Immediate extensions of rings and approximation of roots

Phạm Ngọc Ánh*

July 11, 1997

Dedicated to Professor Irving Kaplansky on his 80^{th} birthday

Mathematical Institute
Hungarian Academy of Sciences
Reáltanoda u. 13-15, Budapest, H-1053 Hungary
e-mail: anh@math-inst.hu

current address:
Department of Mathematics, University of Bielefeld
POB 100131, 33501 Bielefeld, Germany
e-mail: anh@Mathematik.Uni-Bielefeld.de

Abstract

In this work we generalize the notion of immediate extensions of valued fields introduced in Krull's work [21] to arbitrary commutative rings. We show that every ring possesses maximal immediate extensions and point out the importance of linear compactness in the investigation of maximal immediate extensions of local rings as one of the natural classes of maximally complete rings. Another important, in some sense exceptional class of maximally complete rings is that of von Neumann regular rings. A necessary and sufficient condition is given for the linear compactness of any maximal immediate extension of a valuation ring. Newton's Approximation Method is used to give a short direct proof of Hensel's Lemma and to extend a result of Lang [22] for finding simple roots of polynomials over linearly compact local rings. A somewhat weaker version of Lang's condition for the existence of simple roots whose congruence classes are multiple roots modulo the maximal ideal is given by using an idea of Kaplansky. Results of Teichmüller, Kaplansky and Cohen on the existence of coefficient fields are also generalized to local linearly compact rings. In the last section Kaplansky's method [14], as a non-linear version of Newton Approximation, is used to find (possibly multiple) roots of polynomials, and this leads to an extension of Mac Lane's characterization of algebraically closed linearly compact valued fields to linearly compact valuation rings. Moreover, Kaplansky's result on the uniqueness of maximal immediate extensions of valued fields is reformulated as a result on the location and the approximation of roots, the classical topic for real polynomials. Therefore this reformulation yields new aspects of Kaplansky's result and points out at the same time a natural analogy between linearly compact valued fields satisfying Hypothesis A and the field $\mathbb{R}$ of real numbers or more generally, complete real closed fields.

*Diese Arbeit entstand zum Teil im Rahmen eines Studienaufenthaltes als Humboldtstipendiat an der Universität Bielefeld. Der Verfasser bedankt sich sowohl bei der Alexander-von-Humboldt Stiftung als auch bei Prof. Claus Michael Ringel für die Unterstützung. Die Arbeit ist teilweise auch von der Ungarischen Stiftung für wissenschaftliche Forschung Projekt Nr. T23434 unterstützt worden.

1 Introduction

The origins of this work can be found in the works of Krull [21], Ostrowski [28] and Kaplansky [14] on maximal extensions of valued fields. It is well known that the value group and the residue field do not characterize a valued field. For example, for every field K of prime characteristic p and every totally ordered group Γ, there are three fields which have K as residue field and Γ as value group: F_1 of characteristic 0, F_2 and F_3 of characteristic p. Note that in contrast to the existence of F_2 and F_3, for which we can simply take the quotient field of the group algebra $K\Gamma$ and the much bigger field $K[[\Gamma]]$ of generalized power series, respectively, the existence of F_1 is highly non-trivial. For the construction of F_1 one can use Krull's "Aufspaltungsprozess" to reduce the problem to the case when Γ is archimedean or equivalently F_1 is of rank 1, and then construct F_1 with the help of Witt rings. This observation led probably to the creation of the theory of maximal valued fields in [21]: these are those fields which are maximal in the set of valued fields with given residue field and value group. The theory of maximal valued fields culminated in the work [14] of Kaplansky. Roughly speaking, Kaplansky's result states that one can get all maximally complete valued fields by using Krull's Aufspaltungsprozess and Witt rings in the unequal characteristic case and generalized power series fields in the equal characteristic case. Kaplansky discussed also the uniqueness problem of maximal extensions and reduced it to the existence problem of roots of certain polynomials. Then he solved this question almost completely by giving both sufficient conditions for the existence of roots and a transfinite approximation of roots which is in fact an extension or, more precisely, a non-linear version of the Newton Approximation for valued fields; this approximation is also suitable for finding multiple roots. We shall call this method Kaplansky's method. As was already observed by Kürschák, the inventor of valuation theory, valued fields are not generalizations of ordered (i.e., formally real) fields: their theory is in some sense paralell to that of the latter. Together with Hensel's Lemma, Kaplansky's method presents a way to locate roots of an equation over maximally complete valued fields, in parallel to the classical topic of location of real roots. For real polynomials the Newton Approximation gives a very effective way to approximate single roots. It may be less effective, however, for multiple roots. In the parallel theory, approximation of single roots in (maximally) complete fields can be done by the help of Hensel's Lemma. However, the Newton Approximation does not work for multiple roots in maximally complete valued fields and very little is known about their approximation. We will show in this work that Kaplansky's method provides a powerful direct method to approximate multiple roots in some maximally complete valued fields which can be considered as natural analogues of the field of real numbers or, more generally, complete real closed fields. Of course, there is an indirect handy way to use Newton's method after eliminating multiple roots by dividing out the greatest common divisor of a polynomial and its derivative for characteristic 0 or by reducing a polynomial to its separable form for prime characteristic.

In the present paper we shall extend the notion of immediate extension to arbitrary commutative rings and discuss the close relationship between linear compactness and maximal extensions of local rings. Our definition of immediate extension can be extended to all associative rings, but that will be done in a subsequent paper.

In the last two sections we discuss the approximation of (possibly multiple) roots by using the methods of Newton and Kaplansky. Some of our results are of a general character, i.e. they hold for all linearly compact local rings, or even for non-commutative rings. We cannot extend, however, our discussion of Kaplansky's method to local rings. We can do it only for linearly compact valuation rings or even for a more special class of linearly compact valuation rings, although we are convinced of the broader usability of this method.

Throughout this paper, all rings are assumed to be commutative with an identity, and all modules to be unitary. Subrings of a given ring are assumed to have the same identity as the overring. We will denote by $E(M)$ the injective hull of a module M.

2 Immediate Extensions

The principal ideals of a ring R form a partially ordered monoid Γ, or more precisely, Γ_R under the ideal multiplication where the partial ordering is reverse inclusion. This monoid will be called the *monoid of divisibility* of R. More generally, if R is a subring of a ring S, then the principal (i.e. cyclic) R-submodules of S form also a partially ordered monoid $\Gamma_{S|R}$ or simply Γ_S if there is no confusion of the meaning, under the multiplication introduced by S where the partial ordering is also reverse inclusion. This monoid will be called the *monoid of divisibility of R with respect to the overring S*. For example, if K is a valued field with the valued group Φ and the corresponding valuation ring R, then the monoid Γ_R of divisibility of R is exactly the positive cone Φ^+ and the monoid Γ_K of divisibility of R with respect to K is exactly Φ. In analogy to valuation theory we shall use the additive notation for these monoids of divisibility of R. Moreover, if x is any element of R or S, then we will denote by $|x|$ the element induced by x in these monoids, i.e., $|1| = 0, |0| = \infty$. Note that through this work, 0 has different meanings, for example, it may be the zero element of the rings, modules or the the identity of the monoid of divisibility, but its meaning is fortunately, always clear in the particular cases.

Example. Let R be any ring and Q be the classical quotient ring of R, i.e. Q is a localization of R at the set of non-zero-divisors. Then R is a subring of Q, Γ_R is a submonoid of $\Gamma_{Q|R}$ and even the latter is the classical quotient monoid of Γ_R at the set $\{|x|\}$ where x runs over all non-zero-divisors of R, as it is routine to check. Furthermore, if R is a domain, then Q is the quotient field of R and we have in this case the usual divisibility for domains.

The following definition is inspired by the investigations of Krull [21] on immediate extensions of valued fields.

Definition. Let R be a subring of a ring T. T is said to be an *immediate extension* of R if their monoids of divisibility are equal, i.e., the natural map $Ra \mapsto Ta, a \in R$ is a bijecttion between the sets of principal ideals of R and T, respectively, and for every maximal ideal $\mathcal{M}$ of T the equality $T = R + \mathcal{M}$ hols. R is called *maximally complete* or simply *maximal* if it has no proper immediate extensions. An overring T of R is called *maximal immediate extension* or simple *maximal completion* of R if T is an immediate extension of R and T is maximally complete.

Remark. According to Krull [21] the notions of immediate extensions and maximal completions are due to Schmidt. However, Schmidt did never publish his related results on maximal immediate extensions of valued fields. In the famous work [21] Krull created the notion of perfectness which is somewhat weaker than the linear compactness and proved Hensel's Lemma for perfect fields. The notion of perfect valued fields can be considered in some sense as the forerunner of the notion of linear compactness. Krull pointed out that maximally complete valued fields are perfect and generalized power series fields are maximally complete. His method was improved by Ostrowski. In [28] Ostrowski proved that a real-valued field is maximally complete if and only if every pseudo-convergent sequence has a pseudo-limit which is the equivalent condition for the linear compactness of the corresponding valuation domain. The theory of maximally complete fields is summarized in Kaplansky's works. In his very early papers [14], [15] Kaplansky eliminated the restrictions in Ostrowski's works, emphasized the natural appearance of generalized power series fields in the structure theory of maximal complete valued fields and discussed the analytical isomorphism problem for maximal completions of valued fields via the approximation of roots.

It is obvious from definition that R is relatively divisible in any immediate extension T of R, or in other words, R is an AR-submodule of $_RT$. For the definition of *relative divisibility* and *AR-submodule* we refer to Warfield's work [37]. However, R is in general not a pure submodule of T. If R is in addition a principal ideal ring, i.e.. a ring whose ideals can be generated by one element, or a valuation ring, then R is trivially a pure submodule of T. The following almost trivial statement makes the situation clear.

Proposition 2.1 *Let R be a subring of T. T is an immediate extension of R and R is a pure submodule of T if and only if the compositions of the maps $I \mapsto IT, J \mapsto J \cap R, I \lhd R, J \lhd T$ are the identity on the lattices of ideals of R and T, respectively and for each maximal ideal $\mathcal{M}$ of T the equality $R + \mathcal{M} = T$ holds.*

Proof. The sufficiency follows trivially from the definition of immediate extensions and pure submodules. For the necessity it is enough to see that $I = IT \cap R$ for every ideal $I \lhd R$. Consider $r = \sum_{i=1}^{n} r_i t_i \in IT \cap R$. Since R is a pure submodule of T, we have $(\sum_{i=1}^{n} Tr_i) \cap R = \sum_{i=1}^{n} Rr_i \subseteqq I$. Hence $r \in I$ and our proof is complete. □

Remark. For every ideal J of an immediate extension T of R we have obviously $J = (J \cap R)T$. Therefore $\mathcal{N} = R \cap \mathcal{M}$ is a maximal ideal of R for every maximal ideal $\mathcal{M}$ of T. Conversely, if $\mathcal{N}$ is any maximal ideal of R and $\mathcal{M} = T\mathcal{N}$, then $\mathcal{M}$ is a maximal ideal of T. For if $t \notin \mathcal{M}$, then we have $Tt = Tr$ with an element $r \notin \mathcal{N}$. Therefore we have $R = Rr + \mathcal{N}$ by the maximality of $\mathcal{N}$. Thus $T = (Rr + \mathcal{N})T \subseteqq Tt + \mathcal{M}$ and hence $T = Tt + \mathcal{M}$, i.e., $\mathcal{M}$ is indeed a maximal ideal of T. This means that the maps $I \mapsto IT, J \mapsto J \cap R$ establish a bijection between the sets of maximal ideals of R and T, rerspectively.

Proposition 2.2 *Let T be an immediate extension of R. For every maximal ideal $\mathcal{M}$ of T the local ring $T_{\mathcal{M}}$ is an immediate extension of the local ring $R_{\mathcal{N}}$ where $\mathcal{N} = R \cap \mathcal{M}$. Moreover we have the equality $T = R_{\mathcal{N}}T$.*

Proof. ¿From the previous remark we know that $\mathcal{N}$ is a maximal ideal of R. The equality $\mathcal{N} = R \cap \mathcal{M}$ implies the inclusion $R \setminus \mathcal{N} = R \setminus \mathcal{M} \subseteqq T \setminus \mathcal{M}$. Consequently we have the canonical map from $R_{\mathcal{N}}$ in $T_{\mathcal{M}}$. It is well-known that $a \in R$ maps to 0 in T_M if and only if $at = 0$ for some $t \notin \mathcal{M}$. The equality $Tt = Tr$ for some $r \in R$ implies $r \notin \mathcal{N}$ and $at = 0$ if and only if $ar = 0$. Therefore $R_{\mathcal{N}}$ is a subring of $T_{\mathcal{M}}$. From $T_{\mathcal{M}} = T + T_{\mathcal{M}}M = R + T_{\mathcal{M}} = R_{\mathcal{N}} + T_{\mathcal{M}}\mathcal{M}$ we obtain easily that $T_{\mathcal{M}}$ is in fact an immediate extension of $R_{\mathcal{N}}$

For the last equality of the statement it suffices to see that every element $t \notin \mathcal{M}$ is invertible in $R_{\mathcal{N}}T$. Since $Tt = Tr$ for some $r \in R$ we deduce $r \notin \mathcal{N}$ and $r = tv$ for some $v \in T$. Thus together with r, t is also invertible in $R_{\mathcal{N}}T$ which finishes the proof. □

Remark. The last equality of Proposition 2.2 shows that as modules over R the module $T_{\mathcal{M}}$ is the localization of $_RT$ with respect to the maximal ideal $\mathcal{N}$ of R. Therefore we use sometimes the notation $T_{\mathcal{N}}$ for $T_{\mathcal{M}}$, too, or more precisely, we make no difference between $T_{\mathcal{N}}$ and$T_{\mathcal{M}}$.

Proposition 2.3 *Let R be a local ring with the maximal ideal $\mathcal{N}$, and T be an immediate extension of R. Then as modules over R the injective hull E of the simple T-module is isomorphic to a submodule of the injective hull of the simple R-module*

Proof. In view of the remark before Proposition 2.2 $\mathcal{M} = T\mathcal{N}$ is the maximal ideal of T. The isomorphic factor fields $R/\mathcal{N} \cong T/\mathcal{M}$ imply that it is enough to see that ${}_RE$ is an essential extension of its socle. For if $0 \neq x \in E$, then there is $0 \neq t \in T$ such that tx is a non zero element in the socle of E. Since $Tt = Tr$ for some $r \in R$, we have $t = rb, r = at$ for some $a, b \in T$. Consequently $t(1 - ab) = 0$ and hence $1 - ab \in \mathcal{M}$ which ensures that both a and b are units in T. Thus $0 \neq rx$ belongs to the socle of E and therefore our proof is complete. □

Corollary 2.1 *Let T be an immediate extension of a local ring R. Then the cardinality of T is bounded by κ^κ where κ is the cardinality of the minimal injective cogenerator of R.*

Corollary 2.2 *Let T be an immediate extension of a ring R. Then the cardinality of T is bounded.*

Proof. It is well-known that an element $t \in T$ is 0 if and only if t maps to 0 in $T_\mathcal{M}$ for every maximal ideal $\mathcal{M}$ of T. Hence the statement follows directly from Corollary and the remark before Proposition 2.3. □

In view of Corollary 2.2, as a trivial consequence of Zorn's Lemma we obtain:

Theorem 2.1 *Every ring admits a maximal immediate extension.*

Remark. Krull's original proof for the existence of maximal completions of a valued field F with a valued group Γ and a residue field K was based on his ingenious construction of an injective map from F to a generalized power series field $K[[\Gamma]]$. His proof is a nice demonstration of the importance of power series fields in the theory of maximal completions. Simplifications of Krull's proof can be found in [12], [30]. For valuation rings or even for Prüfer rings the observation in Warfield's work [37] that an immediate extension is a pure-essential extension of the ring implies another possibility to deduce an existence of maximal completions. The idea to use injective modules for the existence of maximal completions is mentioned in Kaplansky's Lecture Notes [18], too.

The injective hull of a simple T-module is in general not injective as an R-module. For some particular cases, however, T is a flat R-module and hence injective T-modules are also injective R-modules.

Proposition 2.4 *If R is a valuation ring and T is an immediate extension of R, then T is a flat R-module.*

Proof. It is enough to see that the canonical map $aR \underset{R}{\otimes} T \longrightarrow aT : a \otimes t \mapsto at$ is injective. Assume $at = 0$. Since $t = ur$ for some $r \in R$ and a unit $u \in T$ as it was checked in Proposition 2.3, we have $ar = 0$. Therefore $a \otimes t = a \otimes ru = ar \otimes u = 0$, which implies that ${}_RT$ is flat. □

Since an R-module M is flat if and only if $M_\mathcal{M}$ is a flat $R_\mathcal{M}$-module for every maximal ideal $\mathcal{M}$ of R, the previous statement and a remark before Proposition 2.3 implies

Corollary 2.3 *Let R be a Prüfer ring and T is an immediate extension of R, then ${}_RT$ is flat.*

Recall that a ring R (not necessarily without zero-divisors) is said to be a *Prüfer ring* if the lattice of ideals of R is distributive.

Corollary 2.4 *If R is a principal ideal ring, T is an immediate extension of R, then ${}_RT$ is flat.*

In [8] Corner constructed an uncountable immediate extension of the ring $\mathbb{Z}$ of intergers which played a key role in his proof of the representation of some rings as endomorphism rings of abelian groups. His results suggest the following two problems:

Problem 1. Let T be a maximal completion of a principal ideal domain R. Is it true that $T_{\mathcal{M}}$ is a complete DVR for every maximal ideal $\mathcal{M}$ of T?

Problem 2. Characterize maximal completions of particular rings, for example, of Dedekind domains, UFD or noetherian domains.

The following two statements give some classes of maximally complete rings.

Proposition 2.5 *Every von Neumann regular ring R is maximally complete.*

Recall that a ring is said to be *von Neumann regular* if for every element x there is an element a with $x = xax$.
Proof. Let T be any immediate extension of a von Neumann regular ring R. We have to show $T = R$. Since localizations of von Neumann regular rings at maximal ideals are fields, in view of the remark before Proposition 2.3 we have $R_{\mathcal{N}} = T_{\mathcal{N}}$ for every maximal ideal $\mathcal{N}$ of R. Therefore we have $R = T$ which completes the proof. $\square$

Proposition 2.6 *If R is a linearly compact ring, then R is maximally complete*

Proof. Since every linearly compact ring is a finite direct sum of local rings by the result of Zelinsky [41], we can assume without loss of generality that R is a local ring. Since a linearly compact local ring has a Morita duality in view of the result in [1], it is the endomorphism ring of its minimal injective cogenerator which has no proper faithful submodules. Therefore Proposition 2.3 implies that R is maximally complete. $\square$

For valuation ring Proposition 2.6 is proved by Vicknair [36]. It is clear from Propsition 2.5 that maximal completions of a rings are in general not linearly compact. It is a very exciting problem to describe local rings whose maximal completions are linearly compact.

The notion of pseudo-convergent sequences can be extended naturally to local rings as follows. Let $||$ denote the canonical map from a local ring R to its monoid Γ_R of divisibility. A set $\{r_\rho\}_{\rho\in A}$ of elements from R is called a *pseudo-convergent sequence* provided A is a well-ordered set without a last element, the set $\{|r_\rho|\}$ is a totally ordered subset of Γ and $|r_\sigma - r_\rho| < |r_\tau - r_\sigma|$ for all $\rho < \tau < \sigma$. It is routine to check that for any index ρ the value $|r_\sigma - r_\rho|$ is a constant, denoted by γ_ρ, for all $\sigma > \rho$. The ideal $\{r \in R \mid \forall\rho : |r| > \gamma_\rho\}$ is called the *breadth* of $\{r_\rho\}$. An element $r \in R$ is called a *pseudo-limit* of $\{r_\rho\}$ if $|r - r_\rho| = \gamma_\rho$ for all ρ. As in the case of valued fields one can easily check that for a pseudo-convergent sequence $\{r_\rho\}$ the corresponding sequence $\{|r_\rho|\}$ is ultimately (i.e. from some point on) monotone increasing or stationary.

Proposition 2.7 *Let T be an immediate extension of a local ring R. Then for every element $t \in T \setminus R$ there is an element $r \in R$ with $|r| = |t| < |t - r|$*

Proof. By assumption there is $x \in R$ with $|t| = |x|$. Since T is a local ring, we have $t = ux$ for some unit $u \in T$. The equality $T = R + \mathcal{M}$ where $\mathcal{M}$ is the maximal ideal of$T$, implies $u = v + y$ for some $v \in R, y \in \mathcal{M}$. Therfore v is a unit in R. For $r = vx$ we get $|t| = |r|$ and $t - r = ux - vx = yx$ which implies $|t - r| > |t| = |r|$. $\square$

We are now in position to prove another main result of this section, a generalization of Theorem 1 in [14].

Theorem 2.2 *Let T be an immediate extension of a local ring R and let $||$ denote the canonical map from R and T to the monoid Γ of divisibility of R. Then every element in $T\backslash R$ is a pseudo-limit of a pseudo-convergent sequence of elements of R without a pseudo-limit in R.*

Proof. Let $t \in T \setminus R$. Then $0 \neq t$ and $|t-r| \neq \infty$ for all $r \in R$. Consider the set of all well-ordered sequences $\{r_\rho\}$ in R satisfying

1. $|r_\rho| = |t|$ for all ρ
2. $|r_\sigma - r_\rho| < |r_\tau - r_\sigma|$ for all $\rho < \sigma < \tau$
3. $|t - r_\rho| = |r_{\rho+1} - r_\rho|$, if r_ρ is not the last element of the sequence

In view of Proposition 2.7 this set is non-empty. We define in this set a partial order as follows. A sequence $\{r_\rho\}$ is smaller than a sequence $\{r_\sigma\}$ if the former is a subsequence of the latter. It is easy to check that this partial ordering satisfies the conditions of Zorn's Lemma. Hence there is a maximal sequence, say $\{r_\rho\}$, with the above described 3 properties. First we claim that $\{r_\rho\}$ has no last element. For if r_τ is the last element, then $|t-r_\tau| = |(t-r_\rho)+(r_\rho-r_\tau)| \geqq |r_\rho-r_\tau|$ by conditons 2 and 3. Moreover, by Propositon 2.7 there is an element $b \in R$ such that $|t-r_\tau| = |b| \leqq |t-(b+r_\tau)|$. Putting $r_{\tau+1} = r_\tau + b$ we obtain a greater sequence, which is impossible. Thus $\{r_\rho\}$ does not admit the last element. For every index ρ choose an element $b_\rho \in R$ of value $|r_{\rho+1} - r_\rho|$. If the intersection $\cap(r_\rho + Rb_\rho)$ is non-empty, then it is easy to see that the sequence $\{r_\rho\}$ is extendable to a greater sequence by adjoining any new element of this intersection which contradicts to the maximality of $\{r_\rho\}$. Thus $\{r_\rho\}$ has no pseudo-limit in R and our proof is complete. □

Remark. The above result suggests the following question. Let R be a maximally complete local ring. Is it true that R is *weakly linearly compact* in the sense that every linearly ordered of cosets of ideals of R with respect to inclusion has a non-empty intersection?

3 Linear compactness and maximal completions of valuation rings

In this section we investigate the linear compactness of maximal completions of valuation rings. As the main result we show that the linear compactness of maximal completions is equivalent to the existence of a commutative ring structure on the injective hull of rings as modules over themselves. This requirement is trivial for domains because their injective hulls are quotient fields which admit trivially a commutative ring structure.
The following statement gives some information on endomorphism rings of injective modules over valuation rings. The idea of the proof is taken from the joint work [20] of Klatt and Levy.

Proposition 3.1 *Let E be an injective module over a valuation ring R and S be the endomorphism ring od ${}_RE$. Then E is a linearly compact uniserial module over S and hence S is a linearly compact ring with respect to the finite topology.*

Proof. Let u, v be any two non-zero elements of E. Since R is a valuation ring, we can assume without loss of genarality that $ann_R u = \{r \in R \mid ru = 0\} \subsetneqq ann_R v$. Therefore it is easy to check that the map $Ru \longrightarrow Rv : ru \mapsto rv$ is well-defined and even an R-homomorphism. Hence it extends to an element s of S by the injectivity of ${}_RE$. Consequently $v = us$ holds. This implies exactly that E is a uniserial module over S. Moreover, we have the equality $uS = ann_E ann_R u$ for every element $u \in E$. For if $v \in ann_E ann_R u$, then we have the inclusion $ann_R u \subsetneqq ann_R v$ and hence $v \in Us$ which justifies our claim. For the linear compactness of the S-module E let $\{u_\alpha + E_\alpha\}$ be any filter base of cosets of S-submodules E_α of E. Since E is a uniserial module, by possibly thowing away some of the cosets $u_\alpha + E_\alpha$ we may assume that $\{E_\alpha\}$ is a strictly descending chain of S-submodules of E. Put $I_\alpha = ann_R E_\alpha$ and consider the map

$$\phi : \sum I_\alpha \longrightarrow E : \sum i_\alpha \mapsto \sum i_\alpha u_\alpha.$$

It is routine to check that this map is well-defined and in fact an R-module homomorphism. Therefore the R-injectivity of E implies the existence of an element $u \in E$ such that

$$\phi(\sum i_\alpha) = (\sum i_\alpha)u = \sum i_\alpha u_\alpha.$$

In particular, we have $i_\alpha u_\alpha = i_\alpha u$ for every $i_\alpha \in I_\alpha$, i.e,

$$u - u_\alpha \in ann_E I_\alpha = ann_E ann_R E_\alpha.$$

Choose an E_β with $E_\beta \subsetneqq E_\alpha$ and an elelment $v_\alpha \in E_\alpha \setminus E_\beta$. Hence the inclusion $E_\beta \subsetneqq v_\alpha S \subseteqq E_\alpha$ holds. From the previous result we obtain

$$\begin{aligned} u - u_\alpha = (u_\beta - u_\alpha) + (u - u_\beta) \in E_\alpha + ann_E ann_R E_\beta \subseteqq \\ E_\alpha + ann_E ann_R v_\alpha = E_\alpha + v_\alpha S = E_\alpha. \end{aligned}$$

This shows that E is a linearly compact S-module. Since S is exactly the ring of those hypercontinuous endomorphisms of the abelian group E which commute with the elements of R, the linear compactness of S with respect to the finite topology is an obvious consequence of [24] Theorem 7. Hence our proof is complete. □

For the particular case when $E = E(R)$ is the injective hull of R we have

Corollary 3.1 *Let E be the injective hull of R and S be the endomorphism ring of E. The linearly ordered set of S-submodules of E with respect to inclusion is order-isomorphic to the linearly ordered set Γ_{R_Z}, the monoid of divisibility of the classical ring R_Z of quotients of R where Z is the prime ideal of zero-divisors.*

Proof. First we claim that every cyclic S-submodule of E can be generated by some element of R. For if $0 \neq x \in E$, then there are $a, b \in R$ such that $0 \neq b = ax$. This implies $ann_R a \subseteqq ann_R b$. There are two possibilities. If $ann_R a$ is strictly contained in $ann_R b$, then $b = ac$ for some $c \in R$ and therefore we can easily deduce that $ann_R c = ann_R x$. Hence x, c generate the same S-submodule of E by the injectivity of E. If $ann_R a = ann_R b$ and r is any non-zero element of R, then for elements $u, v \in R$ with $0 \neq ua = vx$ we have $rvx = uax = ub \neq 0$ because $0 \neq ua$ and $ann_R a = ann_R b$. Note that the existence of u, v follows from the fact that E is an essential extension of R. This implies that the S-submodule of E is E and hence it can be generated by $1 \in R$. Since two elements $a, b \in R$ generate the same S-submodule of E if and only if $ann_R a = ann_R b$ we obtain either $a = rb$ or $b = ra$ for some non-zero-divisor $r \in R$. This observation makes our proof complete. □

Remark. This assertion raises the following problem. Let T be an overring of a valuation ring R which is an essetial extension of R as an R-module, i.e., T is an R-submodule of $E(R)$. Let $||$ denote the canonical map from R to Γ_{R_Z}. This map can be extended naturally to a map $|| : T \mapsto \Gamma_{R_Z}$ by putting $|t| = |a| - |b|$ for $0 \neq bt = a \in R$ for some $b \in R$. It is routine to check that this map is well-defined and it is compatible with the R-structure of T, i.e., $|rt| = |r| + |t|, |t_1 + t_2| \geqq min\{|t_1|, |t_2|\}$ for every $r \in R; t, t_1, t_2 \in T$. The problem is to find reasonable condition on T or much better on R such that this map is a valuation map, i.e., $\forall t, s \in T : |ts| = |t| + |s|$ hold.

Corollary 3.2 *Let E be the injective hull of a valuation ring R. If E admits a commutative ring structure which extends one of R, then E is a valuation ring which is a maximal immediate extension of the localization R_Z at the prime ideal Z of zero divisors.*

Proof. Let S be the endomorphism ring of E. The condition implies that E can be considered as a subring of S. Let L be the annihilator of $1 \in R \subseteqq E$ in S. Since 1 generates E as a module over S, we obtain $S = E \oplus L$. This shows that E is also a uniserial module over itself and every E-submodule of E is at the same time an S-submodule of E. Therefore E is a valuation ring, Since E is linearly compact over S we obtain from the above consequence that E is an immediate extension of R_Z. $\square$

Remark. This assertion is trivial for valuation domains because their injective hulls are isomorphic to the fields of quotients and hence they admit a ring structure which extends the original ones. On the other hand, Krull's "Aufspaltungsprozess" ([21] pages 171-173) implies that the subring of E consisting of the inverse image of the valuation ring of the factor field E/ZE in T, is an immediate extension of R as it is easy to check.

We are now ready to give the main result of this section which provides a necessary and sufficient condition for the linear compactness of any maximal immediate extension of a valuation ring.

Theorem 3.1 *The injective hull of every immediate extension T of a valuation ring R admits a commutative ring structure which extends one of T if and only if every maximal immediate extension of R is linearly compact.*

Proof. If V is a linearly compact valuation ring and Z is the set consisting of zero-divisors, then Z is a prime ideal and the localization V_Z (i.e., the classical quotient ring) of V at Z is also a linearly compact valuation ring. In V_Z every element is either a unit or a zero-divisor. Therefore by the result of Klatt-Levy (see e.g. [9], [20]) V_Z is a self-injective ring which shows the necessity of the theorem.

For the sufficiency we have to show the linear compactness of every maximal immediate extension of R. It is obviously enough to see that if R is not linearly compact, then there is a proper immediate extension of R. In view of the previous corollary we can assume that R_Z is linearly compact where Z is a prime ideal of zero-divisors of R as usual. Thus if R is not linearly compact, then there exists $a_\alpha \in R$ and ideals I_α of R such that $\cap(a_\alpha + I_\alpha) = \varnothing$ yet $\{a + I_\alpha\}$ is a filter base. Since R is a valuation ring, we can assume by possibly throwing away some of the cosets $a_\alpha + I_\alpha$ that the index set is well-ordered and $a_\alpha + I_\alpha \supsetneqq a_\beta + I_\beta$ if $\alpha < \beta$. Consequently $|a_\alpha|$ are constant, i.e., equal for all α and $\{a_\alpha\}$ is in fact a pseudo-convergent sequence. Thus we can assume without loss of generality also that $I_\alpha = R(a_\beta - a_\alpha)$ for all $\alpha < \beta$. Since $\{a_\alpha + R_Z I_\alpha\}$ is clearly a filter base of cosets of the ideals $R_Z I_\alpha$ in R_Z, too, there is $c = b^{-1}c \in \cap(a_\alpha + R_Z I_\alpha)$ for some $\alpha \in R, b \in R \setminus Z$ by the linear compactness of R_Z. Hence $a_\alpha - b^{-1}a \in R_Z I_\alpha$ for all α. Since $R_Z I \cap R = \{r \mid \exists t \in I : ann_R r \supseteqq ann_R t\}$ holds for any ideal $I \lhd R$, as it is easy to verify, we have $R_Z I_\alpha \cap R \subsetneqq I_\beta$ for $R_Z I_\alpha \subsetneqq R_Z I_\beta$. Therefore if $\cap(R_Z I_\alpha \cap R) = \cap I_\alpha$, then $a \in \cap(b a_\alpha + I_\alpha)$ and hence $\cap(a_\alpha + I_\alpha) \neq \varnothing$, a contradiction. Thus we can assume that all $R_Z I_\alpha$ are equal by throwing the superfluous ideals. Passing from $\{a_\alpha\}$ to $\{a_\beta - a_\rho, \beta > \rho\}$ if it is necessary, we may assume that all a_α belong to the fix ideal $R_Z I_\alpha$ of R_Z. Choosing an element $s \in R$ with $|s| = |a_\alpha|$ and write $a_\alpha = s s_\alpha$ we obtain a new pseudo-convergent sequence $\{s_\sigma\}$ such that all s_α and all the differences $s_\alpha - s_\beta$ are non-zero-divisors in R. Hence if R is not linearly compact, then there are elements $a_\alpha \in R$ and ideals $I_\alpha \lhd R$ such that the filter base $\{\bar{a}_\alpha + \bar{I}_\alpha\}$ has an empty intersection, where the index set is well-ordered, $a_\alpha + I_\alpha \subsetneqq a_\beta + I_\beta$ for $\alpha > \beta$ and $^-$ donotes the canonical morphism from R to the factor ring R/Z. Thus all the elements $\bar{a}_\alpha$ are different. Now we can follow the classical construction of immediate extensions of valued fields as it is described in [5]. For any primitive polynomial $f \in R[X]$ the elements $f(a_\alpha)$ are all non-zero-divisors from some index on, otherwise the nonzero polynomial $\bar{f} \in \bar{R}[X]$ would have infinitely many different roots which is impossible. We have two cases.

Case 1. For every polynomial $f \in R[X]$ the values $|f(a_\alpha)|$ are all equal from some index α an. Let P be the set of all $f \in R[X]$ such that the common value $|f(a_\alpha)|$ from some index on is contained in the set $|Z| = \{|r| \mid r \text{ is a zero-divisor}\}$, i.e., all $f(a_\alpha)$ are zero-divisors from some index on. Then P is a prime ideal of $R[X]$ as it is easy to check. In view of the observation made before this paragraph a polynomial f belongs to P if and only if it can be written as a product with one factor from the prime ideal Z of zero-dividors in R. Let $Q = R[X]_P$ be the localization of $R[X]$ at P. The factor field Q/PQ is clearly the function field $F(X)$ where F is the factor field R_Z/ZR_Z as it is routine to verify. Every element of Q is of the form $c^{-1}d(f(X))^{-1}g(X)$ where $c, d \in R$, $c \notin Z$ and $f(X), g(X) \in R[X]$ are primitive polynomials. This implies that Q is a valuation ring. In view of case 1 in [5] Theorem 12.6 we know that Q/PQ is an immediate extension of the valued field R_Z/ZR_Z. Using Krull's "Aufspaltungsprozes" let T be the inverse image of the valuation domain of Q/PQ under the canonical map from Q to Q/PQ. We claim that T is an immediate extension of R. Since $Q \setminus P$ is obviously a set of non-zero-divisors by the definition of Q, R is a subring of T. From the construction of T the equality $T = R + T\mathcal{M}$ holds trivially. Every element t of T is of the form $t = (f(X))^{-1}g(X)b$ where $f(X) \in R[X] \setminus P, b \in R$ and $g(X)$ is a primitive polynomial. Hence $g(X)$ is a unit in Q, too. We have two possibilities. Firstly, b is a zero-divisor. In this case, we can write $b = cd; c, d \in R$, such that $|d|$ is the common value of $|f(a_\alpha)|$ from some index on. If $u \in R$ such that $|u| = |g(a_\alpha)|$ form some index on, then the elements $d^{-1}f(X), u^{-1}g(X)$ are units in T by the construction. Therefore t and $cu \in R$ generate the same ideal in T. All these facts imply that T is indeed an immediate extension of R.

Case 2. There is a primitive polynomial $f \in R[X]$ such that there exists no index α for which all $|f(a_\beta)|, \beta > \alpha$ are equal. In this case as in the proof of [5] Theorem 12.6 we can see that a primitive polynomial f of the possibly smallest degree n is an irreducible polynomial and $n \geqq 2$. Since we can throw from f all the terms which have the coefficients from Z, we can assume that all coefficients in f are non-zero-divisors. Let a be the highest coefficient in f. Then a is a non-zero-divisor. From the fact that the degree of f is at least 2, we obtain that the new filter base $\{a^{n-1}a_\alpha + I_\alpha\}$ satisfies the conditions required for $\{a_\alpha + I_\alpha\}$. By passing from $\{a_\alpha\}$ to $\{a^{n-1}a_\alpha\}$, if it is necessary, we can assume that f is even a monic polynomial. Consider now the factor ring $R[x]$ of $R[X]$ by the ideal generated by f. Here x is the image of X. Every element of $R[x]$ can be written uniquely as a value of a polynomial $g \in R[X]$ at the substitution $X = x$ of a degree $< n$. Moreover, in this case the values $|g(a_a)|$ are equal from some index on by the minimality of n. In $R[x]$ the set P of all $g(x)$ such that $deg(g) < n$ and the common value of $|g(a_\alpha)|$ from some index on is a zero-divisor, is again a prime ideal. The localization of $R[x]$ at P is again a valuation ring with the factor field Q/PQ is a finite field extension $F[x]$ of the factor field $F = R_Z/ZR_Z$. Again by Case 2 in [5] Theorem 12.6 $F[x]$ is an immediate extension of the valued field F. As in the Case 1 we can verify that the inverse image T of the valuation domain of $F[x]$ under the map $Q \longrightarrow F[x]$ is an immediate extension of R. Therefore our proof is complete. $\square$

We conclude this section with some remarks on the examples of valuation rings which are not homomorphic images of valuation domains. All existing examples are based on the original example and idea of Fuchs and Salce [10]. Let M be a nonstandard divisible uniserial module over a valuation domain D. Then the trivial extension $R = D \ltimes M$ is a valuation ring which is not a homomorphic image of a valuation domain. Let B be any maximal immediate extension of D. Then ${}_BB$ is linearly compact and D is a pure D-submodule of B. Hence $M \cong D \underset{D}{\otimes} M$ is a subset of $B \underset{D}{\otimes} M$. It is easy to check that $B \underset{D}{\otimes} M$ is a uniserial divisible module over B. Consequently, it must be standard by the linearly compactness of B. Therefore the trivial extension $B \ltimes (B \underset{D}{\otimes} M)$ is a linearly compact maximal immediate extension of R and is a homomorphic image of a valuation domain. However, we are not able to prove that all maximal immediate extensions of R are both linearly compact and homomorphic images of valuation domains. This suggests the following two problems

Problem 3. Is it true that maximal immediate extensions of valuation rings are homomorphic images of valuation domains?

Problem 4. Are maximal immediate extensions of valuation rings henselian?

Note that if R is a linearly compact valuation ring such that $R = D \oplus N$ for some subring D and the minimal prime ideal $N \lhd R$ is nilpotent, then one can see in the similar way as it was done in the proof of Lemma 3 [11] that R is a factor of a valuation domain.

4 Approximation of roots I: Newton's method for simple roots

In [14] Kaplansky deduced that the uniqueness of maximal completions or in the other words, the existence of an analytical isomorphism between maximal completions of a valued field is equivalent to the solubility of some equations and he introduced Hypothesis A as a pair of conditions which ensure the uniqueness of maximal immediate extensions. In fact he gave also a transfinite algorithm to find a root. We will investigate his method in the next section when we use his idea to approximate multiple roots. We devote the following two sections to the investigation of the solubility of polynomials over linearly compact rings. The classical process in (algebraic) number theory is to solve the equation $f(X) = 0$ firstly modulo p, where p is the (unique) prime element, then lift the solution step by step over the powers of p, and form finally their limit which will be probably a root. In this process we have three cases. The first simplest case is that when $f(X) = 0$ has a simple root modulo p. The second, relatively simple case is that when $f(X) = 0$ has a simple root whose congruence class is a multiple root of $f(X) = 0$ modulo p. The most complicated one is the case when $f(X) = 0$ has a multiple root. The case of simple root modulo p is solved completely by Hensel for p-adic integers. His celebrated result, the so-called Hensel's Lemma plays a decisive role in algebra and number theory. In this section we use Newton Approximation Method to deal with the first two cases for linearly compact local rings. First we give a short proof for the (probably well-known) fact that Hensel's Lemma holds for linearly compact local rings. Our proof is a simple application of Newton Approximation explained in Lang's works [22], [23]. As a non-direct elegant proof we can deduce obviously Hensel's Lemma for linearly compact local rings (in the discrete topology) from the beautiful treatment [2] of Azumaya on henselian rings. Note that in [13] (see [29], too) Hartmann gave a direct proof of Hensel's Lemma for linearly compact valued fields. His result is another inspiration to giving a direct proof of Hensel's Lemma for linearly compact rings which emphasizes the use of Newton Approximation. The case of simple roots which are multiple roots modulo the maximal ideals is discussed by using the ideas of Kaplansky and Lang. As applications we investigate the existence of coefficient fields in some linearly compact local rings as well as Hypothesis A for linearly compact valued fields.

For the sake of convenience let us recall the characteristic-independent Taylor Formula for polynomials (see for example [4] pp. A. IV. 6-8). Let $f \in R[X]$ be an arbitrary polynomial of degree n. Subsituting $X + Y$ for X we obtain the polynomial $f(X + Y) \in R[X, Y] = R[X][Y]$ which may therefore be written as a polynomial in Y with coefficients in $R[X]$

$$f(X + Y) = f(X) + f_1(X)Y + \ldots + f_i(X)Y^i + \ldots + f_n(X)Y^n \tag{1}$$

Here f_1 is by definition the (first, usual) derivative f' of f. The polynomial $f_i, i \geqq 2$ is called the *i-th formal derivative* of f. For the sake of unified notation we will sometimes denote $f(X)$ by

$f_0(X)$. The above equality is called the general *Taylor Formula* of f. Substituting $X = a$, $Y = b$ for $a, b \in R$ we obtain the Taylor Formula or in the other words, the Taylor expansion for f as

$$f(a+b) = f(a) + f_1(a)b + \ldots + f_n(a)b^n. \tag{2}$$

In view of the Taylor Formula for f we can deduce the Taylor Formula for f_i, which we will often use in the next section, as follows:

$$f_i(X+Y) = f_i(X) + (i+1)f_{i+1}(X)Y + \ldots + \binom{n}{n-i} f_n(X)Y^{n-i}. \tag{3}$$

Now we are able to prove Hensel's Lemma for linearly compact local rings. Note that in the next assertion the element $k \in K$ can be 0, too, i.e., the image of an element in the maximal ideal of R.

Proposition 4.1 (Hensel's Lemma for linearly compact rings) *Let R be a local linearly compact ring with the maximal ideal $\mathcal{M}$ and the factor field $K = R/\mathcal{M}$. Let $^-$ denote the canonical homomorphism from R to K.If the image $\bar{f}$ of a polynomial $f \in R[X]$ has a simple root k in K, i.e., $\bar{f}(k) = 0, \bar{f}_1(k) \neq 0$ where f_1 is the derivative of f, then there is $a \in R$ with $f(a) = 0, \bar{a} = k$.*

Proof. For every $a \in R$ we denote by I_a the ideal generated by $f(a)$. Consider the set $\mathcal{F}$ of all pairs (a, I_a) such that $a \in R$ and $\bar{a} = k$. The set $\mathcal{F}$ is obviously non-empty and $I_a \subseteq \mathcal{M}$ by the assumption that $\bar{f}(\bar{a}) = \bar{f}(k) = 0$. We define a partial ordering in $\mathcal{F}$ by putting $(a, I_a) \leqq (b, I_b)$ if and only if $f(a) \in I_b$, i.e., $I_a \subseteq I_b$ and $a - b \in I_b$. If (a_α, I_{a_α}) is a chain in $\mathcal{F}$, then the linear compactness of R implies $\cap(a_\alpha + I_{a_\alpha}) = I \neq \varnothing$. For any $a \in I$ and each index α we have $\bar{a} = k$ and $c_\alpha = a - a_\alpha \in I_{a_\alpha}$. Consequently $f(a) = f(a_\alpha + c_\alpha) = f(a_\alpha) + d$ where d is a sum of products which have c_α as one factor. Therefore $f(a) \in I_{a_\alpha}$. This shows $(a, I_a) \leq (a_\alpha, I_{a_\alpha})$ for all α. Thus Zorn's Lemma implies the existence of a minimal pair (a, I_a). We claim $I_a = 0$, i.e. $f(a) = 0$ and hence a is a root of f with $\bar{a} = k$. In fact, $\bar{a} = k$ implies $0 \neq \bar{f}_1(k) = \bar{f}_1(\bar{a})$. Hence $f_1(a)$ is a unit because R is a local ring. Following Newton Approximation let $b = a - f_1^{-1}(a)f(a)$. By Taylor Formula we have

$$f(b) = f(a) + f_1(a)(-f_1^{-1}(a)f(a)) + \ldots + f_n(a)(-f_1^{-1}(a)f(a))^n \subseteq R(f(a))^2 \subseteq I_a$$

where n is the degree of f and $f_i, i > 1$ is the i-th formal derivative of f. Therefore the minimality of (a, I_a) implies $R(f(a))^2 = Rf(a)$, i.e., $f(a) = r(f(a))^2$ holds for some $r \in R$. Thus we have $f(a)(1 - rf(a)) = 0$. Since $\bar{f}(\bar{a}) = \bar{f}(k) = 0$ holds by the assumption, we have $f(a) \in \mathcal{M}$. Consequently, $1 - rf(a)$ is invertible and hence $f(a) = 0$. Thus our proof is complete. □

Remark. This proof works also for non-commutative linearly compact rings if we consider only polynomials with coefficients from the ring $\mathbb{Z}$ of integers or from the centre of R, respectively. I_a denotes, of course, in this case only an one-sided ideal of R generated by $f(a)$, say, left ideal if R is left linearly compact. Observing the fact that a commutes with both $f(a)$ and $f_i(a), i = 1, 2, \ldots, n$ and thus $I_a a \subseteq I_a$, $a(f_1(a))^{-1} = (f_1(a))^{-1}a$ hold, we can carry out word for word the above proof to linearly compact non-commutative rings. It works even in the case where R is not a local ring and $\mathcal{M}$ denotes the Jacobson radical of R. In this case we must require concretely that $\bar{f}_1(k)$ is invertible in $\bar{R}$. As an interesting application we re-obtain the lifting of idempotents for linearly compact rings in the very conceptual way by considering the polynomial $f = X^2 - X \in \mathbb{Z}[X]$. If we have $a \in R$ with $a^2 - a \in \mathcal{M}$, i.e., the image $\bar{a}$ of a in the factor ring $\bar{R} = R/\mathcal{M}$ is an idempotent (a root of f), then $(2a-1)^2 = 1 + 4(a^2 - a)$ is invertible and hence $f_1(a) = f'(a) = 2a - 1$ is also invertible. If R has no identity element, the standard method of replacing inverve by quasi-inverse will carry the proof through. (Note that even in the case of non-unital rings we may compute formally with identity if we consider integers as the endomorphisms of the abelian group R. This observation simplifies essentially the computation with quasi-inverses.) Using the techniques in [19] the above proof can be extended naturally to non-discrete linearly compact rings. Another inspiring idea for the above proof of Hensel's Lemma comes back to Zelinsky [41]. In fact, Zelinsky's method for lifting idempotents is based also on Newton Approximation although he did not stated it explicitly. This comment holds also for the proof of Leptin's result [25] (4.1) on lifting idempotents modulo transfinitely nilpotent ideals in linearly compact rings where he left as an open problem the lifting idempotents modulo the radical. If we assume only that $f(a) = a^2 - a$ is nilpotent, then $4(a^2 - a)$ is also nilpotent. Consequently $(1 - 2a)^2$ and hence $2a - 1$ are invertible. In this case Newton Approximation gives after finite steps an idempotent $b^2 = b$ of the form $b = \sum c_i a^i, c_i \in \mathbb{Z}[b]$. This observation offers one easy way to understand the notion of SBI-rings created by Kaplansky [16]. In some sense, the proof of Theorem 2.2 is based also on Newton Approximation. Camillo's work [6] gives another approach of lifting idempotents.

Using the idea of Lang [22] we can use Newton Approximation to find roots of polynomials in the case when the derivative is not necessarily invertible, i.e., when the polynomials has simple roots which are multiple roots modulo the maximal ideal.

Proposition 4.2 *Let R be a local linearly compact ring and $f \in R[X]$. If there is an element $c \in R$ such that $f(c) \in R(f_1(c))^2$ and $Rf(c) \neq R(f_1(c))^2$ where $f_1 = f'$ is the (first) derivative of f, then there is an element $a \in R$ with $f(a) = 0$, i.e., f has a root in R.*

Proof. If $f_1(c)$ is a unit, then the assumption implies that $\bar{c}$ is a root of $\bar{f}$ where $^-$ denotes the canonical map from R to the factor field $K = R/\mathcal{M}$ of R by the maximal ideal $\mathcal{M}$. In this case the statement is exactly Hensel's Lemma. Thus we can assume without loss of generality that $f_1(c) \in \mathcal{M}$. Again as in the proof of Hensel's Lemma we denote by I_a for every $a \in R$ the ideal generated by $f(a)$. Consider the set $\mathcal{F}$ of all pairs (a, I_a) such that

1. $|f_1(c)| = |f_1(a)|$
2. $|f(a)| > 2|f_1(a)|$,

where $||$ denotes the canonical map from R to its monoid Γ of divisibility. By assumption $\mathcal{F}$ is trivially non-empty for $(c, I_c) \in \mathcal{F}$. As in the previous proof we define a partial ordering be setting $(a, I_a) \leq (b, I_b)$ if $I_a \subseteq I_b$ and $a - b \in I_b$. If $\{(a_\alpha, I_{a_\alpha})\}$ is a chain in $\mathcal{F}$, then the linear compactness of R implies $I = \cap(a_\alpha + I_{a_\alpha}) \neq \varnothing$. For any$a \in I$ and each index α we have $a - a_\alpha \in I_{a_\alpha}$. Hence Taylor Formula implies $f(a) = f(a_\alpha + (a - a_\alpha)) = f(a_\alpha) + d$ where d is a sum of products which have $a - a_\alpha$ as one factor. This implies $f(a) \in I_{a_\alpha}$. Applying again Taylor Formula for f_1 we have $f_1(a) = f_1(a_\alpha) + d_1$ where d_1 is a sum of products which have $a - a_\alpha$ as one term. Since $a - a_\alpha \in I_{a_\alpha} = Rf(a_\alpha)$ and $Rf(a_\alpha) \subsetneqq R(f_1(a_\alpha))^2$ we deduce $Rf_1(a) = Rf_1(c)$ and thus $(a, I_a)\, in\mathcal{F}$ and it is obviously a lower bound of the chain (a_α, I_{a_α}). Therefore there is a minimal pair $(a, I_a) \in \mathcal{F}$ in view of the Zorn's Lemma. We claim $I_a = 0$, i.e., $f(a) = 0$ and hence a is a root of f which completes the proof. By the choice of a we have $bf_1(a) = f(a)$ for some $b = tf_1(a), t \in R$ such that $Rb \neq Rf_1(a)$. Put $d = a - b$. By Taylor Formula we have

$$f(d) = f(a) - f_1(a)b + \ldots + (-1)^n f_n(a)b^n = (f_2(a) + \ldots)b^2 \in Rb^2.$$

This shows $f(d) = sb^2 = stbf_1(a) = stf(a) \in Rf(a)$, i.e., $I_d \in I_a$. On the other hand, applying Taylor Formula for f_1 we obtain $f_1(d) = f_1(a) + br$ for some $r \in R$. By $Rb \subsetneqq Rf_1(a)$ we have $Rf_1(d) = Rf_1(a) = Rf_1(c)$. Hence $(d, I_d) \in \mathcal{F}, (d, I_d) \lneqq (a, I_a)$ holds. Consequently the minimality of (a, I_a) implies $I_a = I_d$, i.e., $Rf(a) = Rf(d)$. This means $t(f_1(a))^2 = bf_1(a) = f(a) = uf(d) = usb^2 = ust^2(f_1(a))^2$ which implies $t(f_1(a))^2(1 - ust) = 0$. By $Rb = Rtf_1(a) \subsetneqq Rf_1(a)$ we deduce that t is contained in the maximal ideal $\mathcal{M}$ of R. Henceforth $1 - ust$ is invertible which shows $0 = t(f_1(a))^2 = f(a)$ and thus we are done. □

This result yields the following generalization of Lang's result [22], [23] for complete valuation domain of rank 1.

Corollary 4.1 *Let R be a linearly compact valuation ring and $f \in R[X]$. If there exists an element $c \in R$ such that $|f(c)| > 2|f_1(c)|$ where f_1 is the derivative of f, then f has a root in R.*

In the following statement we apply Kaplansky's idea in the proof of Theorem 5 [14] to modify a very strong condition $|f(c)| > 2|f_1(c)|$ in the previous consequence and in Proposition 4.2. As usual we denote by $||$ the canonical map from R to its monoid Γ of divisiblity.

Proposition 4.3 *Let R be a linearly compact local ring. Let $f \in R[X]$ of a degree n with the Taylor expansion $f(X+Y) = f(X) + f_1(X)Y + \ldots + f_n(X)Y^n$. If there is an element $c \in R$ with $|f(c)| > |f_1(c)|$, $|f_i(c)| > |f_1(c)|, i = 1, \ldots, n$, then f has a root in R.*

Proof. By assumption we have $f(c) = uf_1(c)$, $u \in \mathcal{M}$ where $\mathcal{M}$ is the maximal ideal of R. We denote again by I_x the ideal of R generated by an element $x \in R$. Consider now the set $\mathcal{F}$ of all pairs (a_α, I_{b_α}) with $a_\alpha \in R$, $b_\alpha \in \mathcal{M}$ such that

1. $|f_1(c)| = |f_1(a_\alpha)|$
2. $|f_i(a_\alpha)| > |f_1(c)|, i = 2, \ldots, n$
3. $f(a_\alpha) = b_\alpha f_1(a_\alpha)$.

$\mathcal{F}$ is non-empty, for $(c, I_u) \in \mathcal{F}$ by the assumption. We define again in $\mathcal{F}$ a partial ordering by setting $(a_\alpha, I_{b_\alpha}) < (a_\beta, I_{b_\beta})$ if $|b_\alpha| > |b_\beta|$ and $a_\alpha - a_\beta \in I_{b_\beta}$. Let now $\{(a_\alpha, I_{b_\alpha})\}$ be a chain in $\mathcal{F}$. The linear compactness of R implies that the intersection $\cap(a_\alpha + I_{b_\alpha}) = I \neq \varnothing$ and hence for an element $a \in I$ we have $d_\alpha = a - a_\alpha \in I_{b_\alpha}$ for all α. Consequently we have

$$f(a) = f(a_\alpha) + f_1(a_\alpha)d_\alpha + \ldots + f_n(a_\alpha)d_\alpha^n \in Rf(a_\alpha)$$

which implies $|f(a)| \geqq |f(a_\alpha)|$ and hence $|f(a)| > |f(a_\alpha)|$ for all α. Since

$$f_1(a) = f_1(a_\alpha) + 2f_2(a_\alpha)d_\alpha + \ldots + nf_n(a_\alpha)d_\alpha^{n-1}$$

we have $|f_1(a)| = |f_1(a_\alpha)| = |f_1(c)|$. Similarly we can check that $|f_i(a)| > |f_1(c)|$ for all $i = 2, \ldots, n$. Thus Zorn's Lemma implies the existence of the minimal pair (a, I_b) in $\mathcal{F}$. We claim $f(a) = 0$. For if $f(a) \neq 0$, then by the choice of (a, I_b) we have $f(a) = rf_1(a)$, $f_i(a) = v_i f_1(a)$ for some elements $r, v_2 \ldots, v_n \in \mathcal{M}$. Therefore

$$f(a-r) = f(a) - f_1(a)r + \ldots + (-1)^n f_n(a)r^n = f_1(a)r(v_1 r + \ldots + v_n r^{n-1}) \in \mathcal{M}f(a)$$

holds, i.e. we have $|f(a-r)| > |f(a)|$. Furthermore, by using Taylor Formula for $f_i(a-r), i = 1, \ldots, n$ it is easy to check $|f_1(c)| = |f_1(a)| = |f_1(a-r)|$ and $|f_i(a-r)| > |f_1(c)|$ for all $i = 2, \ldots, n$. These facts altogether imply that the pair $(a-r, I_r)$ is smaller than the pair (a, I_b) which is impossible by the minimality of the latter. Consequently we have $f(a) = 0$ and thus our proof is complete. □

We close our investigation on approximation of simple roots with an application of Hensel's Lemma. We generalize the result of Teichmüller [34], Kaplansky [15] und Cohen [40] on the existence of coefficient fields as follows. Let R be a local linearly compact ring with the maximal ideal $\mathcal{M}$. For each ordinal α we define

$$\mathcal{M}^\alpha = \begin{cases} \bigcap\limits_{\beta<\alpha} \mathcal{M}^\beta, & \alpha \text{ is a limit ordinal} \\ (\mathcal{M}^{\alpha-1})^2, & \text{otherwise.} \end{cases}$$

We say that R is *transfinite-nilpotent* if $\mathcal{M}^\alpha = 0$ for some ordinal α. The smallest ordinal with this property is called the *(transfinite-)nilpotence degree* of R. The notion of valuation fields with discrete value group created by Kaplansky in [15] can be charcterized as the corresponding valuation ring is transfinite-nilpotent. By Krull's Theorem every local noetherian ring is transfinite-nilpotent, too.

Theorem 4.1 *A local linearly compact ring R with the maximal ideal $\mathcal{M}$ and the factor field $K = R/\mathcal{M}$ admits a coefficient field F, i.e., R has a subfield F isomorphic to K if either the characteristic of K is 0 or R is transfinite-nilpotent and $pR = 0$ where the prime number p is the characteristic of K.*

Proof. The case when the characteristic of K is 0, is the obvious consequence of both Hensel's Lemma and Zorn's Lemma, as it is easy to check. For the case of the prime characteristic we prove firstly the following assertion which is a slight generalization of a result in [40] page 304. According to the foote-note in [40] the idea of the proof is due to A. Geddes.

Proposition 4.4 *Let I be an ideal contained in the maximal ideal $\mathcal{M}$ of a ring R satisfying $pR = 0$ for some prime number p. If the factor ring R/I has a coefficient field F, then R/I^2 has a coefficient field E such that E maps to F under the canonical map from R/I^2 to R/I*

Proof. Without loss of generality we can assume that $I^2 = 0$. Let A be the inverse image of F under the canonical map $R \longrightarrow R/I$. It is routine to check that A is then a local ring with the maximal ideal$I$. Thus our results folows directly from the result in [40] page 304. $\square$

Now we can complete the proof of Theorem 4.1 by transfinite induction as follows. Let α be the nilpotence degree of R. The case $\alpha = 2$ is the special case of Proposition 4.5. Assume that the statement is true for all $\beta < \alpha$. We have two cases. If α is not a limit ordinal, i.e., $\alpha = \beta + 1$ for some β, then the existence of coefficient field is a trivial application of Proposition 4.5. If α is a limit ordinal, then the linear compactness of R implies $R = \varprojlim R/\mathcal{M}^\beta$ where β runs over all ordinal smaller than α. We will construc for each $\beta < \alpha$ a coefficient field F_β such that F_β maps to F_γ under the canonical map from $R/\mathcal{M}^\beta$ to $R/\mathcal{M}^\gamma$ if $\gamma < \beta$. For $\beta = 0$ let $F_0 = K$. If $\beta = \gamma + 1$, then the existence of F_β follows from Proposition 4.5. If β is a limit ordinal, then by the linear compactness of R we have $R/\mathcal{M}^\beta = \varprojlim R/\mathcal{M}^\gamma$, $\gamma < \beta$ and hence the inverse limit $F_\beta = \varprojlim F_\gamma$ is a coefficient field of F_β satisfying the required condition. Now again the linear compactness of R implies that the inverse limit $\varprojlim F_\beta, \beta < \alpha$ is a coefficient field of R which finishes the proof of Theorem 4.1. $\square$

5 Approximation of roots: Kaplansky's method for multiple roots

In [26] Mac Lane proved that a linearly compact valued field is algebraically closed if and only if the value group is divisible and the residue field is algebraically closed. However, Mac Lane's proof is a typical existence-proof and it seems hopeless to extend his proof to linearly compact valuation rings. In this section we show that Kaplansky's method is appropriate to find multiple roots. Note that all polynomials in Kaplansky's work [14] have simple roots. The main aim of this section is the extension of Mac Lane's theorem to linearly compact valuation ring by using Kaplansky's method. In contrast to the proof of Mac Lane, Kaplansky's method offers a "quasi-constructive" way to locate roots. Moreover, we give a very effective criterion for starting the construction at an element a of a linearly compact valuation ring satisfying Hypothesis A. This criterion points out an analogy between complete real closed fields and maximally complete valued fields satisfying Hypothesis A. The basic tool for the approximation of roots is the notion of pseudo-convergent sequences. For linearly compact valued fields and valuation rings it is much easier and more convenient to deal with pseudo-convergent sequences than with filter bases of cosets of submodules, although it is wellknown that their convergences are equivalent to the linear compactness of valuation rings and hence define this class of fields and rings. However, we have to make some preparations.

Since the monoid of divisibility of a valuation ring is always a Rees factor semigroup of the non-negative cone of an ordered group by some ideal according to the result of Shores [32], Lemma 4 in [14] can be reformulized as follows

Lemma 5.1 *Let $\beta_1, \ldots, \beta_m$ be any elements of the monoid Γ of divisibility of a valuation ring, and further let $\{\gamma_\rho\}$be a well-ordered, monotone increasing sequence of elements of Γ, without a last element. Let $t_1, \ldots, t_m$ be distinct positive integers. then from some index μ on we have either*

$$\beta_i + t_i\gamma_\rho = \infty, \quad 1 \leqq i \leqq m,$$

or

$$\beta_i + t_i\gamma_\rho > \beta_k + t_k\gamma_\rho, \quad i \neq k, \quad 1 \leqq i \leqq m$$

for some integer k, $\quad 1 \leqq k \leqq m$.

Remark. The monoid of divisibility of a valuation ring can be characterized purely in the language of semigroup theory, as it is done in Clifford's work [7] (see for example [32], too).

Following Kaplansky [14] we say that a property defined on a well-ordered set holds ultimately if it holds from some point on. As a first application of Lemma 5.1 we prove

Proposition 5.1 *Let $\{a_\rho\}$ be a pseudo-convergent sequence in a valuation ring R and $\gamma_\rho = |a_{\rho+1} - a_\rho|$. If $f \in R[X]$ is of degree n such that $\{f_i(a_\rho)\}$ are ultimately either constant or pseudo-convergent for all $i = 1, \ldots, n$, then either there is an integer h, $\quad 0 < h \leqq n$ such that the inequalities*

$$\infty \neq |f_h(a_\rho)| + h\gamma_\rho < |f_i(a_\rho)| + i\gamma_\rho$$

hold ultimately for all $0 < i \neq h$ and $|f_h(a_\rho)|$ is ultimately constant or

$$|f_i(a_\rho)| + i\gamma_\rho = \infty$$

holds ultimately for all $i > 0$. In particular, $\{f(a_\rho)\}$ is either ultimately pseudo-convergent or constant.

Remark. Even in the case when the residue field of R is of prime characteristic p and R is a linearly compact valued field satisfying Hypothesis A and $\{|f(a_\rho)|\}$ is ultimately monotone increasing, h is in general not a power of p. The reason is that the sequence $\{|f_i(a_\rho)|\}$ can be monotone increasing for some derivative f_i of f.

Proof. It is well-known that $\{|b_\rho|\}$ is ultimately either constant or monotone increasing for every pseudo-convergent sequence $\{b_\rho\}$. By Lemma 5.1 in the set A of positive integers i such that $\{|f_i(a_\rho)|\}$ is ultimately constant, there is an integer h such that either

$$\forall i \in A, \quad i \neq h \quad \beta_i + i\gamma_\rho > \beta_h + h\gamma_\rho \neq \infty$$

or

$$\forall i \in A, \quad \beta_i + i\gamma_\rho = \infty$$

hold ultimately, where β_i is the common value of $\{|f_i(a_\rho)|\}$ for each $i \in A$. Note that A is non-empty because f_n is a non-zero constant. Now we can use the descending induction to prove the statement for this integer h in the first case. The assertion is clearly true for $j = n$. Assume that it is true for all $j > k > 0$. By the choice of h we have to work only in the case when $\{|f_i(a_\rho)|\}$ is ultimately monotone increasing. Therefore for sufficiently large $\sigma > \rho$ we have by Taylor Formula

$$|f_k(a_\sigma)| = |f_k(a_\rho) + (k+1)f_{k+1}(a_\rho)(a_\sigma - a_\rho) + \ldots + \binom{n}{n-k} f_n(a_\rho)(a_\sigma - a_\rho)^{n-k}| > \quad |f_k(a_\rho)|,$$

whence $|f_k(a_\rho)| = |\binom{i}{i-k} f_i(a_\rho)(a_\sigma - a_\rho)^{i-k}|$ for some $i > k$. This equality implies trivially by induction hypothesis that

$$|f_k(a_\rho)| + k\gamma_\rho \geqq |f_i(a_\rho)| + i\gamma_\rho > \beta_h + h\gamma_\rho \tag{4}$$

holds. Thus our assertion is also true for k which completes the proof for the first case. Therefore we obtain ultimately from the Taylor formula for f

$$|f(a_\sigma) - f(a_\rho)| = |f_1(a_\rho)(a_\sigma - a_\rho) + \ldots + f_n(a_\rho)(a_\sigma - a_\rho)^n| = |f_h(a_\rho)| + h\gamma_\rho,$$

which implies that $\{f(a_\rho)\}$ is ultimately pseudo-convergent.

Consider now the case when

$$\beta_i + i\gamma_\rho = \infty \tag{5}$$

hold ultimately for all $i \in A$. In this case we can see as above by using Taylor formula and the descending induction that $|f_i(a_\rho)| + i\gamma_\rho = \infty$ holds ultimately for all $i = n, n-1, \ldots, 1$ and hence $\{f(a_\rho)\}$ is ultimately constant. Thus our proof is complete. □

We are now in position to prove one of our main results concerning pseudo-convergent sequences. Note that an ultimately constant sequence is by definition not an ultimately pseudo-convergent one.

Theorem 5.1 *Let $\{a_\rho\}$ be an arbitrary pseudo-convergent sequence in a valuation ring R. Then $\{f(a_\rho)\}$ is ultimately either pseudo-convergent or constant for any polynomial $f \in R[X]$.*

Proof. The statement is obviously true for constant and linear polynomials as it is trivial to check. Assume that the statment is true for all polynomials of degree smaller than n, $n > 1$. If $f \in R[X]$ is any polynomial of degree n, then all the formal derivatives f_i, $i > 0$ of f are of degree smaller than n. Therefore all the sequences $\{f_i(a_\rho)\}$, $i > 0$ are ultimately either pseudo-convergent or constant by the induction hypothesis. Consequently, our claim follows from Proposition 5.1, which completes the proof. □

The combination of Proposition 5.1 and Theorem 5.1 imply the following consequence.

Corollary 5.1 *Let $\{a_\rho\}$ be a pseudo-convergent sequence in a valuation ring R and $\gamma_\rho = |a_{\rho+1} - a_\rho|$. If $f \in R[X]$ is a polynomial of degree n, then there is an integer $1 \leqq h \leqq n$ such that either*

$$\infty \neq |f_h(a_\rho)| + h\gamma_\rho < |f_i(a_\rho)| + i\gamma_\rho$$

hold ultimately for all $0 < i \neq h$ and $|f_h(a_\rho)|$ is ultimately constant or

$$|f_i(a_\rho)| + i\gamma_\rho = \infty$$

holds ultimately for all $i > 0$. Furthermore we have ultimately either

$$\infty \neq |f(a_\rho)| < |f_i(a_\rho)| + i\gamma_\rho \tag{6}$$

or

$$\infty \geqq |f(a_\rho)| \leqq |f_i(a_\rho)| + i\gamma_\rho \tag{7}$$

for all $i > 0$.

Proof. In view of Theorem 5.1 and Proposition 5.1 either there is a positive integer h such that

$$\infty \neq |f_h(a_\rho)| + h\gamma_\rho < |f_i(a_\rho)| + i\gamma_\rho$$

for all $i \neq h$ or

$$|f_i(a_\rho)| + i\gamma_\rho = \infty$$

hold ultimately for all $i > o$. In the first case $\{f(a_\rho)\}$ is ultimately pseudo-convergent. Therefore (6) or (7) hold concerning that $\{|f(a_\rho)|\}$ is ultimately constant or monotone increasing, respectively. In the case when

$$|f_i(a_\rho)| + i\gamma_\rho = \infty$$

holds ultimately, then $\{f(a_\rho)\}$ is ultimately constant. Thus (6) or (7) hold concerning that $\{|f(a_\rho)|\}$ is ultimetely non-zero constant or o, respectively. □

In what follows $\{a_\rho\}$ denotes always a pseudo-convergent sequence in a valuation ring R and $f \in R[X]$ is a polynomial of degree n with the Taylor expansion

$$f(X+Y) = f(X) + f_1(X)Y + \ldots + f_n(X)Y^n.$$

Put $\gamma_\rho = |a_{\rho+1} - a_\rho|$ and let $I = \{r \in R \mid \forall\rho : |r| > \gamma_\rho\}$ be the breadth of $\{a_\rho\}$. Since the main aim of this section is to find roots of f with the help of the approximation we will assume always the non-trivial case when $\{f(a_\rho)\}$ is ultimately different from 0. Of course, one can formulize the results of this section in such a form which includes the case when $\{f(a_\rho)\}$ is ultimately 0, too. However, it is in fact a tedious exercise and does not provide more information. The following simple observation plays a decisive role in the theory of pseudo-convergent sequences.

Proposition 5.2 *If a is any pseudo-limit of $\{a_\rho\}$ and put $\beta = |f(a)|$, then we have ultimately either*

$$\beta = |f(a_\rho)| < |f_i(a_\rho)| + i\gamma_\rho$$

for all $i = 1, \ldots, n$ and hence $\{|f(a_\rho)|\}$ is ultimately constant or $\{|f(a_\rho)|\}$ is ultimately monotone increasing and the inequality $\beta > |f(a_\rho)|$ holds ultimately.

Proof. Since we have ultimately

$$|f(a_\rho) - f(a)| = |f_1(a)(a_\rho - a) + \ldots + f_n(a)(a_\rho - a)^n| = |f_k(a)| + k\gamma_\rho$$

for some integer k in view of Lemma 5.1 and Taylor Formula, we obtain the assertion by observing that the right member of the equality is monotone increasing. □

The most important role of this simple claim is probably that it gives a very handy test to check whether the sequence $\{|f(a_\rho)|\}$ is ultimately monotone increasing. In particular we have the following consequence

Corollary 5.2 *If a is a pseudo-limit of $\{a_\rho\}$ and $f \in R[X]$, then $|f(a_\rho)|$ is ultimately monotone increasing if and only if the inequality*

$$max\{\frac{|f(a)| - |f_i(a)|}{i}\} > \gamma_\rho$$

holds ultimately where i runs over all indices satisfying $|f(a)| > |f_i(a)|$.

Remark. The monoid of divisiblity of a ring does in general admit neither the subtraction nor the division. Fortunately, in the case of valuation ring, as we have already noted, it is a Rees factor of a nonnegative cone of an ordered group. Since every ordered group is embedded in a divisible ordered group, the expression in the above consequence is meaningful. For the sake of simplicity and avoiding logical problems we make the following conventention concerning ∞:

$$\infty - \lambda = \infty = \frac{\infty}{n}, \quad \forall \lambda \in \Gamma \cup \{\infty\}, \quad \forall n \in \mathbb{N}$$

Proof. By Theorem 5.1 $|f(a_\rho)|$ is ultimately either constant or monotone increasing. For the sufficiency assume indirectly that $|f(a_\rho)|$ is ultimately constant. From Taylor Formula and Lemma 5.1 there is an integer k, $1 \leqq k \leqq n$ such that for all $i \neq k$ we have ultimately

$$|f(a)| < |f(a_\rho) - f(a)| = |f_1(a)(a_\rho - a) + \ldots + f_n(a)(a_\rho - a)^n| = \\ |f_k(a)| + k\gamma_\rho < |f_i(a)| + i\gamma_\rho$$

whence the inequality

$$max\{\frac{|f(a)| - |f_i(a)|}{i}\} < \gamma_\rho$$

holds which is impossible by the assumption. For the necessity we observe that Taylor Formula and Lemma 5.1 imply again ultimately the equality

$$|f(a_\rho) - f(a)| = |f_1(a)(a_\rho - a) + \ldots + f_n(a)(a_\rho - a)^n| = |f_k(a)| + k\gamma_\rho$$

for some fix integer k. Since $\{|f(a_\rho)|\}$ is ultimately monotone increasing by the assumption and the right expression of the above equality is monotone increasing, we obtain $|f(a_\rho)| < |f(a)|$. Therefore we have

$$|f(a)| > |f(a_\rho) - f(a)| = |f_k(a)| + k\gamma_\rho$$

from which the necessity follows trivially. Thus the assertion is checked. □

Corollary 5.3 *If a is a pseudo-limit of $\{a_\rho\}$ and $f \in R[X]$, then $|f(a_\rho)|$ is ultimately constant if and only if the inequality*

$$max\{\frac{|f(a)| - |f_i(a)|}{i}\} < \gamma_\rho$$

holds ultimately where i runs over all indices satisfying $|f(a)| > |f_i(a)|$.

We assume in the following two statements that the values $|f_i(a_\rho)|$ are ultimately (not necessarily non-zero) contants and we denote these common values by β_i for all $i = 1, \ldots, n$ as well as that the sequence $\{|f(a_\rho)|\}$ is ultimately monotone increasing. The following two claims are reformulisations of Lemmas 8 and 9 of Kaplansky [14].

Lemma 5.2 *There exists an integer $h, 1 \leqq h \leqq n$ such that*

$$\beta_i + t_i\gamma_\rho > \beta_h + h\gamma_\rho \neq \infty, \quad i \neq h, \quad 1 \leqq i \leqq n$$

holds ultimately and hence

$$f(a_\rho) = \beta_h + h\gamma_\rho.$$

This h is a power p if the characteristic of K is a prime number p.

Lemma 5.3 *If a is any pseudo-limit of $\{a_\rho\}$, then*

$$|f(a)| > \beta_h + h\gamma_\rho.$$

We are now able to extend Mac Lane's result to valuation rings

Theorem 5.2 *Let R be a linearly compact valuation ring with monoid Γ of divisibility and factor field $K = R/\mathcal{M}$ where $\mathcal{M}$ is the maximal ideal of R. Let $^-$ denote the canonical map from R to K. Then R is algebraically closed (i.e., every primitive polynomial $f \in R[X]$ such that $\bar{f} \in K[X]$ is not a constant polynomial, has a root in R) if and only if K is algebraically closed and Γ is divisible (i.e. $n\Gamma = \Gamma$ for every natural number $n \in \mathbb{N}$).*

Proof. The necessity is clear because the divisibility of Γ follows obviously from that fact that the polynomials $X^n - a, a \in \mathcal{M}$ have always roots in R by the assumption.

For the sufficiency let $f \in R[X]$ be an arbitrary primitive polynomial such that $\bar{f} \in K[X]$ is not a constant polynomial. Since R is linearly compact and hence Henselian, we can assume without loss of generality that f is monic. By assumption there is an element $a \in R$ such that $|f(a)| > 0$. Let n be the degree of f and consider the Taylor expansion of f

$$f(X+Y) = f(X) + f_1(X)Y + \ldots + f_n(X)Y^n.$$

By the extra condition on f, f_n is a constant polynomial which is a unit in R. Therefore there exists

$$\lambda = max\{\frac{|f(a)| - |f_i(a)|}{i}\},$$

where i runs over all integers such that $|f(a)| > |f_i(a)|$. Moreover $\lambda > 0$ holds trivially. Since Γ is divisible, there is a monotone increasing sequence $\{\gamma_\rho\}$ of elements $\gamma_\rho \in \Gamma$ with $\gamma_\rho < \lambda$.

If $\forall\rho : |a| > \gamma_\rho$, then choose for each ρ an element $a_\rho \in R$ with $|a_\rho| = \gamma_\rho$ we obtain a pseudo-convergent sequence $\{a_\rho\}$ with a as a pseudo-limit such that

$$\forall\rho : \gamma_\rho = |a_{\rho+1} - a_\rho|.$$

Thus the corresponding sequence $\{|a_\rho|\}$ is monotone increasing. In this case the approximation gives a root b such that the value $|b|$ is greater than $|a|$.

If $|a| < \gamma_\rho$ holds ultimately, then the image of a in the factor ring R/I_ρ is ultimately non-zero where for each index ρ we denote by I_ρ the ideal of R consisting of all elements r with $|r| \geqq \gamma_\rho$. In this case the linear compactness of R implies the existence of a pseudo-convergent sequence $\{a_\rho\}$ with a as a pseudo-limit such that

$$\forall\rho : \gamma_\rho = |a_{\rho+1} - a_\rho|.$$

It is easy to check that the corresponding sequence $\{|a_\rho|\}$ is ultimately the constant $|a|$. In this case the approximation gives a root b with $|b| = |a|$.

In both cases we have always a pseudo-convergent sequence $\{a_\rho\}$ with a as a pseudo-limit such that

$$\forall \rho : \gamma_\rho = |a_{\rho+1} - a_\rho|.$$

This condition implies by Corollary 5.2 that $\{|f(a_\rho)|\}$ is ultimately monotone increasing. Moreover by Corollary 5.1 we have

$$\lambda = max\{\frac{|f(a)| - |f_h(a_\rho)|}{h}\} \tag{8}$$

where h runs over those integers such that $\{|f_h(a_\rho)|\}|$ is ultimately constant and $|f(a)| > |f_h(a_\rho)|$ holds ultimately.

Now we can apply Kaplansky's method. By assumption there is an element $c \in R$ with $|c| = \lambda$. For any element $z \in R$ we have

$$f(a + cz) = \sum_{i=0}^{n} f_i(a) c^i z^i.$$

¿From the definition of λ we have for all $i = 1, \ldots, n$

$$|f_i(a)| + i|c| = |f_i(a)| + i\lambda \geqq |f(a)|.$$

Therefore for all $i = 1, \ldots, n$ there exist elements $d_i \in R$ such that $f_i(a)c^i = d_i f(a)$. Put $g = \sum_{i=0}^{n} d_i X^i \in R[X]$ with $d_0 = 1$. By assumption K is algebraically closed, thus $\bar{g}$ has a root k because it is obviously a non-constant polynomial by the choice of c. If z_1 is any representative of k in R, we then have $|f(a + cz_1)| > |f(a)|$ and a pseudo-limit $a + cz_1$ of $\{a_\rho\}$. Put now $b_0 = a$ and $b_1 = a + cz_1$. From the above process we can select a transfinite sequence of elements b_μ of R such that

1. each b_μ is a pseudo-limit of $\{a_\rho\}$
2. the sequence $\{|f(b_\mu)| = \lambda_\mu\}$ is monotone increasing
3. $|b_\nu - b_\mu| = max\{\frac{\lambda_\mu - |f_i(b_\mu)|}{i}\}$ where i runs over the indices for which $\lambda_\mu > |f_i(b_\mu)|$.

If ξ is a limit ordinal, then in view of the equality (8) and Corollary 5.1 conditions (2) and (3) show that $\{b_\mu, \mu < \xi\}$ is pseudo-convergent. The linear compactness of R implies that this sequence has pseudo-limits in R. Corollary 5.1 and Proposition 5.2 show that any pseudo-limit of this sequence is a good choice for b_ξ.
If ξ is not a limit ordinal, then we can construct $b_{\xi+1}$ as we constructed b_1 from b_0. Thus for an ordinal ψ greater then the cardinality of R we obtain that $b = b_\psi$ is a (possibly multiple) root of f which completes the proof of the theorem. □

Remark. Note that there are in general several pseudo-convergent sequences $\{a_\rho\}$ having a as their pseudo-limit which satisfy the required additional condition in the proof of Theorem 5.2. In view of computation a good choice of a sequence $\{a_\rho\}$ determines mainly the effect of the approximation.

We conclude our work with the discussion of Hypothesis A. It is clear that Hypothesis A can be stated without difficulties for valuation rings. For the sake of completeness we recall this notion as follows.

Hypothesis A. *Let R be a valuation ring with residue field K of prime characteristic p and monoid Γ of divisibility. Hypothesis A is the following pair of conditions:*

1. *Any polynomial f of the form*

$$f = X^{p^n} + a_1 X^{p^{n-1}} + \ldots + a_{n-1}X^p + a_n X + a_{n+1} \tag{9}$$

with coefficients in K, has a root in K.

2. *Γ is p-divisible, i.e. $p\Gamma = \Gamma$.*

It is obviuos that R satisfies Hypothesis A if the first condition of Hypothesis A holds for R, i.e. every polynomial f of the form (9) with coefficients in R has a root in R. In [38] Whaples proved an interesting result which says that a field of characteristic p satisfies Condition 1 of Hypothesis A if and only if it is *p-algebraically closed*, i.e., p is co-prime to the degree of all finite extensions. Using this result Vámos proved that a linearly compact valued field satisfied Hypothesis A if and only if it is p-algeraically closed. It is an exciting question to investigation the close relation between Hypothesis A and its first condition for linearly compact valued fields and rings. In what follows we devote to the solubility of polynomials over linearly compact valuation rings satisfying Hypothesis A. First we generalize a result of Kaplansky [14]. The main idea of [14] Theorem 5 says that some class of polynomials over maximally valued fields satisfying Hypothesis A, is soluble. However, the condition defining this class, is in some sense not natural. In paralell to the location of real roots we give in the next assertion a very effective criterion for the existence of roots of polynomials in a maximally complete valued fields and ring satisfying Hypothesis A which explains the natural analogs of these rings and fields to the field of real number or more generally to complete real closed fields. The class of polynomials included in Proposition 5.3 is much larger than the class of polynomials considered in the proof of [14] Theorem 5 and it is in fact their abstract description. For the sake of simplicity, we denote by $D_j f_i$ the j-th formal derivative $\binom{i+j}{j} f_{i+j}$ of the i-formal derivative f_i of a polynomial f.

Proposition 5.3 *Let R be a linearly compact valuation ring with monoid Γ of divisibility and factor field $K = R/\mathcal{M}$ where $\mathcal{M}$ is the maximal ideal of R. Let $^-$ denote the canonical map from R to K. Assume that R satisfies Hypothesis A. Let $f \in R[X]$ be a polynomial of degree n with the Taylor expansion $f(X+Y) = f(X) + f_1(X)Y + \ldots + f_n(X)Y^n$. If $a \in R$ such that*

$$\lambda = max\{\frac{|f_i(a)| - |D_j f_i(a)|}{j}\} < max\{\frac{|f(a)| - |f_i(a)|}{i}\},$$

where in left side of the inequality i runs over all integers between 1 and n while j runs over all integers with $|f_i(a)| \geq |D_j f_i(a)|$, and in the right side of the inequality i runs over all integers with $|f(a)| > |f_i(a)|$, then f has a root in R.

Proof. We sketch only the proof. If a is a root of some f_i, then $\lambda = \infty$ holds, which contradicts to the assumption because ∞ is the greatest element of Γ. Therefore $|f_i(a)| < \infty$ for all $i > 0$. The assumption implies now in view of Corollary 5.2 the existence of a pseudo-convergent sequence $\{a_\rho\}$ with a as a pseudo-limit such that $\{|f(a_\rho)|\}$ is ultimately monotone increasing, but $\{|f_i(a_\rho)|\}$ is ultimately constant for all $i > 0$. Consequently, in view of Lemmas 5.2 and 5.3 one can completes the proof by approximation as in the proof of [14] Theorem 5. □

Remark. As we have noted it already in the proof of Theorem 5.2, the value $|b|$ of a root b of f is in general greater than the value $|a|$. The situation in [14] Theorem 5 is completely different: the corresponding pseudo-convergent sequence constructed in this statement, has always a constant value.

Together with Lemmas 5.2 and 5.3 the proof of [14] Theorem 5 yields also the following technical, but very useful condition on the existence of roots.

Proposition 5.4 *Let R be a linearly compact valuation ring satisfying Hypothesis A and $f \in R[X]$.If $\{a_\rho\}$ is a pseudo-convergent in R such that $\{|f(a_\rho)|\}$ is ultimately monotone increasing and $\{|f_i(a_\rho)|\}$ is ultimately (not necessarily non-zero) constant for all i, then f has a root which is a pseudo-limit of $\{a_\rho\}$.*

Let R be a linearly compact valuation ring of prime characteristic p, i.e, $pR = 0$ holds. Let $f \in R[X]$ be any polynomial of the form (9). Since we have the equality

$$f(X+Y) = Y^{p^n} + a_1 Y^{p^{n-1}} + \ldots + a_{n-1}Y^p + a_n Y + f(X),$$

it follows that all the i-th formal derivative $f_i, (i = 1, \ldots, p^n)$ of f are constants. Therefore Proposition 5.4 implies obviously the following statement.

Proposition 5.5 *Let R be a linearly compact valuation ring of prime characteristic p with factor field K and monoid Γ of divisibility. Then R satisfies Hypothesis A if and only if every equation of the form*

$$f = X^{p^n} + a_1 X^{p^{n-1}} + \ldots + a_{n-1}X^p + a_n X + a_{n+1} = 0$$

with coefficients in R is soluble, i.e., R satisfies Condition 1 of Hypothesis A.

Remark. For linearly compact valued fields the above assertion is a consequence of the results of Whaples [38] and Vámos [35]. However, it seems hard to extend both the notion of p-algebraic closedness and the proof of Whaples and Vámos to valuation rings.

For linearly compact valued fields of characteristic 0 the p-algebraic closedness does not answer the solubility of some equations, for example equations of the form $X^p - X - b = 0$ in Hilbert's Theorem 90. In what follows we will investigate the solubility of some concrete equation with the help of Kaplansky's Methods. Since the i-th formal derivative $f_i, (i = 2, \ldots, p)$ of $f(X) = X^p - aX - b \in R[X]$ are $\binom{p}{i} X^{p-i}$, we obtain that the sequence $\{|\binom{p}{i} a_\rho^{p-i}|\}$ is ultimately constant if $\{|a_\rho|\}$ is. This observation is crucial in the proof of the following statement.

Proposition 5.6 *Let R be a linearly compact valuation ring with the factor field K of the prime characteristic p. If every equation of the form $X^p - aX - b = 0$ is soluble in K and the monoid of divisibility Γ of R is p-divisible, i.e., $p\Gamma = \Gamma$ if and only if every equation of the form $X^p - aX - b = 0$ is soluble in R.*

Proof. We have to check only the necessity. This will be done in several steps.

1. We can assume that p is a non-zero element of R, otherwise the statement follows trivially from Proposition 5.4.

2. R is perfect, i.e., every element of R has a p-root in R. Since Γ is p-divisible, it is enough to see that every invertible element has a p-root. This can be showed as in the proof of Lemma 15 [14] with the help of approximating roots by observing that all the formal derivatives of $X^p - a$, a a unit in R, are monomials $\binom{p}{i} X^{p-i}$.

3. Consider now an arbitrary equation $f(X) = X^p - aX - b = 0$ with $a, b \in R$. In what follows we can assume by Step 2 without loss of generality that $a \neq 0$. From the assumption there is an element $c \in R$ with $|f(c)| > 0$. If $d \in R$ with $|d| > O$, then

$$|f(c+d)| = |(c+d)^p - a(c+d) - b| = |f(c) - ad + pr| > O$$

holds for some element $r \in R$. This shows that $|f(c+d)| > 0$ for all $d \in R$ with $|d| > 0$.

4. We claim that there exists an element $c \in R$ such that $|f(c)| > 0$ and c is not a root of f_1. Assume indirectly that $|f(u)| > 0$ implies $f_1(u) = 0$. Step 3 together with the fact that Γ is p-divisible, ensure the existence of a pseudo-convergent sequences $\{u_\rho\}$ of arbitrarily small breadth such that $\{|u_\rho|\}$ is a constant, $|f(u_\rho)| > 0$ and hence $f_1(u_\rho) = 0$ for all ρ by the indirect assumption. Since $|f_i(u_\rho)| = |\binom{p}{i} u_\rho^{p-i}|$ is a constant for all $i > 1$, by applying Lemma 5.1 to the Taylor's formula

$$f_1(u_\sigma) = f_1(u_\rho) + 2f_2(u_\rho)(u_\sigma - u_\rho) + \ldots + pf_p(u_\rho)(u_\sigma - u_\rho)^{p-1}$$

we obtain $|p| + (p-1)\gamma_\rho = \infty$ holds ultimately where $\gamma_\rho = |u_{\rho+1} - u_\rho|$. Since Γ is p-divisible and γ_ρ can choosen be arbitrarily small, we obtain $|p| + \gamma = \infty$ for all $0 \neq \gamma \in \Gamma$. This shows that pR is the minimal ideal of R. We have here two cases.

Case 1. u_ρ is not a unit, i.e. $pu_\rho = 0$. Consequently, we $f_1(u_\rho) = p{u_\rho}^{p-1} - a = a \neq 0$, which contradicts to $f_1(u_\rho) = 0$.

Case 2. u_ρ is a unit. $f_1(u_\rho) = p{u_\rho}^{p-1} - a = 0$ implies $|p| = |a|$. On the other hand, the equalities $|f_i(u_\rho| + i\gamma_\rho = \infty$ show that $f(u_\rho)$ are equal for all ρ. However, for every $d \in R$ with $|d| > 0$ we have the equality $f(c+d) = (c+d)^p - a(c+d) - b = f(c) + d^p$ which shows that $f(u_\rho)$ are not equal for all ρ at the appropriate choice of u_ρ, because by Step 3 and the p-divisibility of Γ there exist pseudo-convergent sequences u_ρ of arbitrary small breadth. This contradiction show that it is always possible to choose an element $c \in R$ such that $|f(c)| > 0$ and c is not a root of f_1.

5. Since pseudo-limits of a pseudo-convergent sequence consisting of roots of a polynomial is always a root by Proposition 5.2, we deduce the existence of a root d of $f_1(X)$ such that

$$\delta = |c - d| = max\{|c - r| \forall r \in R : f_1(r) = 0\} \neq \infty.$$

Since

$$f(c) - f(d) = f_1(d)(c-d) + \ldots + f_n(d)(c-d)^n$$

and $f_1(d) = 0$ we obtain $\lambda = |f(c) - f(d)| \geqq 2\delta$. Since Γ is divisible, there is a monotone increasing sequence $\{\gamma_\rho\}$ of elements $\gamma_\rho \in \Gamma$ with $|c| \leqq |c-d| =$ ¡$\gamma_\rho < \lambda$. Since $|c| < \gamma_\rho$ holds for all ρ, the image of c in the factor ring R/I_ρ is non-zero where for each index ρ we define I_ρ as an ideal of R consisting of all elements r with $|r| \geqq \gamma_\rho$. In this case the linear compactness of R implies the existence of a pseudo-convergent sequence $\{c_\rho\}$ with c as a pseudo-limit such that

$$\forall \rho : \gamma_\rho = |c_{\rho+1} - c_\rho|.$$

In this case $|c_\rho| = |c|$ for all ρ. The sequence $\{|f_1(c_\rho)|\}$ is ultimately constant, otherwise a straightforward application of Kaplansky's approximation would implies an existence of a root of $f_1(X)$ which is a pseudo-limit of $\{c_\rho\}$ which is impossible by the choice of d and $\{c_\rho\}$. Now it is routine to use Kaplansky's approximation for finding to root of f which completes the proof. □

The proof of the above statement can be repeated to obtain the following result.

Proposition 5.7 *Let R be a linearly compact valuation ring with the factor field K of the prime characteristic p. If R satisfies Hypothesis A, then every equation of the form $X^{p^n} - aX - b = 0$ is soluble in R.*

References

[1] P. N. Anh, *Morita duality for commutative rings*, Comm. Algebra **18** (1990), 1781–1788.

[2] G. Azumaya, *On maximally central algebras*, Nagoya Math. J. **2** (1951), 119-150

[3] N. Bourbaki, *Commutative algebra*, Springer Verlag 1989.

[4] N. Bourbaki, *Algebra II Chapters 4-7*, Springer Verlag 1990.

[5] W. Brandal, *Commutative rings whose finitely generated modules decompose*, LNM 723, Springer Verlag 1979.

[6] V. Camillo, *What lifts*, Comm. Algebra **24** (1996), 2003–2010.

[7] A. H. Clifford, *Naturally totally ordered commutative semigroups*, Amer. J. Math. **76** (1954), 631-646.

[8] A. L. S. Corner, *Every countable reduced torsion-free ring is an endomorphism ring*, Proc. London Math. Soc. **13** (1963), 687 –710.

[9] C. Faith, *Algebra II Ring theory*, Grundlehren Math. Wiss 191, Springer Verlag 1976.

[10] L. Fuchs and L. Salce , *Modules over valuation rings*, Marcel Dekker, New York 1985.

[11] L. Fuchs and S. Shelah, *Kaplansky 's problem on valuation rings*, PAMS **105** (1989), 25–30.

[12] K. Gravett, *Note on a result of Krull*, Cambridge Philos. Soc. Proc. **52** (1956), 379.

[13] P. Hartmann, *Ein Beweis des Henselschen Lemmas für maximal bewertete Körper*, Arch. Math. **37** (1981), 163–168.

[14] I. Kaplansky, *Maximal fields with valuations*, Duke Math. J. **9** (1942), 303–321.

[15] I. Kaplansky *Maximal fields with valuations II*, Duke Math. J. **12** (1945), 243–248.

[16] I. Kaplansky, *Fields and Rings*, University of Chicago Press 1970.

[17] I. Kaplansky, *Commutative Rings*, University of Chicago Press 1974.

[18] I. Kaplansky, *Topics in commutative ring theory*, Department of Mathematics, University of Chicago 1974.

[19] A. Kertész, *Lectures on artinian rings*, Akadémiai Kiadó, Budapest 1987.

[20] G. B. Klatt and L. S. Levy, *Pre-self-injective rings*, TAMS **137** (1969), 407–419.

[21] W. Krull *Allgemeine Bewertungstheorie*, J. reine angew. Math. **167** (1932), 160–196.

[22] S. Lang, *On quasi-algebraic closure*, Annals of Mathematics **55** (1952), 373 – 390.

[23] S. Lang, *Algebra*, third Edition Addition-Wesley 1993.

[24] H. Leptin, *Linear kompakte Moduln und Ringe*, Math. Z. **62** (1955), 241–267.

[25] H. Leptin, *Linear kompakte Moduln und Ringe II*, Math. Z. **66** (1957), 289–327.

[26] S. Mac Lane, *The universality of formal power series fields*, Bull. Amer. Math. Soc. **45** (1939), 888–890.

[27] N. Narici, E. Beckenstein and G. Bachman, *Functional analysis and valuation theory*, Marcel Dekker 1971.

[28] A. Ostrowski, *Untersuchungen zur arithmetischen Theorie der Körper*, Math. Z. **39** (1935), 269–404.

[29] S. Priess-Crampe, *Angeordnete Strukturen: Gruppen, Körper, projektive Ebene*, Ergebnisse in Math. 98, Springer Verlag 1983.

[30] F. J. Rayner, *Gravett's proof of a theorem of Krull*, Quart. J. Math. Oxford **24** (1973), 409–410.

[31] O. F. G. Schilling, *The theory of valuations*, AMS Mathematical Surveys **4** 1950.

[32] T. Shores, *On generalized valuation rings*, Michigan Math. J. **21** (1974), 405–409

[33] O. Teichmüller, *Differentialrechnung bei Charakteristik p*, J. reine angew. Math. **175** (1936), 89-99.

[34] O. Teichmüller, *Diskrete bewertete perfekte Körper mit unvollkommenem Restklassenkörper*, J. reine angew. Math. **176** (1937), 141-152.

[35] P. Vámos, *Kaplansky fields and p-algebraically closed fields*, preprint.

[36] J. P. Vicknair, *On valuation rings*, Rocky Mountain J. Math **17** (1987), 55 –58.

[37] R. B. Warfield, *Purity and algebraic compactness for modules*, Pacific J. Math **28** (1969), 699–719.

[38] G. Whaples, *Galois cohomology of additive polynomials and n^{th} power mappings of fields*, Duke Math. J. **24** (1957), 143–150.

[39] G. Whaples, *Algebraic extensions of arbitrary fields*, Duke Math. J. **24** (1957), 201–204.

[40] O. Zariski and P. Samuel *Commutative Algebra II*, Springer Verlag 1975.

[41] D. Zelinsky, *Linearly compact modules and rings*, Amer. J. Math **75** (1953), 79–90.