularization. In this section we show that Theorems 3.2 and 3.3 can be obtained as quite easy logical consequences of Theorem 3.1.

Proof of Theorem 3.2. Suppose that Theorem 3.2 is false. Then there exist $n, N, d, \alpha \in \mathbb{N}$ such that for each $j \in \mathbb{N}$ there is a field k_j and polynomials $f_j = (f_{j1}, \ldots, f_{jm_j}) \in k_j[X, Y]$ of degree $\leqq d$ ($X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_N)$) and $\bar{y}_j = (\bar{y}_{j1}, \ldots, \bar{y}_{jN}) \in k_j[X]$ satisfying $f_j(X, \bar{y}_j) \equiv 0 \bmod (X)^j$ but there is no $y_j = (y_{j1}, \ldots, y_{jN}) \in k_j[X]^\sim$ satisfying $f_j(X, y_j) = 0$ and $y_j \equiv \bar{y}_j \bmod (X)^\alpha$. Since the number of linearly independent polynomials over k_j in X and Y of degree $\leqq d$ is uniformly bounded in terms of n, N and d, the m_j can be taken to be bounded and hence without loss of generality we may assume that $m_j = m$ $\forall_j \in \mathbb{N}$. Let D be a nonprincipal ultrafilter on $\mathbb{N}$ and let $A^* = (\prod_{j \in \mathbb{N}} k_j[X]^\sim)/D$, $k^* = (\prod_{j \in \mathbb{N}} k_j)/D$. We consider $k^*[X]$ as a subring of A^*, identifying X with $[(X, X, \ldots, X, \ldots)]$. Let $\bar{y}_i(j) = \bar{y}_{ji} \in k_j[X]$ for $j \in \mathbb{N}$, let $\bar{\bar{y}}_i = [\bar{y}_i(j)_{j \in \mathbb{N}}] \in A^*$ for $i = 1, \ldots, N$ and let $\bar{\bar{y}} = (\bar{\bar{y}}_1, \ldots, \bar{\bar{y}}_N)$. Similarly let f be the element of $k^*[X, Y]$ which corresponds to the sequence $f_j \in k_j[X, Y]$. (The coefficients of f correspond to the sequences of coefficients of

the f_j). Notice that f has degree $\leqq d$. Since D contains all cofinite subsets of $\mathbb{N}$ we have $f(X,\bar{y})\equiv 0 \bmod (X)^j$ for all $j\in\mathbb{N}$ but there is no $y\in A^*$ such that $f(X,y)=0$ and $y\equiv\bar{y} \bmod (X)^\alpha$. (This follows immediately from the definition of the ultraproduct A^* as in our proof of Theorem 2.1). Let $(X)^\infty=\bigcap_{i\in\mathbb{N}}(X)^i\subset A^*$, let $A_1=A^*/(X)^\infty$, and let π be the projection of A^* onto A_1. Since π is injective on $k^*[X]\subset A^*$, we consider $k^*[X]$ as a subring of A_1. We need the following

Lemma 3.4. *With the above notation, $A_1\cong k^*[\![X]\!]$ (as $k^*[X]$ algebras).*

Proof. We define a map λ: $A^*\to k^*[\![X]\!]$ as follows. Let $a=[a(j)]\in A^*$, $a(j)\in k_j[X]^\sim\subset k_j[\![X]\!]$. Write $a(j)=\sum_h a_h(j)\,X^h$, where h is a multi-index and $a_h(j)\in k_j$. Set $\lambda(a)=\sum_h [a_h(j)]X^h\in k^*[\![X]\!]$. It is easy to verify that λ: $a\mapsto\lambda(a)$ is a homomorphism with kernel $(X)^\infty$. Moreover λ maps $(\prod_{j\in\mathbb{N}} k_j[X])/D\subset A^*$ onto $k^*[\![X]\!]$. Indeed let $b=\sum_h [b_h(j)]\,X^h\in k^*[\![X]\!]$ with $b_h(j)\in k_j$. Set $a(j)=\sum_{h\leqq j} b_h(j)\,X^h\in k_j[X]$, then $\lambda([a(j)])=b$. This finishes the proof of Lemma 3.4.

From now on we identify A_1 with $k^*[\![X]\!]$. We have that $f(X,\pi(\bar{y}))=\pi f(X,\bar{y})=0$ since $f(X,\bar{y})\equiv 0 \bmod (X)^j$ for all $j\in\mathbb{N}$. Hence by Theorem 3.1 there is a $y\in k^*[X]^\sim$ such that $f(X,y)=0$ and $y\equiv\pi(\bar{y}) \bmod (X)^\alpha$. Since A^* is Henselian (cf. § 1(ii)) there is a $k^*[X]$-homomorphism ψ: $k^*[X]^\sim\to A^*$ (cf. [15, Thm. 43.5, p. 181]). By the uniqueness of the Henselization, $\pi\circ\psi$ is the identity on $k^*[X]^\sim$. Thus we have $f(X,\psi(y))=0$ (in A^*) and $\psi(y)\equiv\bar{y} \bmod (X)^\alpha$, since $\pi\psi(y)=y\equiv\pi(\bar{y}) \bmod (X)^\alpha$. But we know there is no such $\psi(y)\in A^*$. This contradiction proves Theorem 3.2.

Remarks. (i) Using similar methods one can also show that the degree of $y_1,\ldots,y_N$ over $k[X]$ can be bounded by a function of n, N, d and α. (If $P(X,Y)$ is the minimal polynomial of $y\in k[X]^\sim$ over $k[X]$ we define the degree of y to be the total degree of $P(X,Y)$.)

(ii) The results of Hermann and Stolzenberg (i.e. [2, Thm. 6.5]) used in Artin's proof of Theorem 3.2 can also be established quite efficiently using model theoretic techniques. (See [20] and the references there in.)

Proof of Theorem 3.3 Suppose that Theorem 3.3 is false for the excellent Henselian discrete valuation ring R. Then there exist $f=(f_1,\ldots,f_m)\in R[X,Y]$ and $\alpha\in\mathbb{N}$ such that for each $j\in\mathbb{N}$ there is a $\bar{y}_j\in R[X]$ satisfying $f(X,\bar{y}_j)\equiv 0 \bmod (p,X)^j$, but there is no $y\in R[X]^\sim$ satisfying $f(X,y)=0$ and $y\equiv\bar{y}_j \bmod (p,X)^\alpha$ ($X=(X_1,\ldots,X_n)$, $Y=(Y_1,\ldots,Y_N)$). Again let D be any nonprincipal ultrafilter on $\mathbb{N}$, let $A^*=(R[X]^\sim)^{\mathbb{N}}/D$ and $R^*=R^{\mathbb{N}}/D$. Notice that $R[X]\subset A^*$, and $R^*\subset A^*$.

Let $\bar{y}=(\bar{y}_1,\ldots,\bar{y}_N)\in A^*$ where $\bar{y}_i$ corresponds to the sequence $\{\bar{y}_{ji}\}_{j\in\mathbb{N}}$. Then as above $f(X,\bar{y})\equiv 0 \bmod (p,X)^j$ for all $j\in\mathbb{N}$, but there is no $y\in A^*$ such that $f(X,y)=0$ and $y\equiv\bar{y} \bmod (p,X)^\alpha$. Let $\mathfrak{p}^\infty=\bigcap_{j\in\mathbb{N}}(p)^j\subset R^*$, $\mathfrak{M}^\infty=\bigcap_{j\in\mathbb{N}}(p,X)^j\subset A^*$. Let $R_1=R^*/\mathfrak{p}^\infty$ and $A_1=A^*/\mathfrak{M}^\infty$. We denote both the projection of A^* onto A_1 and the projection of R^* onto R_1 by π. Notice that $R\subset R_1$ and $R[X]\subset A_1$. We need the following

Lemma 3.5. *With the above notation, R_1 is a complete discrete valuation ring and $A_1 \cong R_1[\![X]\!]$ (as $R[X]$ algebras).*

Proof. R_1 is a discrete valuation ring (cf. § 2), moreover R_1 is complete by § 1(iii). We define a map λ: $A^* \to R_1[\![X]\!]$ as follows. Let $a=[a(j)] \in A^*$, $a(j) \in R[X]^\sim \subset R[\![X]\!]$. Write $a(j)=\sum_h a_h(j) X^h$, where h is a multi-index and $a_h(j) \in R$. Set $\lambda(a)=\sum_h \pi([a_h(j)]) X^h \in R_1[\![X]\!]$. As in the proof of Lemma 3.4, it is easy to verify that λ: $a \mapsto \lambda(a)$ is a surjective homomorphism with kernel $\mathcal{M}^\infty$. This proves Lemma 3.5.

From now on we identify A_1 with $R_1[\![X]\!]$. We have $f(X, \pi(\bar{y}))=\pi f(X, \bar{y})=0$ (in A_1). Since R_1 is complete it is excellent and hence by Theorem 3.1 there is a $y \in R_1[X]^\sim$ such that $f(X, y)=0$ and $y \equiv \pi(\bar{y}) \bmod (p, X)^\alpha$. For rings R_2 such that $R \subseteq R_2 \subseteq R_1$, let $R_2[X]^\sim$ denote the intersection of all the Henselian rings dominated by $R_1[\![X]\!]$ which contain $R_2[X]$. (If $R_2[X]$ is normal, then its Henselization is isomorphic with $R_2[X]^\sim$, see [15, Thm. 43.5, p. 181].) For some $R_2 \subseteq R_1$ which is finitely generated over R we have that $y \in R_2[X]^\sim$. By Lemma 2.2 (with $V=\mathbb{Z}$) there is a lifting of R_2 into $R^* \subset A^*$. This lifting extends naturally to a lifting ψ of $R_2[X]$ into A^* (i.e. ψ is the identity on $R[X]$ and $\pi \circ \psi$ is the identity on $R_2[X]$). Since A^* is Henselian (cf. § 1(ii)), ψ extends to a lifting ψ of $R_2[X]^\sim$ into A^* (cf. [15, Thm. 43.5, p. 181]. Indeed ψ extends to the derived normal ring of $R_2[X]$).

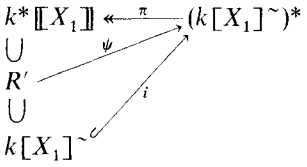
Hence $f(X, \psi(y))=0$ in A^* and $\psi(y) \equiv \bar{y} \bmod (p, X)^\alpha$, since $\pi\psi(y)=y \equiv \pi(\bar{y}) \bmod (p, X)^\alpha$. But we know (above) that no such $\psi(y) \in A^*$ exists. This contradiction proves Theorem 3.3.

§ 4. In this section we prove some extensions of Theorems 3.1, 3.2 and 3.3. Theorem 4.1 is due to M. Artin (oral communication) and can also be proved quite easily without using ultrapowers. It seems that a proof of Theorem 4.2 using conventional methods would be rather difficult.

Theorem 4.1. *Let k be a field, $X=(X_1, \ldots, X_n)$ and let $A=k[X]^\sim$ (the henselization of $k[X]$ at the maximal ideal (X)). Let $Y=(Y_1, \ldots, Y_N)$, $f=(f_1, \ldots, f_m) \in k[X, Y]$, $\bar{y}=(\bar{y}_1, \ldots, \bar{y}_N) \in k[\![X]\!]$ with $\bar{y}_1, \ldots, \bar{y}_r \in k[\![X_1]\!]$. Suppose that $f(X, \bar{y})=0$, and let $\alpha \in \mathbb{N}$. Then there exists $y=(y_1, \ldots, y_N) \in A$ with $y_1, \ldots, y_r \in k[X_1]^\sim$ such that $f(X, y)=0$ and $y_i \equiv \bar{y}_i \bmod (X)^\alpha$ for $i=1, \ldots, N$.*

Proof. First notice that the theorem as stated follows immediately from the special case where $\alpha=0$, i.e., the congruence condition $y \equiv \bar{y} \bmod (X)^\alpha$ falls away. Indeed the congruence condition can be replaced by polynomial equations.

Let $\bar{y}$ be given. Freeze $\bar{y}_1, \ldots, \bar{y}_r \in k[\![X_1]\!]$. Then by Theorem 3.1 for $R=k[\![X_1]\!]$ we know that there exist $y_{r+1}, \ldots, y_N \in k[\![X_1]\!]\,[X_2, \ldots, X_n]^\sim$ such that $f(X, \bar{y}_1, \ldots, \bar{y}_r, y_{r+1}, \ldots, y_N)=0$. Choose a subring R' of $k[\![X_1]\!]$ which is finitely generated over $k[X_1]^\sim$ and such that $\bar{y}_1, \ldots, \bar{y}_r \in R'$ and $y_{r+1}, \ldots, y_N \in R'[X_2, \ldots, X_n]^\sim$ (the intersection of all the henselian rings dominated by $k[\![X]\!]$ and containing $R'[X_2, \ldots, X_n]$). Now by Lemma 2.2 (with $R=k[X_1]^\sim$ and $V=\mathbb{Z}$) there is a lifting ψ such that the following diagram commutes, where * denotes the ultrapower with respect to a nonprincipal ultrafilter D.



(i is the natural embedding and π is the projection of $(k[X_1]^\sim)^*$ onto $k^*[\![X_1]\!]$ $=(k[X_1]^\sim)^*/(X_1)^\infty$ where $(X_1)^\infty=\bigcap_{i\in N}(X_1)^i$ as in the proofs of Theorems 3.2 and 3.3.) Moreover ψ extends naturally to a lifting ψ of $R'[X_2,\ldots,X_n]^\sim$ into $(k[X]^\sim)^*$ (cf. proof of Theorem 3.3). Since $\bar{y}_1,\ldots,\bar{y}_r\in R'$ we have that $\psi(\bar{y}_1),\ldots,\psi(\bar{y}_r)\in(k[X_1]^\sim)^*$. Also

$$f(X,\psi(\bar{y}_1),\ldots,\psi(\bar{y}_r),\psi(y_{r+1}),\ldots,\psi(y_N))=\psi f(X,\bar{y}_1,\ldots,\bar{y}_r,y_{r+1},\ldots,y_N)=0$$

(in $(k[X]^\sim)^*$). The theorem (for $\alpha=0$) follows now from the definition of ultrapower.

Theorem 4.2. *Using the notation of Theorem 3.2, there is an integer valued function* $\beta'=\beta'(n,N,d,\alpha)$ *such that if* k *is any field and* $f=(f_1,\ldots,f_m)\in k[X,Y]$ $(X=(X_1,\ldots,X_n)$ *and* $Y=(Y_1,\ldots,Y_N))$ *have degrees* $\leqq d$ *and there exists a* $\bar{y}=(\bar{y}_1,\ldots,\bar{y}_N)\in k[X]$ *with* $\bar{y}_1,\ldots,\bar{y}_r\in k[X_1]$ *satisfying* $f(X,\bar{y})\equiv 0 \bmod (X)^{\beta'}$, *then there exists* $y=(y_1,\ldots,y_N)\in k[X]^\sim$ *with* $y_1,\ldots,y_r\in k[X_1]^\sim$ *satisfying* $f(X,y)=0$ *and* $y_i\equiv\bar{y}_i \bmod (X)^\alpha$ *for* $i=1,\ldots,N$.

Proof. The proof of Theorem 4.2 is identical with our proof of Theorem 3.2 above except that one uses Theorem 4.1 in place of Theorem 3.1.

Theorem 4.3. *Using the notation of Theorem 3.3, there exists an integer valued function* $\nu'(\alpha)$, *depending on* f, *such that given polynomials* $\bar{y}=(\bar{y}_1,\ldots,\bar{y}_N)\in R[X]$ *with* $\bar{y}_1,\ldots,\bar{y}_r\in R$ *such that* $f(X,\bar{y})\equiv 0 \bmod (p,X)^{\nu'(\alpha)}$, *there are elements* $y=(y_1,\ldots,y_N)\in R[X]^\sim$, *with* $y_1,\ldots,y_r\in R$, *solving the equations* $f(X,Y)=0$ *and such that* $y\equiv\bar{y} \bmod (p,X)^\alpha$.

Proof. We use the notation of our proof of Theorem 3.3. Suppose 4.3 is false, then there is $\bar{y}=(\bar{y}_1,\ldots,\bar{y}_N)\in A^*$ with $\bar{y}_1,\ldots,\bar{y}_r\in R^*$ such that $f(X,\bar{y})\equiv 0$ $\bmod (p,X)^j$ for all $j\in\mathbb{N}$, and such that there is no $y\in A^*$, with $y_1,\ldots,y_r\in R^*$, satisfying $f(X,y)=0$ and $y\equiv\bar{y} \bmod (p,X)^\alpha$. We have $f(X,\pi(\bar{y}))=0$ (in A_1). By Theorem 3.1 (for $R_1[X]^\sim$ and the variables $Y_{r+1},\ldots,Y_N$) there are $y_{r+1},\ldots,y_N$ $\in R_1[X]^\sim$ such that $f(X,\pi(\bar{y}_1),\ldots,\pi(\bar{y}_r),\ y_{r+1},\ldots,y_N)=0$ and $y_i\equiv\pi(\bar{y}_i)$ $\bmod (p,X)^\alpha$, $i=r+1,\ldots,N$. Choose a ring $R_2\subset R_1$, which is finitely generated over R and such that $\pi(\bar{y}_1),\ldots,\pi(\bar{y}_r)\in R_2$ and $y_{r+1},\ldots,y_N\in R_2[X]^\sim$. Now proceed as in our proof of 3.3.

Remark. The implication "theorem of type 3.1 for fields implies theorem of type 3.2" is true for any constraints on the y_i, for example:

$$\bar{y}_1,\ldots,\bar{y}_{r_1}\in k[\![X_1]\!],\ \bar{y}_{r_1+1},\ldots,\bar{y}_{r_2}\in k[\![X_1,X_2]\!],\ldots,$$

$$\bar{y}_{r_{s-1}+1},\ldots,\bar{y}_{r_s}\in k[\![X_1,\ldots,X_s]\!]$$

and similar constraints on the y_i. The same proofs works.

§ 5. In this section we give some counterexamples to problems related to § 4. In [4], using the methods of [5], it was shown that there is a polynomial $f(X_1, X_2, Y_1, \ldots, Y_4) \in \mathbb{C}[X, Y]$ such that there exist $\bar{y}_1 \in \mathbb{C}[\![X_1]\!]$, $\bar{y}_2 \in \mathbb{C}[\![X_2]\!]$ and $\bar{y}_3, \bar{y}_4 \in \mathbb{C}[\![X_1, X_2]\!]$ satisfying $f(X_1, X_2, \bar{y}_1, \ldots, \bar{y}_4) = 0$ but there exist no $y_1 \in \mathbb{C}\{X_1\}$, $y_2 \in \mathbb{C}\{X_2\}$, and $y_3, y_4 \in \mathbb{C}\{X_1, X_2\}$ (the convergent power series rings) such that $f(X_1, X_2, y_1, \ldots, y_4) = 0$.

The key idea of [5] is to notice that one can define composition of power series as follows: Let $h(X_1)$, $g(X_2)$, $\sigma(X_2)$ be fixed. (The notation $h(X_1)$ means that h depends only on X_1.) Then

$$h(\sigma(X_2)) = g(X_2) \quad \text{iff } \exists S(X_1, X_2)[h(X_1) = g(X_2) + S(X_1, X_2)(X_1 - \sigma(X_2))].$$

One direction is immediate by substitution, the other follows by Taylor's theorem.

It is easy to write down a functional equation which only has a divergent solution e.g. $f(X + X^2) = 2f(X) + X$. The example in [4] is obtained by applying the above observation to translate such a functional equation into a system of polynomial equations $f(X, Y) = 0$ with the constraints that some of the Y's depend only on X_1 and some only on X_2. This shows that the statement of the form 3.1 with such constraints on the Y's is false. We shall show that some weaker assertions are also false.

(i) Let k be a field and consider the equation

$$f(cX) - f(X) = (1 - X)^{-1} - 1, \quad c \in k,\ c \neq 0.$$

This equation has only the solutions $f(X) = a_0 + \sum_{n=1}^{\infty} \frac{X^n}{c^n - 1}$, $a_0 \in k$. Hence if we let k be the algebraic closure of the field $\mathbb{F}_p$ with p elements, then this equation (with unknowns X, c) will have no solution in $k[\![X]\!]$ with $c \in k$, $c \neq 0$, but it will have solutions mod X^j for all $j \in \mathbb{N}$. The above device can be used to translate this equation into a system (1) of polynomial equations over $k[X_1, X_2]$ with constraints that some of the Y's depend only on X_1 and some only on X_2, which have a solution mod$(X_1, X_2)^j$ for all $j \in \mathbb{N}$ but have no solution in $k[\![X_1, X_2]\!]$:

$$\begin{aligned}
&Y_2 = Y_1 + Y_4(X_1 - X_2)\\
&Y_3 = Y_1 + Y_5(Y_6 X_1 - X_2)\\
&(Y_3 - Y_2)(1 - X_1) = 1 - (1 - X_1)\\
&Y_6\, Y_7 = 1\\
&Y_1,\ Y_6 \in k[\![X_2]\!]\\
&Y_2,\ Y_3,\ Y_6 \in k[\![X_1]\!].
\end{aligned} \tag{1}$$

($Y_6 \in k[\![X_1]\!] \cap k[\![X_2]\!] = k$ stands for c; $Y_6 \neq 0$, since $Y_6\, Y_7 = 1$; $Y_1 = f(X_2)$, $Y_2 = f(X_1)$ and $Y_3 = f(cX_1)$.)

(ii) In this example we give a system (3)–(7) of equations and constraints which has no solution in $\mathbb{Q}[\![X]\!]$ but has a solution in $\mathbb{Q}[\![X]\!]/(X)^j$ for all $j \in \mathbb{N}$.

Let k be a field. We consider the differential equation

$$c^2 X_1 \frac{\partial f}{\partial X_1}(X_1, X_2) - X_2 \frac{\partial f}{\partial X_2}(X_1, X_2) = \sum_{i,j \geqq 1} X_1^i X_2^j = [(1-X_1)^{-1}-1][(1-X_2)^{-1}-1]. \tag{2}$$

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

$$f(X_1, X_2) = a_0 + \sum_{i,j \geqq 1} \frac{X_1^i X_2^j}{c^2 i - j}, \quad a_0 \in k.$$

(This is easy to verify.) We use the above device to translate this into a system of equations (3)–(7) over the ring $\mathbb{Q}[X_1, \ldots, X_5]$ with some of the variables depending on only X_1 and X_2 and some on only X_3, X_4, X_5:

$$Y_1(X_1, X_2) \equiv Y_2(X_3, X_4, X_5) \bmod (X_1 - X_3, X_2 - X_4) \tag{3}$$

$$Y_2(X_3, X_4, X_5) \equiv Y_1(X_1, X_2) + X_5 Y_5(X_1, X_2) + X_5^2 Y_6 \bmod (X_3 - X_1 - X_5, X_4 - X_2) \tag{4}$$

$$Y_2(X_3, X_4, X_5) \equiv Y_1(X_1, X_2) + X_5 Y_7(X_1, X_2) + X_5^2 Y_8 \bmod (X_3 - X_1, X_4 - X_2 - X_5) \tag{5}$$

$$Y_9 \in k \quad \text{(i.e. } Y_9(X_1, X_2) = Y_{10}(X_3, X_4, X_5)\text{)} \quad \text{and} \quad Y_9 Y_{11} = 1 \tag{6}$$

$$Y_9^2 X_1 Y_5 - X_2 Y_7 = [(1-X_1)^{-1}-1][(1-X_2)^{-1}-1]. \tag{7}$$

(The variables Y_6 and Y_8 depend on all the X's.) Suppose (3)–(7) are satisfied. Set $f(X_1, X_2) = Y_1(X_1, X_2)$.

Set $X_1 = X_3$, $X_2 = X_4$ in (3) to conclude that $Y_2(X_3, X_4, X_5) = f(X_3, X_4)$. Setting $X_3 = X_1 + X_5$ and $X_4 = X_2$ in (4) we get that $f(X_1 + X_5, X_2) = f(X_1, X_2) + X_5 Y_5(X_1, X_2) + X_5^2 Y_6$. From this it follows that $Y_5 = \frac{\partial f}{\partial X_1}(X_1, X_2)$. Similarly $Y_7 = \frac{\partial f}{\partial X_2}(X_1, X_2)$. Then Eq. (7) is just Eq. (2), with $c = Y_9$. Finally (6) guarantees that $c \neq 0$ and $c \in k$. Thus it is clear that this system (3)–(7) has no solution over $\mathbb{Q}[\![X]\!]$ but has a solution in $Q[\![X]\!]/(X)^j$ for all $j \in \mathbb{N}$. Notice that the system has a solution in $k[\![X]\!]$ if $k \supsetneqq \mathbb{Q}$.

(iii) There is no $v \in \mathbb{N}$ having the following property: if the system (3)–(7) has a solution $\bar{y}$ in $\mathbb{C}[\![X]\!]/(X)^v$, then the system has a solution y in $\mathbb{C}[\![X]\!]$ with $y \equiv \bar{y} \bmod (X)$. On the other hand it is well-known (cf. introduction) that any system (with any constraints), which has a solution in $\mathbb{C}[\![X]\!]/(X)^j$ for all $j \in \mathbb{N}$, has a solution in $\mathbb{C}[\![X]\!]$. It is easy to see that there is not such a v. Indeed for every v there is a $c_v \in \mathbb{Q}$ such that the system (3)–(7) has a solution $\bmod (X)^v$ with $Y_9 = c_v$ but there is no solution in $\mathbb{C}[\![X]\!]$ with $Y_9 \equiv c_v \bmod (X)$ ($\Rightarrow Y_9 = c_v$).

Similar remarks are true for algebraically closed fields of infinite transcendence degree over $\mathbb{F}_p$ (cf. (i)).

Remark. A more interesting case is that of the example in the remark after Theorem 4.3. We don't know what happens in this case.

§ 6. In this section we prove the computability of the function β of Theorem 3.2. The computability of β follows by examining Artin's proof of Theorem 3.2 (i.e. Thm. 6.1 of [2]) (cf. [5] and [13]). Although our proof of Theorem 3.2 is not at all effective one can still prove the computability of β using some concepts from logic, without analyzing the proof of 3.2 but using only the existence of β. We shall assume that the reader is familiar with the concepts of a first order language and a first order theory. We shall also use the Gödel completeness theorem for first order logic which says that if T is a first order theory and φ a first order statement (in the language of T) then φ is provable from T if and only if φ is true in every model of T.

Theorem 6.1. *There is a computable function β which satisfies the condition of Theorem* 3.2.

Proof. We shall describe a first order theory T below by describing which structures are models of T. It will be apparent from our description what the appropriate first order language and axioms for T are. We shall not make these explicit. Fix n, N, d, α (as in Theorem 3.2). A model for T will be a local ring A, containing a field k, with maximal ideal $\mathcal{M}$ generated by $X_1, \ldots, X_n$ satisfying: (i) The X_i are algebraically independent over k (this will be an infinite set of axioms – one for each degree). (ii) $A/\mathcal{M}^i = k[X_1, \ldots, X_n]/(X_1, \ldots, X_n)^i$ for each $i = 1, 2, \ldots$. (iii) A is henselian.

Let Φ_β be the first order statement which says that the integer β satisfies the conditions of Theorem 3.2 for A, i.e. if $f(X, Y)$ is a system of polynomials of degrees $\leqq d$ in $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_N)$ over k and $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_N) \in k[X]$ (of degrees $\leqq \beta$) satisfy $f(X, \bar{y}) \equiv 0 \bmod \mathcal{M}^\beta$, then there is a $y = (y_1, \ldots, y_N) \in A$ such that $f(X, y) = 0$ and $y \equiv \bar{y} \bmod \mathcal{M}^\alpha$. Let $\beta_0 = \beta(n, N, d, \alpha)$. (By this we mean the value shown to exist in Theorem 3.2). We shall show that if A is any model of T then Φ_{β_0} is true in A. Let A be a model of T and let f be any system of polynomials of degrees $\leqq d$ in X and Y over k_A (the field in A). Let $\bar{y} \in k_A[X]$, and suppose that $f(X, \bar{y}) \equiv 0 \bmod \mathcal{M}^{\beta_0}$ (in A). Now by (i) and (ii) $A \supset k_A[X]$ and $f(X, \bar{y}) \equiv 0 \bmod (X)^{\beta_0}$ in $k_A[X]$. Hence by Theorem 3.2 there exists a $y \in k_A[X]^\sim$ such that $f(X, y) = 0$ and $y \equiv \bar{y} \bmod (X)^\alpha$. By (iii), $A \supset k_A[X]^\sim$ and hence $y \in A$, $f(X, y) = 0$ and $y \equiv \bar{y} \bmod \mathcal{M}^\alpha$. This shows that Φ_{β_0} is true in A.

Now by the Gödel Completeness Theorem there is a proof of Φ_{β_0} from T. We calculate a function β as follows. The axioms of T are effectively given so we can generate the (first order) theorems of T. Do this in some fixed way until one obtains a theorem of the form Φ_β. (We know this eventually happens since Φ_{β_0} is a theorem of T). Let Φ_{β_1} be the first such to occur. Set $\beta(n, N, d, \alpha) = \beta_1$.

Remarks. (i) The function β' of Theorem 4.2 can be shown to be computable by a similar proof.

(ii) The above method of proof, while establishing the computability of β gives a hopelessly inefficient algorithm.

References

1. Artin, M.: On the solutions of analytic equations, Inventiones Math. **5**, 277–291 (1968)
2. Artin, M.: Algebraic approximation of structures over complete local rings, Publ. Math. I.H.E.S., **36**, 23–58 (1969)
3. Ax, J., Kochen, S.: Diophantine problems over local fields I, II Amer. J. Math, **87**, 605–648 (1965); III Ann. Math. (2), **83**, 437–456 (1966)
4. Becker, J.: A counterexample to Artin approximation with respect to subrings, Math. Ann., **230**, 195–196 (1977)
5. Becker, J., Lipshitz, L.: Remarks on the elementary theories of formal and convergent power series, Fund. Math., (to appear)
6. Birch, B.J., McCann, K.: A criterion for the p-adic solubility of Diophantine equations, Quart. J. Math., Oxford, (2) **18**, 59–63 (1967)
7. Chang, C.C., Keisler, H.J.: Model Theory, North Holland, Amsterdam
8. Delon, F.: Résultats D'indécidabilité Dans Les Anneaux de Séries Formelles, (to appear)
9. Eršov, Yu.: On the elementary theory of maximal normed fields, Algebra i Logika, **4**, 31–69 (1965); **5**, 8–40 (1966); **6**, 31–37 (1967)
10. Greenberg, M.J.: Rational points in henselian discrete valuation rings, Publ. Math. I.H.E.S. **31**, 59–64 (1966)
11. Greenberg, M.J.: Strictly local solutions of diophantine equations, Pacific J. Math., **51**, 143–153 (1974)
12. Lang, S.: Algebra, Addison-Wesley, 1965
13. Lascar, D.: Caractère effectif des théorèmes d'approximation de M. Artin, (to appear)
14. Luxemburg, W.A.J.: A General Theory of Monads. In: Applications of Model Theory to Algebra, Analysis and Probability, pp. 18–86, New York: Holt Rinehart and Winston, 1969
15. Nagata, M.: Local rings, Interscience, New York, 1962
16. Nerode, A.: A decision method for p-adic integral zeros of diophantine equations, Bull. A.M.S. **69**, 513–517 (1963)
17. Pfister, G., Popescu, D.: Die strenge Approximationseigenschaft lokaler Ringe, Inventiones Math., **30**, 145–174 (1975)
18. Popescu, D.: A strong approximation theorem over discrete valuation rings, Rev. Roum. Math. Pures et Appl. Tome XX, No **6**, 659–692 (1975)
19. Robinson, A.: Elementary embeddings of fields of power series, Journal of Number Theory, **2**, 237–247 (1970)
20. van den Dries, L.: Model Theory of Fields, Thesis, Utrecht, June 1978

Received July 19, 1978

Current addresses: J. Becker, State University of New York, Buffalo, New York 14214
L. Lipshitz, Purdue University, West Lafayette, Indiana, 47907, USA
L. van den Dries, University of Utrecht, Utrecht, Netherlands