Chapter 9

Hensel's Lemma

9.1 Equivalent forms of Hensel's Lemma

For a valued field $\mathbf{K} = (K, v)$, the property of being henselian is equivalent to a variety of criteria for polynomials $f \in \mathcal{O}_{\mathbf{K}}[X]$ to admit a zero in $\mathcal{O}_{\mathbf{K}}$. Recall that Lemma 5.7 has shown that the residue map induces a bijection between the integral roots of f in $\tilde{K}$ and the roots of $\overline{f}$ in $\tilde{\overline{K}}$ (counted with multiplicity). In view of this, the question arises under which conditions a root of $\overline{f}$ lying in $\overline{K}$ is the residue of some root of f which lies in K. If $b \in K$ is such that $\overline{b}$ is a root of $\overline{f}$, then b could be thought of as an **approximative root**. Hensel's original lemma showed how an approximative root can be **refined to a root of** f. This refinement procedure works in every spherically complete field (see Theorem 7.44 and Exercise **??**). However, Hensel's Lemma turned out to be one of a lot of properties of a valued field which are equivalent to the property of being henselian, the proof not anymore involving a refinement procedure. Note that the condition $f \in \mathcal{O}_{\mathbf{K}}[X]$ implies that also the derivative f' and all higher derivatives of f are polynomials in $\mathcal{O}_{\mathbf{K}}[X]$ (cf. Lemma 24.59). Since the residue map is a ring homomorphism, we have $\overline{f'} = \overline{f}\,'$ (and the same holds for the higher derivatives).

Theorem 9.1 *For a valued field $\mathbf{K} = (K, v)$, the property of being henselian is equivalent to each of the following properties:*

1) *If $f = 1 + X + X^2 g(X)$ with $g(X) \in \mathcal{M}_{\mathbf{K}}[X]$, then f admits a root in $\mathcal{O}_{\mathbf{K}}$ whose residue is -1.*

2) *Every monic polynomial $f = X^n + c_{n-1}X^{n-1} + \ldots + c_1 X + c_0 \in \mathcal{O}_{\mathbf{K}}[X]$ with $\overline{c}_{n-1} \neq 0$ and $\overline{c}_{n-2} = \ldots = \overline{c}_0 = 0$ admits a linear factor $X + c$ in $\mathcal{O}_{\mathbf{K}}[X]$ such that $\overline{c} = \overline{c}_{n-1}$.*

3) *For every monic polynomial $f \in \mathcal{O}_{\mathbf{K}}[X]$ the following holds: if $\overline{f}$ has a simple root $\overline{b} \in \overline{K}$, then f admits a root $a \in \mathcal{O}_{\mathbf{K}}$ such that $\overline{a} = \overline{b}$.*

4) ("Hensel's Lemma")
For every monic polynomial $f \in \mathcal{O}_{\mathbf{K}}[X]$ the following holds: if $b \in \mathcal{O}_{\mathbf{K}}$ satisfies $vf(b) > 0$ and $vf'(b) = 0$, then f admits a root $a \in \mathcal{O}_{\mathbf{K}}$ such that $\overline{a} = \overline{b}$.

5) ("Newtons Lemma")
For every monic polynomial $f \in \mathcal{O}_{\mathbf{K}}[X]$ the following holds: if $b \in \mathcal{O}_{\mathbf{K}}$ satisfies $vf(b) > 2vf'(b)$, then f admits a root $a \in \mathcal{O}_{\mathbf{K}}$ such that $v(a - b) > vf'(b)$.

6) (strong "Hensel's Lemma")

For arbitrary polynomials $f \in \mathcal{O}_K[X]$ the following holds: if there is a factorization $\overline{f} = \overline{g}\overline{h}$ such that $\overline{g}$ is relatively prime to $\overline{h}$ in the polynomial ring $\overline{K}[X]$, then there exist polynomials $g, h \in \mathcal{O}_K[X]$ reducing to $\overline{g}, \overline{h}$ respectively and such that $f = gh$ and $\deg g = \deg \overline{g}$.

7) ("Krasner's Lemma")
Assume v to be extended to the separable closure K^{sep} of K. Then for every element $a \in K^{\mathrm{sep}}$ the following holds: if $b \in K^{\mathrm{sep}} \setminus K$ satisfies

$$v(a-b) > \max\{v(a-\sigma a) \mid \sigma \in \operatorname{Gal} K \wedge \sigma a \neq a\}\ , \tag{9.1}$$

then $a \in K(b)$.

3′), 4′), 5′) *The same as 3), 4), 5) respectively, but without the hypothesis that f be monic.*

2″), 3″), 4″), 5″) *The same as 2), 3), 4), 5) respectively, but with the additional hypothesis that f be separable over K.*

Proof: During this proof, let $f = c_nX^n + c_{n-1}X^{n-1} + \ldots + c_0$ always be a polynomial in $K[X]$ with roots $a_1, \ldots, a_n \in \tilde{K}$, so that

$$f = c_n \prod_{i=1}^{n} (X - a_i)\ .$$

• We show that 6) holds in every henselian field (K, v). This field admits a unique extension of v to $\tilde{K}$ which we will again denote by v. Hence, $v\sigma = v$ for all $\sigma \in \operatorname{Gal} K$. Now let f, $\overline{g}$ and $\overline{h}$ satisfy the hypothesis of 6). Observe that the condition that $\overline{g}$ be prime to $\overline{h}$ implies that $\overline{g} \neq 0 \neq \overline{h}$, so $\overline{f} = \overline{g}\overline{h} \neq 0$. We define

$$\tilde{g} := \prod_{i\in J} (X - a_i) \in \tilde{K}[X] \quad \text{with} \quad J = \{i \mid 1 \leq i \leq n \wedge va_i \geq 0 \wedge \overline{g}(\overline{a_i}) = 0\}\ .$$

Since (K, v) is assumed to be henselian, from (7.2) we know that every $\sigma \in \operatorname{Gal} K$ induces an automorphism $\overline{\sigma} \in \overline{K}$ (the reduction of σ) via $\overline{\sigma}\,\overline{a} = \overline{\sigma a}$. For all $\sigma \in \operatorname{Gal} K$ and all $i \in J$ we have that $v\sigma a_i = va_i \geq 0$ and that $\overline{\sigma}\,\overline{a_i}$ is again a root of $\overline{g}$, which yields that σa_i is again a root of $\tilde{g}$. So the set of roots of $\tilde{g}$ is closed under the application of all $\sigma \in \operatorname{Gal} K$. Since $\overline{g}$ is assumed to be prime to $\overline{h}$, our construction yields that $\tilde{g}$ and $\tilde{h} = c_n \prod_{i\notin J}(X - a_i) \in \tilde{K}[X]$ have no common root in $\tilde{K}$. Since $f = \tilde{g}\tilde{h} \in K[X]$, part b) of Lemma 24.1 now shows that $\tilde{g}, \tilde{h} \in K[X]$.

We know from Lemma 5.7 that for every root $\overline{a}$ of $\overline{g}$ of multiplicity m, there are at least m roots (counted with multiplicity) of f whose residue is equal to $\overline{a}$. From our choice of $\tilde{g}$ it thus follows that $\overline{g}$ divides $\overline{\tilde{g}} = \prod_{i\in J}(X - \overline{a_i})$. On the other hand, $f = \tilde{g}\tilde{h}$ yields that $\overline{\tilde{g}}$ divides $\overline{f} = \overline{g}\overline{h}$. By assumption, $\overline{g}$ and $\overline{h}$ are prime to each other, so by part a) of Lemma 24.1 they have no common root in $\tilde{\overline{K}}$. Since by our choice of $\tilde{g}$, every root of $\overline{\tilde{g}}$ is also a root of $\overline{g}$, this shows that $\overline{\tilde{g}}$ must divide $\overline{g}$. This proves that $\deg \tilde{g} = \deg \overline{g}$ and that there is $c \in \mathcal{O}_{\mathbf{K}}^{\times}$ such that $g := c\tilde{g} \in \mathcal{O}_{\mathbf{K}}[X]$ reduces to $\overline{g}$. We find that $f = gh$ with $h = c^{-1}\tilde{h} \in \mathcal{O}_{\mathbf{K}}[X]$ by virtue of Gauß' Lemma.

• If 6) holds, then also 1) holds: If $f = 1 + X + X^2g(X)$ with $g(X) \in \mathcal{M}_{\mathbf{K}}[X]$, then the reduction of f is just $\overline{f} = 1 + X =: \overline{g}$, so the factor g obtained by 6) is linear, and the corresponding root of f has residue -1.

• If 1) holds, then also 5′) and thus also 5) and 5″) hold: Let f, b satisfy the assumptions of 5). We may assume that $f(b) \neq 0$ because otherwise we are done. Then it follows from the assumption $vf(b) > 2vf'(b)$ that also $f'(b) \neq 0$. We employ a substitution used by P. Ribenboim in [RIB17]. From the Taylor Expansion (24.13) we infer the equality

$$f(X+Y) = f(X) + f'(X)Y + \tilde{g}(X,Y)Y^2$$

with $\tilde{g}(X,Y) \in \mathcal{O}_{\mathbf{K}}[X,Y]$. We set $X = b$ and $Y = f(b)f'(b)^{-1}Z$ with Z a new variable. We obtain

$$f(b + f(b)f'(b)^{-1}Z) = f(b) + f(b)Z + \frac{f(b)^2}{f'(b)^2}Z^2\,\tilde{g}(b, f(b)f'(b)^{-1}Z)\ .$$

We have $\tilde{g}(b, f(b)f'(b)^{-1}Z) \in \mathcal{O}_{\mathbf{K}}[Z]$. Since $f(b)/f'(b)^2$ is an element of $\mathcal{M}_{\mathbf{K}}$ by hypothesis,

$$g(Z) := f(b)f'(b)^{-2}\tilde{g}(b, f(b)f'(b)^{-1}Z) \in \mathcal{M}_{\mathbf{K}}[Z]\ .$$

We obtain

$$h(Z) := \frac{f(b + f(b)f'(b)^{-1}Z)}{f(b)} = 1 + Z + Z^2 g(Z)\ .$$

By 1), h admits a root $a' \in \mathcal{O}_{\mathbf{K}}$ such that $\overline{a'} = -1$ and thus, $va' = 0$. Consequently, $a := b + f(b)f'(b)^{-1}a'$ is a root of f, and it satisfies $v(a-b) = v(f(b)f'(b)^{-1}a') > vf'(b)$. Furthermore, $vf'(b) \geq 0$ since $f' \in \mathcal{O}_{\mathbf{K}}[X]$ and $b \in \mathcal{O}_{\mathbf{K}}$.

• Hensel's Lemma in the versions 4), 4′), 4″) are special cases of Newton's Lemma 5), 5′), 5″) respectively, because $\overline{a} = \overline{b}$ follows from $v(a-b) > 0$.

• Property 4) and property 3) are equivalent since their hypotheses are, and similarly for 4′) and 3′) as well as for 4″) and 3″). Indeed, for $b \in \mathcal{O}_{\mathbf{K}}$, $\overline{b}$ is a simple root of $\overline{f}$ if and only if $\overline{f}(\overline{b}) = 0$ and $\overline{f}\,'(\overline{b}) \neq 0$. Since $\overline{f}(\overline{b}) = \overline{f(b)}$ and $\overline{f}\,'(\overline{b}) = \overline{f'(b)}$, this in turn is equivalent to $vf(b) > 0$ and $vf'(b) = 0$.

• If property 3) or 3′) holds, then also 2) holds, and if property 3″) holds, then also 2″) holds. For, if $f = X^n + c_{n-1}X^{n-1} + \ldots + c_1X + c_0 \in \mathcal{O}_{\mathbf{K}}[X]$ with $\overline{c}_{n-1} \neq 0$ and $\overline{c}_{n-2} = \ldots = \overline{c}_0 = 0$, then $\overline{f} = X^n + \overline{c}_{n-1}X^{n-1}$ admits $-\overline{c}_{n-1}$ as a simple root.

• If property 2) or 2″) holds, then (K,v) is henselian: Assume that (K,v) admits more than one extension of v to $\tilde{K}$. Then there is already a finite Galois extension $L|K$ admitting more than one extension of v from K to L. Hence, $\mathrm{Z} := (L|K,v)^d$ is a proper extension of K. Let $\sigma_1 = \mathrm{id}, \sigma_2, \ldots, \sigma_n$ be the distinct embeddings of Z into L. From Theorem 7.9 we know that $v\sigma_i \neq v$ on Z for all $i \neq 1$. From Lemma 6.60 we infer the existence of some $a \in L$ such that $va = 0$ and $v\sigma_i a > 0$ for all $i \neq 1$. Consequently, $a \notin K$ is the only root of value 0 of its minimal polynomial over K, and all other roots have value > 0. This minimal polynomial is thus of the form as in the hypothesis of 2), but it is irreducible. We have shown that 2) can not hold if (K,v) is not henselian.

• If (K,v) is henselian, then Krasner's Lemma holds: Assume that $a, b \in K^{\mathrm{sep}}$ satisfy the hypothesis of Krasner's Lemma. We will show that a is fixed by every automorphism $\tau \in \operatorname{Gal} K(b)$; since $K(b)$ is the fixed field of $\operatorname{Gal} K(b)$ in K^{sep} and $a \in K^{\mathrm{sep}}$, this will yield that $a \in K(b)$. Note that $\tau b = b$ because of $\tau \in \operatorname{Gal} K(b)$. Since (K,v) is henselian, the same holds for $(K(b),v)$ by Lemma 7.33, hence we have

$$v(b - \tau a) = v\tau(b-a) = v(b-a) = v(a-b) > \max\{v(a - \sigma a) \mid \sigma \in \mathrm{Gal}\,(K) \wedge \sigma a \neq a\}\ .$$

It follows that

$$v(a-\tau a) \geq \min\{v(a-b), v(b-\tau a)\} > \max\{v(a-\sigma a) \mid \sigma \in \operatorname{Gal} K \wedge \sigma a \neq a\}$$

which yields that $\tau a = a$.

• Krasner's Lemma implies property 2): Assume that f satisfies the assumption of 2). Then in view of Lemma 5.7, there is precisely one root of f in $\tilde{K}$ with residue $-c_{n-1}$, say a, and all other roots have residue zero. Observe that f is separable because $c_{n-1} \neq 0$ and not both n and $n-1$ can be divisible by the characteristic. Hence, $a \in K^{\mathrm{sep}}$. Since all conjugates of a over K are among these roots, we find that $v(a-\sigma a) = va = 0$ for all $\sigma \in \operatorname{Gal} K$ such that $\sigma a \neq a$. On the other hand, $v(a + c_{n-1}) > 0$, hence by Krasner's Lemma, $a \in K(-c_{n-1}) = K$. Now the desired linear factor is $X + c$ with $c = -a$. □

Remark 9.2 In Hensel's Lemma, the root a satisfying $\overline{a} = \overline{b}$ is uniquely determined. This follows from Lemma 5.7 and the hypothesis that $\overline{b}$ be a simple root of $\overline{f}$. The same holds for Newton's Lemma:
Let $f \in \mathcal{O}_{\mathbf{K}}[X]$ and $b \in \mathcal{O}_{\mathbf{K}}$ such that $vf(b) > 2vf'(b)$. Then there is a unique root $a \in \tilde{K}$ of f such that $v(a-b) > vf'(b)$, and if $a \neq b$, then $v(a-b) = vf(b) - vf'(b) > vf'(b)$.
This is seen as follows. From the Taylor expansion (cf. Lemma 24.59), we know that $0 = f(a) = f(b) + f'(b)(a-b) + \tilde{h}(b,a)(a-b)^2$ with $\tilde{h}(b,a) \in \mathcal{O}_{\mathbf{K}}$. Hence, $f(b) = -f'(b)(a-b) - \tilde{h}(b,a)(a-b)^2$. Now assume that $a \neq b$, which means that $v(a-b) < \infty$. From our hypothesis that $v(a-b) > vf'(b)$, it then follows that $vf'(b)(a-b) < v\tilde{h}(b,a)(a-b)^2$. Consequently, $vf(b) = vf'(b)(a-b)$, that is, $v(a-b) = vf(b) - vf'(b)$.

Suppose that also $a' \in \tilde{K}$ is a root of f with $v(a'-b) > vf'(b)$. Then $v(a'-a) \geq \min\{v(a'-b), v(a-b)\} > vf'(b)$. By the Taylor expansion, $0 = f(a') - f(a) = f'(a)(a'-a) + \tilde{h}(a,a')(a'-a)^2$ with $\tilde{h}(a,a') \in \mathcal{O}_{\mathbf{K}}$. From $v(a-b) > vf'(b)$, it also follows that $vf'(a) = vf'(b)$. Hence if $a' \neq a$, then we would have $vf'(a)(a'-a) = vf'(b)(a'-a) < v\tilde{h}(a,a')(a'-a)^2$, and it would follow that $v0 = vf'(a)(a'-a) < \infty$, a contradiction. Hence $a' = a$, as asserted.

Let us give a first application of Hensel's Lemma.

Example 9.3 Consider the Artin-Schreier polynomial $f = X^p - X - c$ over the henselian field (K, v). If $vc > 0$, then $\overline{f} = X^p - X$. This polynomial admits 0 and 1 as simple roots, hence by Hensel's Lemma, f admits roots in K with residue 0 and 1. If $\operatorname{char} \overline{K} = p$, then $X^p - X$ admits precisely the elements of the prime field $\mathbb{F}_p$ as its roots, that is, it splits completely over $\mathbb{F}_p$. In this case, Hensel's Lemma shows that f splits completely over K, the residues of its roots being precisely the elements of $\mathbb{F}_p$.

If $vc = 0$ then $\overline{f} = X^p - X - \overline{c}$ with $\overline{c} \neq 0$. If this polynomial admits a root ϑ in $\overline{K}$ and $\operatorname{char} \overline{K} = p$, then it splits completely over $\overline{K}$, the roots being $\vartheta + i$, $i \in \mathbb{F}_p$. In this case again by Hensel's Lemma, f splits completely over K, the residues of its roots being just $\vartheta + i$, $i \in \mathbb{F}_p$. On the other hand, if $\overline{f}$ does not admit a root in $\overline{K}$ then f does not admit a root in K.

By the preceding discussion, we see that an irreducible Artin-Schreier polynomial $f = X^p - X - c$ over a henselian field (K, v) with Artin-Schreier closed residue field $\overline{K}$ must satisfy $vc < 0$. The case of $vc < 0$ has already appeared in Lemma 6.39, Lemma 6.40 and Lemma 6.41 and will reappear in Examples 11.44, 11.47, 11.56 and 11.59. It will play an important role in several later chapters of this book. ♢

Recall that $\wp(X)$ denotes the Artin-Schreier polynomial $X^p - X$. Further, we set $\wp(S) := \{\wp(a) \mid a \in S\}$ for every $S \subset K$. From the above considerations, we conclude:

Lemma 9.4 *Let (K, v) be as above. Then $\mathcal{M}_{\mathbf{K}} = \wp(\mathcal{M}_{\mathbf{K}}) \subset \wp(K)$. If in addition $\overline{K}$ is Artin-Schreier closed, then $\mathcal{O}_{\mathbf{K}} = \wp(\mathcal{O}_{\mathbf{K}}) \subset \wp(K)$.*

Notes 9.5 The proof that 6) holds in a henselian field is due to Rayner [RAY2]. The proof that 2) implies "henselian" is due to Nagata [NAG1]. Ribenboim's substitution which we have employed for the proof of 1)⇒5) is a refinement of a substitution that was used by Iversen in [IVER] to prove the equivalence of Hensel's Lemma with Newton's Lemma. This trick was apparently rediscovered by L. v. d. Dries in his thesis [VDD1]. It works as follows: Suppose that f and b satisfy the hypothesis of Newton's Lemma. Set $g(X) := f(b + f'(b)X)/f'(b)^2$. Then $vg(0) = vf(b) - 2vf'(b) > 0$. Further, $g'(0) = 1$ by virtue of the chain rule, and thus, $vg'(0) = 0$. Then by Hensel's Lemma, there is $\tilde{a} \in \mathcal{M}_{\mathbf{K}}$ such that $g(\tilde{a}) = 0$. Now $a := b + f'(b)\tilde{a}$ is a root of f which satisfies $v(a - b) > vf'(b)$.

We have found the name "**Hensel's Lemma**" for property 4) as well as for property 6) or slightly different versions of them. We prefer to use it for property 4). Originally, Hensel's Lemma is the lemma which states that every complete valued field of rank 1 satisfies property 4) resp. property 6) (cf. Theorem 9.6 below).

See P. Ribenboim [RIB17] for more equivalent forms of Hensel's Lemma, alternative proofs and a different, possibly historically more adequate naming of the various forms.

Exercise 9.1 *Show the equivalence of 1) of Theorem 9.1 with a weaker version of 2) where $c_{n-1} = 1$ by means of a suitable substitution. What happens to the zeros under this substitution? Try to deduce "henselian" from this weaker version of 2); what is the problem?*

Exercise 9.2 *Show that $\mathbf{K}$ is henselian if and only if it satisfies the following property: If $f \in \mathcal{O}_{\mathbf{K}}[X]$ is irreducible over K and $\overline{f}$ is not constant, then $\overline{f} = \overline{c}\varphi^m$ where $c \in \mathcal{O}_{\mathbf{K}}$ and $\varphi \in \overline{K}[X]$ is an irreducible polynomial such that $\deg f = m \deg \varphi$.*

9.2 Henselian fields

Although we know already from Theorem 7.44 as well as from Theorem 11.27 that every spherically complete valued field is henselian, we will now give alternative proofs on the basis of Theorem 9.1. Let (K, v) be spherically complete.

First proof: We prove that Newton's Lemma (property 5') of Theorem 9.1) holds in (K, v). Take a polynomial $f \in \mathcal{O}[X]$ and $b \in \mathcal{O}$ such that $vf(b) > 2vf'(b)$. In particular, we have that $vf'(b) \neq \infty$, that is, $s := f'(b) \neq 0$. From part a) of Theorem 5.15 we know that f induces a pseudo-linear isomorphism of ultrametric spaces from $b + s\mathcal{M}$ onto $f(b) + s^2\mathcal{M}$. The hypothesis $vf(b) > 2vf'(b) = vs^2$ implies that $f(b) \in s^2\mathcal{M}$, whence $0 \in f(b) + s^2\mathcal{M}$. Thus, there is $a \in b + s\mathcal{M} \subseteq \mathcal{O}$ such that $f(a) = 0$. This element satisfies $v(a - b) > vs = vf'(b)$. □

Second proof: Part b) of Theorem 5.15 can be used to prove a generalization of Newton's Lemma: the multidimensional Newton's Lemma. We will do this in Theorem 9.11 below. The multidimensional Newton's Lemma implies the one-dimensional Newton's Lemma. □

Third proof: Following an idea of S. Prieß-Crampe, we employ the Ultrametric Fixed Point Theorem to prove that (K, v) satisfies property 1) of Theorem 9.1. So let $f = 1 + X + X^2 g(X)$ with $g(X) \in \mathcal{M}[X]$. For every $b \in K$, set

$$\Xi b \;:=\; b - f(b) \;=\; -1 - b^2 g(b)\;. \tag{9.2}$$

By assumption, all coefficients of g have value > 0, hence $vb^2g(b) > 0$ for all $b \in \mathcal{O}$. It follows that Ξ sends $\mathcal{O}$ into $-1 + \mathcal{M}$. Moreover, we can show that Ξ is contractive on $\mathcal{O}$. For $b, c \in \mathcal{O}$, we have

$$\Xi b - \Xi c = c^2 g(c) - b^2 g(b)\;.$$

Applying Lemma 24.59 with $R = \mathcal{M}$ to the polynomial $X^2g(X) \in \mathcal{M}[X]$, we find that there exists $G(X,Z) \in \mathcal{M}[X,Z]$ such that

$$Z^2g(Z) - X^2g(X) = (Z-X)G(X,Z)\;.$$

Evaluating with $Z = b$ and $X = c$, we find $G(c,b) \in \mathcal{M}$, whence $v(\Xi b - \Xi c) = v(c^2g(c) - b^2g(b)) > v(b-c)$. This proves Ξ to be contractive on $\mathcal{O}$. In view of Lemma **??** and our assumption that (K,v) be spherically complete, we may infer from the Ultrametric Fixed Point Theorem 1.12 that Ξ has a fixed point $a \in \mathcal{O}$. It satisfies $a = \Xi a = a - f(a)$ which is the same as $f(a) = 0$. Hence a is a root of f and it satisfies $\overline{a} = -1$ since it lies in $-1 + \mathcal{M}$. □

Fourth proof: The idea is to employ the results of Section 5.6 to prove that (K,v) satisfies Newton's Lemma. Take $f \in \mathcal{O}[X]$ and $b \in \mathcal{O}$ such that $vf(b) > 2vf'(b)$. For every $c \in b + f'(b)\mathcal{M}$, we set

$$\Xi c \;:=\; c - \frac{f(c)}{f'(b)}\;.$$

Since $c \in b + f'(b)\mathcal{M}$, we have that $v(c-b) > vf'(b)$. Hence,

$$\begin{aligned} vf(c) &= v(f(b) + f'(b)(c-b) + (c-b)^2H_f(b,c)) \\ &\geq \min\{vf(b), vf'(b) + v(c-b), 2v(c-b) + vH_f(b,c)\} \\ &> 2vf'(b) \end{aligned}$$

This implies that

$$v(\Xi c - c) \;=\; vf(c) - vf'(b) \;>\; vf'(b)\;,$$

and consequently,

$$\Xi c \;\in\; b + f'(b)\mathcal{M}\;.$$

We show that $\Xi : b + f'(b)\mathcal{M} \mapsto b + f'(b)\mathcal{M}$ is contractive. Take any $c, d \in b + f'(b)\mathcal{M}$. Then (5.11), with c, d in place of y, z respectively, shows that

$$\begin{aligned} v\left(\Xi c - \Xi d\right) &= v\left(c - d - \frac{f(c) - f(d)}{f'(b)}\right) \;=\; v(f(c) - f(d) - f'(b)(c-d)) - vf'(b) \\ &> vf'(b)(c-d) - vf'(b) \;=\; v(c-d)\;. \end{aligned}$$

Since the ball $b + f'(b)\mathcal{M}$ is spherically complete by Lemma 1.21, the Ultrametric Fixed Point Theorem 1.12 shows that Ξ has a fixed point a in $b + f'(b)\mathcal{M}$. This means that

$$0 \;=\; \Xi a - a \;=\; \frac{f(a)}{f'(b)}$$

and consequently, $f(a) = 0$. Since $a \in b + f'(b)\mathcal{M}$, we have that $v(a-b) > vf'(b)$. We have thus proved that (K,v) satisfies Newton's Lemma. □

Historically, the first examples for henselian fields have been the complete fields of rank 1. Recall that "rank 1" just says that the value group is archimedean. If the value group is not isomorphic to $\mathbb{Z}$, then a complete field may not be spherically complete (cf. Example 11.50). So the above proofs do not work in this case. But Theorem 5.16 shows that every complete field of rank 1 satisfies Newton's Lemma. So we have:

Theorem 9.6 *Every complete field of rank 1 is henselian. If (K,v) is a valued field of rank 1, then its completion contains a henselization.*

Here is an alternate proof for this theorem:
Let (K,v) be a complete field of rank 1, and let the notation be as in the first of the above two proofs. Since $G(X,Z)\in\mathcal{M}[X,Z]$, we may write $G(X,Z)=c\cdot\tilde{G}(X,Z)$ with $c\in\mathcal{M}$ and $\tilde{G}(X,Z)\in\mathcal{O}[X,Z]$ (for instance, c may be taken to be the coefficient of smallest value in G). Then we have $\tilde{G}(b,a)\in\mathcal{O}$ and $v(\Xi a-\Xi b)=v(ag(a)-bg(b))\geq vc+v(a-b)$. By induction on n we find that $v(\Xi^{n+1}a-\Xi^n a)\geq nvc+v(\Xi a-a)\geq nvc$. Since vK is archimedean, for every $\alpha\in vK$ there is some $n\in\mathbb{N}$ such that $nvc>\alpha$. Now we can use Lemma 1.13 to infer the existence of the desired fixed point. □

The fact that a valued field is henselian if and only if it satisfies some form of Hensel's Lemma as given in Theorem 9.1 can also be used for the construction of new classes of henselian fields which may not even be complete.

Lemma 9.7 *Take an ascending chain $((K_\nu,v))_{\nu<\lambda}$ of henselian fields. Then their union (K,v) is again a henselian field.*

Proof: Clearly, the union is a valued field. In order to show that it is henselian, it suffices to prove that it has property 4) of Theorem 9.1, that is, it satisfies Hensel's Lemma. Let $f\in K[X]$ and $b\in K$ satisfy the assumptions of Hensel's Lemma. Since $K=\bigcup_{\nu<\lambda}K_\nu$, there must be some $\nu_f<\lambda$ such that K_{ν_f} contains the finitely many coefficients of f and the element b. Since K_{ν_f} is henselian, we know from Theorem 9.1 that it satisfies Hensel's Lemma. Thus, the required root a of f can already be found in the subfield K_{ν_f} of K. This proves that (K,v) is henselian. □

Theorem 9.8 *Every Puiseux series field K over k is henselian with respect to its canonical valuation v_t.*

Proof: K_{ν_f} is henselian with respect to its t_{n_i}-adic valuation (cf. Corollary 11.28). Again by Theorem 9.1, K_{n_i} satisfies Hensel's Lemma. Thus, the required root a of f can already be found in the subfield K_{n_i} of K. This proves that (K,v_t) is henselian. □

Another way of obtaining new henselian fields, again using the equivalence of "being henselian" with "satisfying Hensel's Lemma", will be presented in Section 9.10.

9.3 The Krasner constant

Let (K,v) be any valued field. If $a\in\tilde{K}\setminus K$ is not purely inseparable over K, we choose some extension of v from K to $\tilde{K}$ and define

$$\mathrm{kras}(a,K)\ :=\ \max\{v(\tau a-\sigma a)\mid\sigma,\tau\in\mathrm{Gal}\,K\ \text{ and }\ \tau a\neq\sigma a\}\ \in\ v\tilde{K}$$

and call it the **Krasner constant of** a **over** K. Since all extensions of v from K to $\tilde{K}$ are conjugate, this does not depend on the choice of the particular extension of v. For the same reason, over a henselian field (K,v) our Krasner constant $\mathrm{kras}(a,K)$ coincides with the Krasner constant

$$\max\{v(a-\sigma a)\mid \sigma\in\mathrm{Gal}\,K \text{ and } a\neq\sigma a\}$$

which already appeared in (9.1).

The following theorem states an alternate version of the original Krasner's Lemma (property 7) of Theorem 9.1).

Theorem 9.9 *Assume that $K(a)|K$ is a separable-algebraic extension and $(K(a,b),v)$ is any (possibly transcendental) extension of (K,v) such that*

$$v(b-a) \;>\; \mathrm{kras}(a,K)\;. \tag{9.3}$$

Then for every extension of v from $K(a,b)$ to its algebraic closure $\widetilde{K(a,b)}=\widetilde{K(b)}$, the element a lies in the henselization of $(K(b),v)$ in $(\widetilde{K(b)},v)$.

Proof: Take any extension of v from $K(a,b)$ to $\widetilde{K(b)}$ and denote by $K(b)^h$ the henselization of $(K(b),v)$ in $(\widetilde{K(b)},v)$. Since a is separable-algebraic over K, it is also separable-algebraic over $K(b)^h$. Since for every $\rho\in\mathrm{Gal}\,K(b)^h$ we have that $\rho a=\rho|_{\tilde K}a$ and $\rho|_{\tilde K}\in\mathrm{Gal}\,K$, we find that

$$\begin{aligned}&\{v(a-\rho a)\mid \rho\in\mathrm{Gal}\,K(b)^h \text{ and } a\neq\rho a\}\\ &\subseteq\;\{v(a-\sigma a)\mid\sigma\in\mathrm{Gal}\,K \text{ and } a\neq\sigma a\}\\ &\subseteq\;\{v(\tau a-\sigma a)\mid\sigma,\tau\in\mathrm{Gal}\,K \text{ and } \tau a\neq\sigma a\}\;.\end{aligned}$$

This implies that

$$\mathrm{kras}(a,K(b)^h)\;\leq\;\mathrm{kras}(a,K)\;,$$

and consequently, $v(b-a)>\mathrm{kras}(a,K(b)^h)$. Now $a\in K(b)^h$ follows from the original Krasner's Lemma. □

9.4 The Hensel-Rychlik Property

We will now add two further properties to the list of equivalent forms of Hensel's Lemma. For an arbitrary polynomial f of degree n with roots $a_1,\ldots,a_n$ and leading coefficient c_n, we define the **discriminant of** f to be

$$\mathrm{discr}\,f\;:=\;(-1)^{\frac{n(n-1)}{2}}c_n^{2n-2}\prod_{1\leq i<j\leq n}(a_i-a_j)^2\;=\;c_n^{2n-2}\prod_{i\neq j}(a_i-a_j)\;. \tag{9.4}$$

We will also need the following notion. Two valuations v and w of a field K are called **independent** if they do not admit a non-trivial common coarsening, that is, if there is no non-trivial valuation ring $\mathcal{O}$ of K which contains both $\mathcal{O}_v$ and $\mathcal{O}_w$.

Theorem 9.10 *For a valued field $\mathbf{K} = (K, v)$, the property of being henselian is equivalent to each of the following properties:*

1) (strong "Hensel-Rychlik")

For every polynomial $f \in \mathcal{O}_{\mathbf{K}}[X]$ the following holds: if some $b \in K$ satisfies $vf(b) > v \operatorname{discr} f$, then f admits a root in K.

2) (weak "Hensel-Rychlik")

For every monic polynomial $f \in \mathcal{O}_{\mathbf{K}}[X]$ the following holds: if some $b \in K$ satisfies $vf(b) > v \operatorname{discr} f$, then f admits a root in K.

Proof: Every henselian field (K, v) satisfies property **1)**:
The following proof is due to E. Ehrhardt, a student of F. Lorenz (cf. [LOR2], pp. 112–113).

Assume that f satisfies the assumption of **1)**, and let c denote its leading coefficient. Let $n = \deg f$ and $a_1, \ldots, a_n$ be the roots of f in $\tilde{K}$, enumerated such that $va_1 \leq va_2 \leq \ldots \leq va_n$. Choose $m \leq n$ such that $a_{m+1}, \ldots, a_n$ are precisely the integral roots of f. Since f has integral coefficients, it follows from part b) of Lemma 5.6 that $vc + va_1 + \ldots + va_m \geq 0$. Hence, $vc + va_1 + \ldots + va_i \geq vc + va_1 + \ldots + va_m \geq 0$ for every i. We write

$$\operatorname{discr} f = \prod_{i=2}^{n} d_i \quad \text{with} \;\; d_i := c^2 \prod_{j=1}^{i-1} (a_i - a_j)^2 \; .$$

From our enumeration of the roots it follows that $v(a_i - a_j) \geq va_j$ for every $j \leq i$. Hence for every i,

$$vd_i \;=\; 2vc + \sum_{j=1}^{i-1} 2v(a_i - a_j) \;\geq\; 2(vc + \sum_{j=1}^{i-1} va_j) \;\geq\; 0 \; . \tag{9.5}$$

Further, for every i and k such that $1 \leq i, k, \leq n$,

$$vc + \sum_{\substack{j=1\\ j\neq k}}^{i-1} v(a_i - a_j) \;\geq\; vc + \sum_{\substack{j=1\\ j\neq k}}^{i-1} va_j \;\geq\; vc + \sum_{j=1}^{m} va_j \;\geq\; 0 \, , \tag{9.6}$$

because $a_1, \ldots, a_m$ are precisely the roots of negative value by assumption.

Now suppose that there are two indeces $k < \ell$ such that $v(b - a_k) = v(b - a_\ell) \geq v(b - a_i)$ for every i. Note that $v(a_\ell - a_i) \geq v(b - a_i)$ for every i. From (9.6), we obtain that

$$vd_\ell \;=\; v(c \prod_{\substack{j=1\\ j\neq k}}^{\ell-1} (a_\ell - a_j)) + v(c(a_\ell - a_k)^2 \prod_{\substack{j=1\\ j\neq k}}^{\ell-1} (a_\ell - a_j)) \;\geq\; v(c(a_\ell - a_k)^2 \prod_{\substack{j=1\\ j\neq k}}^{\ell-1} (a_\ell - a_j))$$

and for $i > \ell$,

$$vd_i \;=\; v(c \prod_{j=1}^{i-1} (a_i - a_j)) + v(c \prod_{\substack{j=1\\ j\neq \ell}}^{i-1} (a_i - a_j)) + v(a_i - a_\ell) \;\geq\; v(a_i - a_\ell) \;=\; v(a_\ell - a_i) \; .$$

With these inequalities and with $vd_i \geq 0$ for $i < \ell$, which we know from (9.5), we deduce:

$$\begin{aligned} v \operatorname{discr} f \;&\geq\; v \prod_{i \geq \ell} d_i \;\geq\; v(c(a_\ell - a_k)^2 \prod_{\substack{j=1\\ j\neq k,\ell}}^{n} (a_\ell - a_j)) \\ &\geq\; v(c(b - a_k)(b - a_\ell) \prod_{\substack{j=1\\ j\neq k,\ell}}^{n} (b - a_j)) \;=\; vf(b) \; . \end{aligned}$$

But this contradicts our assumption that $vf(b) > v\mathrm{discr}\, f$. Hence, there is an index k such that $v(a_k - b) > v(a_i - b)$ for all $i \neq k$. All conjugates of a_k over K are among the a_i. Since $v\,\mathrm{discr}\, f < \infty$, we have that $a_i \neq a_j$ for $i \neq j$. This shows that a_k is separable over K. Since Krasner's Lemma holds in every henselian field by Theorem 9.1, we find that $a_k \in K(b) = K$, which shows that f admits a root in K.

The implication **1)⇒2)** is trivial. Now assume that (K, v) is not henselian. We wish to show that **2)** does not hold. Let $L|K$ be a finite Galois extension which admits at least two extensions of v from K to L. Let $v_1, \ldots, v_n$ be all extensions of v to L. If w_i is a coarsening of v_i then let $\overline{w}_i$ denote the valuation induced by v_i on Lw_i (such that $v_i = w_i \circ \overline{w}_i$), let w denote the restriction of w_i to K and $\overline{w}$ the restriction of $\overline{w}_i$ to Kw. We are going to prove that there is some i and a coarsening w_i of v_i such that all extensions of $\overline{w}$ from Kw to Lw_i are independent.

If already $v_1, \ldots, v_n$ are independent, then there is nothing to show. Otherwise, we have a nonempty set $\mathfrak{O}$ of all non-trivial valuation rings of L which contain at least two valuation rings $\mathcal{O}_{v_i}$ and $\mathcal{O}_{v_j}$. Observe that the intersection over every descending chain of rings $\mathcal{O}_\nu$, $\nu < \lambda$, ($\mathcal{O}_\nu \supset \mathcal{O}_\mu$ for $\nu \leq \mu$) is again a member of $\mathfrak{O}$. Indeed, it is a valuation ring since every intersection over a descending chain of valuation rings is a valuation ring, and it satisfies the above condition since there are only finitely many valuation rings $\mathcal{O}_{v_i}$. By Zorn's Lemma, it follows that $\mathfrak{O}$ has minimal elements. Let $\mathcal{O}$ be such a minimal element in $\mathfrak{O}$. After a suitable renumbering, we can assume that precisely $\mathcal{O}_{v_1}, \ldots, \mathcal{O}_{v_m}$ are the valuation rings contained in $\mathcal{O}$ $(1 < m \leq n)$. Let w_1 be the common coarsening of $v_1, \ldots, v_m$ such that $\mathcal{O} = \mathcal{O}_{w_1}$, and let $\overline{w}_1, \ldots, \overline{w}_m$ be the valuations induced by $v_1, \ldots, v_m$ on Lw_1. Further, let w be the restriction of w_1 to K and let $\overline{w}$ denote the valuation induced by v on Kw. Then $\overline{w}_1, \ldots, \overline{w}_m$ are all extensions of $\overline{w}$ from Kw to Lw_1. Indeed, if $\overline{w}'$ is an extension of $\overline{w}$ from Kw to Lw_1, then $v' = w_1 \circ \overline{w}'$ is an extension of v from K to L which satisfies $\mathcal{O}_{v'} \subset \mathcal{O}$ and thus, $v' = v_i$ for some $i \leq m$. By the minimality of $\mathcal{O}$, there is no valuation ring properly contained in $\mathcal{O}$ which contains more than one of these rings. Consequently, there is no proper valuation ring of Lw_1 which contains more than one $\mathcal{O}_{\overline{w}_i}$. This shows that the extensions $\overline{w}_1, \ldots, \overline{w}_m$ are independent, and our claim is proved.

Note that by Lemma 6.61, the extension $Lw_1|Kw$ is normal. Let Z denote the decomposition field of $(Lw_1|Kw, \overline{w}_1)$. Since $m \geq 2$, Z is a proper extension of Kw. Since it is a separable extension, we can choose a primitive element $\overline{a}$ such that $\mathrm{Z} = Kw(\overline{a})$. W.l.o.g. we can choose $\overline{a}$ such that $\overline{w}_i\overline{a} \geq 0$ for all i. Let $\overline{f} \in \mathcal{O}_{(Kw,\overline{w})}[X]$ be the minimal polynomial of $\overline{a}$ over Kw. Set $\alpha := \overline{w}\,\mathrm{discr}\,\overline{f}$. Further, let $\sigma_1 = \mathrm{id}, \sigma_2, \ldots, \sigma_m$ be the distinct embeddings of Z into L. Since the valuations $\overline{w}_1, \ldots, \overline{w}_m$ are independent on Lw_1, it follows that also the extensions of $\overline{w}$ from Kw to Z are independent. From Theorem 7.9 we can thus infer that $\overline{w}_1\sigma_i \neq \overline{w}_1$ on Z for all $i \neq 1$. From the Strong Approximation Theorem (cf. [ZA-SA2], §10, Theorem 18′) we infer the existence of some $\overline{b}_1 \in \mathrm{Z}$ such that such that $\overline{w}_1(\overline{a} - \overline{b}_1) > \alpha$ and $\overline{w}_1\sigma_i\overline{b}_1 > \alpha$ for all $i \neq 1$. For $\overline{b} := \mathrm{Tr}_{\mathrm{Z}|Kw_1}(\overline{b}_1)$ we find:

$$\overline{w}_1(\overline{a} - \overline{b}) \;=\; \overline{w}_1 \left(\overline{a} - \overline{b}_1 - \sum_{i>1} \sigma_i \overline{b}_1 \right) \;\geq\; \min\{\overline{w}_1(\overline{a} - \overline{b}_1), \overline{w}_1\sigma_i\overline{b}_1 \mid i = 2, \ldots, m\} \;>\; \alpha\,.$$

Since $\overline{w}_1\overline{b} \geq 0$ and thus also $\overline{w}_1(\sigma_i\overline{a} - \overline{b}) \geq 0$ for all i, we find that

$$\overline{w}_1\overline{f}(\overline{b}) = \overline{w}_1(\overline{a} - \overline{b}) + \sum_{i>1} \overline{w}_1(\sigma_i\overline{a} - \overline{b}) > \alpha = \overline{w}\,\mathrm{discr}\,\overline{f}\,.$$

Now choose a monic polynomial $f \in \mathcal{O}_{\mathbf{K}}[X]$ whose w-reduction fw is $\overline{f}$. Further, choose $b \in K$ such that $bw = \overline{b}$. It follows that $vf(b) > v\operatorname{discr} f$. But f is irreducible over K since $\overline{f}$ is irreducible over Kw. This proves that **2)** does not hold. □

Note that in contrast to Newton's Lemma, the Hensel-Rychlik property gives a lower bound for the value of $vf(b)$ which does not depend on the approximative root b. Property **2)** is the version of the "Hensel-Rychlik" property used by Ax and Kochen in their important paper [AX–KOC1].

9.5 The multidimensional Newton's Lemma

In this section, we will use to show that a multidimensional Newton's Lemma holds in every spherically complete valued field. From this we will deduce that it also holds in every henselian field.

Theorem 9.11 *Let (K, v) be a spherically complete valued field. Then (K, v) satisfies the multi-dimensional Newton's Lemma:*
Let $f = (f_1, \ldots, f_n)$ be a system of n polynomials in n variables with coefficients in $\mathcal{O}$. Assume that $b \in \mathcal{O}^n$ is such that $vf(b) > 2v \det J_f(b)$. Then there exists a unique $a \in \mathcal{O}^n$ such that $f(a) = 0$ and $v(a - b) = vJ_f^*(b)f(b) - v\det J_f(b) > v\det J_f(b)$.

Proof: The inequality $vf(b) > 2v\det J_f(b)$ implies that $s := \det J_f(b) \neq 0$. Hence by Theorem 5.15, J^*f induces an isomorphism of ultrametric spaces from $b + s\mathcal{M}^n$ into $J^*f(b) + s^2\mathcal{M}^n$, where $J^* = J_f^*(b)$. Since $vf(b) > vs^2$, we have that $f(b) \in s^2\mathcal{M}^n$ and hence also $J^*f(b) \in s^2\mathcal{M}^n$ (since $J^* \in \mathcal{O}^{n\times n}$). That is, $J^*f(b) + s^2\mathcal{M}^n = s^2\mathcal{M}^n$. Therefore, $0 \in J^*f(b) + s^2\mathcal{M}^n$. Since J^*f induces a bijection from $b + s\mathcal{M}^n$ onto $J^*s^{-2}f(b) + \mathcal{M}^n$, there is a unique $a \in b + s\mathcal{M}^n$ such that $J^*f(a) = 0$. Since J^* is invertible, we have that $f(a) = 0 \Leftrightarrow J^*f(a) = 0$. Hence, a is the unique element in $b+s\mathcal{M}^n$ such that $f(a) = 0$. We have that $v(a - b) = v\left(J_f^*(b)f(a) - J_f^*(b)f(b)\right) - v\det J_f(b) = vJ_f^*(b)f(b) - v\det J_f(b) > v\det J_f(b)$. □

Note that like in the one-dimensional case, also in the multi-dimensional case the proof of Newton's Lemma can be reduced by transformation to a simpler case where we would in fact obtain the identity as a pseudo-companion. But as we have already shown that even in the general case we can derive suitable pseudo-linear maps from f, it is much easier to employ them directly in the proof of the multidimensional Newton's Lemma.

We are now going to show that the multidimensional Newton's Lemma holds in every henselian field. Beforehand, we need the following lemma. See [LANG3], Chapter X, §7, Proposition 8 for its proof, which uses the theory of derivations.

Lemma 9.12 *Let $x_1, \ldots, x_n$ be elements in an arbitrary extension field of the field K. Suppose that there are n polynomials $f_1, \ldots, f_n \in K[X_1, \ldots, X_n]$ such that*
1) $f_i(x_1, \ldots, x_n) = 0$ for $i = 1, \ldots, n$,
2) $J_{\underline{f}}(x_1, \ldots, x_n) \neq 0$.
Then the elements $x_1, \ldots, x_n$ are separable algebraic over K.

Theorem 9.13 *A valued field (K,v) is henselian if and only if it satisfies the multidimensional Newton's Lemma.*

Proof: $\Rightarrow$: Let (K,v) be henselian. Take (L,v) to be a maximal immediate extension of (K,v). Then (L,v) is spherically complete. By the foregoing theorem, (L,v) satisfies the multidimensional Newton's Lemma. Denote by $\mathcal{O}$ the valuation ring of K, and by $\mathcal{O}_L$ that of L. Now assume that the hypothesis of the multidimensional Newton's Lemma is satisfied by a system f of polynomials with coefficients in $\mathcal{O}$ and by $b \in \mathcal{O}^n$. It follows that there is a unique $a = (a_1, \ldots, a_n) \in \mathcal{O}_L^n$ such that $f(a) = 0$ and $v(a-b) > v \det J_f(b)$. From the latter, it follows that $v \det J_f(a) = v \det J_f(b)$ and in particular, $\det J_f(a) \neq 0$. Now the previous lemma shows that the elements $a_1, \ldots, a_n$ are separable algebraic over K. On the other hand, for every $\sigma \in \mathrm{Aut}\,(\tilde{K}|K)$, the element $\sigma a = (\sigma a_1, \ldots, \sigma a_n)$ satisfies $f(\sigma a) = \sigma f(a) = 0$ and $v(\sigma a - b) = \min_i v(\sigma a_i - b_i) = \min_i v\sigma(a_i - b_i) = \min_i v(a_i - b_i) = v(a-b) > v \det J_f(b)$ (note that $v\sigma = v$ because (K,v) is henselian). By the uniqueness of a, it follows that $\sigma a = a$ for every $\sigma \in \mathrm{Aut}\,(\tilde{K}|K)$, that is, $a \in K^n$, as required.

$\Leftarrow$: If $n = 1$, then $\det J_f(b) = f_1'(b_1)$, and the assertion is precisely the assertion of the one-dimensional Newton's Lemma. Hence the multidimensional Newton's Lemma implies that (K,v) is henselian. □

9.6 The Implicit Function Theorem

Using the multidimensional Newton's Lemma, one can prove the multidimensional **Implicit Function Theorem**:

Theorem 9.14 *Take a henselian field (K,v) and polynomials*

$$f_1, \ldots, f_n \in \mathcal{O}[X_1, \ldots, X_m, Y_1, \ldots, Y_n] \quad \textit{with } m > 0.$$

Set $Z = (X_1, \ldots, X_m, Y_1, \ldots, Y_n)$ and

$$J(Z) := \begin{pmatrix} \frac{\partial f_1}{\partial Y_1}(Z) & \ldots & \frac{\partial f_1}{\partial Y_n}(Z) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial Y_1}(Z) & \ldots & \frac{\partial f_n}{\partial Y_n}(Z) \end{pmatrix} .$$

Assume that $f_1, \ldots, f_n$ admit a common zero

$$z \;=\; (x_1, \ldots, x_m, y_1, \ldots, y_n) \in \mathcal{O}^{m+n}$$

and that the determinant of $J(z)$ is nonzero. Then for all $(x_1', \ldots, x_m') \in \mathcal{O}^m$ with $v(x_i - x_i') > 2v \det J(z)$, $1 \leq i \leq m$, there exists a unique tuple $(y_1', \ldots, y_n') \in \mathcal{O}^n$ such that $(x_1', \ldots, x_m', y_1', \ldots, y_n')$ is a common zero of $f_1, \ldots, f_n$, and

$$\min_{1\leq i\leq n} v(y_i - y_i') \;\geq\; \min_{1\leq i\leq m} v(x_i - x_i') - v \det J(z) \;.$$

Proof: We observe that the entries of $J(Z)$ and its adjoint matrix $J^*(Z)$ are polynomials in $X_1,\ldots,X_m,Y_1,\ldots,Y_n$ with coefficients in $\mathcal{O}$. We set $b=(x'_1,\ldots,x'_m,y_1,\ldots,y_n)$. Then $J^*(b)$ is the adjoint matrix for $J(b)$, and the entries of both matrices lie in $\mathcal{O}$. In particular, this implies that $vJ^*(b)f(b)\geq vf(b)$.

By assumption, $f_i(z)=0$ for $1\leq i\leq m$. Hence, the condition $v(x_i-x'_i)>2\det vJ(a)$, $1\leq i\leq m$, will imply that

$$\begin{aligned} vf_i(b) &= v\left(f_i(x'_1,\ldots,x'_m,y_1,\ldots,y_n)-f(x_1\ldots,x_m,y_1,\ldots,y_n)\right)\\ &\geq \min_{1\leq i\leq m} v(x_i-x'_i) \;>\; 2v\det J(x_1\ldots,x_m,y_1,\ldots,y_n)\\ &= 2v\det J(x'_1,\ldots,x'_m,y_1,\ldots,y_n) \;=\; 2v\det J(b) \end{aligned}$$

for $1\leq i\leq m$. In particular, $\det J(b)\neq 0$. Hence by the multidimensional Newton's Lemma (which holds by Theorem 9.13), there is a unique common zero $(y'_1,\ldots,y'_n)\in\mathcal{O}^n$ of the polynomials $f_i(x'_1,\ldots,x'_m,Y_1,\ldots,Y_n)$, $1\leq i\leq n$, such that

$$\begin{aligned} \min_{1\leq i\leq n} v(y_i-y'_i) &\geq vJ^*(b)f(b)-v\det J(b) \;=\; vJ^*(b)f(b)-v\det J(z)\\ &\geq \min_{1\leq i\leq m} vf_i(b)-v\det J(z)\\ &\geq \min_{1\leq i\leq m} v(x_i-x'_i)-v\det J(z)\;. \end{aligned}$$

This proves our assertion. □

Exercise 9.3 *Let (K,v) be a valued field of rank 1.*

a) Let $0<\alpha\in vK$ and prove that (K,v) is henselian if and only if it satisfies the following condition

Every monic polynomial $f=X^n+c_{n-1}X^{n-1}+\ldots+c_1X+c_0\in K[X]$ with $vc_{n-1}=0$ and $vc_i>\alpha$, $0\leq i\leq n-2$, admits a linear factor $X+c$ in $\mathcal{O}_{\mathbf{K}}[X]$ such that $\bar{c}=\bar{c}_{n-1}$.

(Hint: use the Approximation Theorem).

b) Prove that (K,v) is henselian if and only if it satisfies the Implicit Function Theorem.

c) Find a transformation that reduces the nultidimensional Newton's Lemma to a nultidimensional Hensel's Lemma.

9.7 An infinite-dimensional Implicit Function Theorem

From our result in Section 1.7 it follows that an infinite power Y^I of an ultrametric space Y can be equipped with an ultrametric u^I (analogous to the minimum valuation) if the value set uY is well ordered. In this case, if (Y,u) is spherically complete, then so is (Y^I,u^I). So we obtain the following corollary to our Main Theorem 1.26 and to Proposition 2.41:

Corollary 9.15 *a) Take two ultrametric spaces (Y,u) and (Y',u'), and an arbitrary index set I. Assume that uY is well ordered, $f:\;Y^I\rightarrow Y'$ is immediate and that (Y,u) is spherically complete. Then f is surjective and (Y',u') is spherically complete.*

b) Take two valued abelian groups (G,v) and (G',v'), and an arbitrary index set I. Assume that vG is well ordered, $b\in G^I$, B is a ball around 0 in G^I, $f:\;G^I\rightarrow G'$ has a pseudo-companion on $b+B$, and that (G,v) is spherically complete. Then f is surjective and (G',v') is spherically complete.

In the case of a valued field (K,v) we cannot do the same since if the valuation is non-trivial, the value group will not be well ordered. If the valuation is not discrete (i.e., its value group is not isomorphic to $\mathbb{Z}$), then not even the value set $v\mathcal{O} := v(\mathcal{O}\setminus\{0\})$ of the valuation ring is well ordered. But we may be interested in infinite systems of polynomials with coefficients in a subring R of $\mathcal{O}$ with well ordered value set $vR := v(R\setminus\{0\})$. We set $\mathcal{M}_R := \{a \in R \mid va > 0\}$.

Note that (R,v) is not necessarily spherically complete, even if (K,v) is. So we will assume that (R,v) is spherically complete.

We generalize the definitions of **minimum valuation** and of **pseudo linear map** in the obvious way. If $a = (a_i)_{i\in I} \in R^I$, then $va := \min_{i\in I} va_i$. If $Y \subseteq R^I$, $0 \neq s \in R$ and f a map from Y into R^I, then f is pseudo-linear with pseudo-slope s if (**??**) holds for all $y,z \in Y$ such that $y \neq z$. We then have the following application of Proposition 2.41 together with Proposition 1.11:

Proposition 9.16 *Take $b \in R^I$ and B a ball in (R^I,v) around 0. Assume that $f : b+B \rightarrow R^I$ is pseudo-linear with pseudo-slope $s \in R$ and that (R,v) is spherically complete. Then f is an isomorphism of ultrametric spaces from $b+B$ onto $fb+sB$.*

If the map is given by an infinite system of polynomials $f = (f_k)_{k\in I}$ in infinitely many variables X_i , $i \in I$, and with coefficients in R, then we may consider the infinite matrix $J_f(b) \in R^{I\times I}$. Note that this matrix has only finitely many non-zero entries in every row. We denote by $R^{(I\times I)}$ all matrices in $R^{I\times I}$ which have only finitely many non-zero entries in every row and every column. If every variable appears only in finitely many f_k, then $J_f(b) \in R^{(I\times I)}$.

If we assume that R is spherically complete, we can consider a larger class of matrices. We denote by $R^{((I\times I))}$ all matrices in $R^{I\times I}$ which for each $\alpha \in vR$ have only finitely many entries of value $\leq \alpha$ in every row and every column. For every two matrices in $R^{((I\times I))}$, their product can be computed and lies again in $R^{((I\times I))}$. It is possible that $J_f(b) \in R^{((I\times I))}$ even when there are variables that appear in infinitely many f_k.

We define $\mathcal{M}_R^{(I\times I)}$ and $\mathcal{M}_R^{((I\times I))}$ analogously and note that $R^{(I\times I)}$, $R^{((I\times I))}$, $\mathcal{M}_R^{(I\times I)}$ and $\mathcal{M}_R^{((I\times I))}$ are all closed under matrix addition and multiplication and under scalar multiplication. Further, $R^{(I\times I)}\mathcal{M}_R^{(I\times I)} \subseteq \mathcal{M}_R^{(I\times I)}$, $\mathcal{M}_R^{(I\times I)}R^{(I\times I)} \subseteq \mathcal{M}_R^{(I\times I)}$, $R^{((I\times I))}\mathcal{M}_R^{((I\times I))} \subseteq \mathcal{M}_R^{((I\times I))}$ and $\mathcal{M}_R^{((I\times I))}R^{((I\times I))} \subseteq \mathcal{M}_R^{((I\times I))}$.

We are not able to use determinants here. Still, we can use our original approach if $J_f(b)$ has an inverse. But we can even work with less than invertibility. Given matrices M, M° in $R^{(I\times I)}$, or in $R^{((I\times I))}$ if R is spherically complete, we will say that M° is a **pseudo-inverse of** M if the matrices $MM^\circ - E$ and $M^\circ M - E$ are in $\mathcal{M}_R^{I\times I}$, where E denotes the $I \times I$-identity matrix.

Actually, we also do not need that the ring R is a subring of a valued field. It suffices to assume that it is a valued abelian group with its multiplication satisfying (V3), and that its value set is a well ordered subset of an ordered abelian group. It then follows that the value set does not contain negative elements. In particular, all entries of $M \in R^{I\times I}$ have value ≥ 0. This implies that $vMa \geq va$ for all $a \in R^I$. Since vR is well ordered, it contains a minimal positive value α_0. If M is in $\mathcal{M}_R^{(I\times I)}$ or in $\mathcal{M}_R^{((I\times I))}$, then all entries of M have value $\geq \alpha_0$. It then follows that $vMa \geq va + \alpha_0 > va$ for all $a \in R^I$.

Lemma 9.17 *Take M, M° in $R^{(I\times I)}$, or in $R^{((I\times I))}$ if R is spherically complete. Assume that M° is a pseudo-inverse of M. Then the following holds:*
1) For all $a \in R^I$, $vMa = va$ and $vM^\circ a = va$; in particular, $M, M^\circ \notin \mathcal{M}_R^{I\times I}$ and the value set vR must contain 0.
2) If M' is in $R^{(I\times I)}$, or in $R^{((I\times I))}$ respectively, such that $M' - M \in \mathcal{M}_R^{I\times I}$, then M° is also a pseudo-inverse of M'.
3) Both M and M° induce immediate embeddings of the ultrametric space R^I in itself with value map id, and the same holds on every ball around 0 in R^I.

Proof: 1): For all $a \in R^I$ we have that $v(MM^\circ a - a) = v((MM^\circ - E)a) > va$ and hence $va = vMM^\circ a \geq vM^\circ a \geq va$. It follows that equality holds everywhere, which gives $vM^\circ a = va$. Interchanging M and M°, we obtain $vMa = va$.

2): We compute: $M'M^\circ - E = (M' - M)M^\circ + MM^\circ - E \in \mathcal{M}_R^{I\times I}$, and similarly for $M^\circ M' - E$.

3): It suffices to show that for every ball B around 0 in R^I, M induces an immediate embedding of B into itself with value map id. Since $vMa = va$ for all $a \in R^I$, we have $MB \subseteq B$ and that M induces an injective map on B with value map id. As M induces a group homomorphism, we only have to show now that for every $a' \in B\setminus\{0\}$ there is $a \in B$ such that (IH1) and (IH2) of Proposition 2.36 hold for M in the place of f. As $vM^\circ a' = va'$, we have that $a := M^\circ a' \in B$. Further, $v(a' - Ma) = v(a' - MM^\circ a') = v(E - MM^\circ)a' > va'$. Finally, if $b \in B$ with $va \leq vb$, then $vMa = va \leq vb = vMb$. □

Proposition 9.18 *Assume that (R, v) is spherically complete. Take any index set I and a system of polynomials $f = (f_k)_{k\in I}$ in variables Y_i, $i \in I$, with coefficients in R. Take $b \in R^I$ and suppose that $J_f(b)$ lies in $R^{((I\times I))}$ and admits a pseudo-inverse in $R^{((I\times I))}$. Then $J_f(b)$ is a pseudo-companion of f on $b + \mathcal{M}_R^I$, and f is an isomorphism from $b + \mathcal{M}_R^I$ onto $f(b) + \mathcal{M}_R^I$ with value map id. The system f has a zero on $b + \mathcal{M}_R^I$ (which then is unique) if and only if $vf(b) > 0$.*

Proof: Since $J = J_f(b)$ has a pseudo-inverse, we know from the previous lemma that J induces an immediate embedding of $\mathcal{M}_R^I$ in itself with value map id.

Take $\varepsilon_1, \varepsilon_2 \in \mathcal{M}_R^I$. An infinite-dimensional version of the multidimensional Taylor expansion gives the infinite-dimensional analogue of (5.12) and (5.13), with $s = 1$. We obtain that for $y = b + \varepsilon_1$ and $z = b + \varepsilon_2$ in $b + \mathcal{M}_R^I$ with $y \neq z$,

$$v(f(y) - f(z) \;-\; J(y - z)) \;>\; v(y - z) \;=\; vJ(y - z)\;.$$

This proves that J is a pseudo-companion of f on $b + \mathcal{M}_R^I$. From Proposition 2.41 we infer that f induces an embedding of $b + \mathcal{M}_R^I$ in $f(b) + J\mathcal{M}_R^I \subseteq f(b) + \mathcal{M}_R^I$ with value map $\varphi = \mathrm{id}$.

The remaining assertions now follow from Proposition 2.41 and Theorem 1.26. □

Now we can prove an **infinite-dimensional Implicit Function Theorem**:

Theorem 9.19 *Take any index sets I and I' and a system of polynomials $f = (f_k)_{k\in I}$ in variables X_j, $j \in I'$, and Y_i, $i \in I$, with coefficients in R, and such that each variable*

Y_i appears in only finitely many f_k. Assume that (R,v) is spherically complete. Set $Z = (X_j, Y_i \mid j \in I', i \in I)$ and

$$J(Z) := \left(\frac{\partial f_k}{\partial Y_i}(Z)\right)_{k,i\in I} .$$

Assume that the polynomials f_k, $k \in I$, admit a common zero $z = (x_j, y_i \mid j \in I', i \in I)$ in $R^{I'\cup I}$ such that $J(z)$ admits a pseudo-inverse in $R^{((I\times I))}$. Then for all $(x'_j)_{j\in I'} \in R^{I'}$ with $v(x_j - x'_j) > 0$ there exists a unique $(y'_i)_{i\in I} \in R^I$ such that $z' = (x'_j, y'_i \mid j \in I', i \in I)$ is a common zero of the polynomials f_k , $k \in I$, and

$$\min_{i\in I} v(y_i - y'_i) \geq \min_{j\in I'} v(x_j - x'_j) .$$

Proof: We set $\tilde{z} := (x'_j, y_i \mid j \in I', i \in I)$ and observe that our condition that $v(x_j - x'_j) > 0$ implies that $v\left(\frac{\partial f_k}{\partial Y_i}(\tilde{z}) - \frac{\partial f_k}{\partial Y_i}(z)\right) > 0$. From part 2) of Lemma 9.17 it thus follows that the pseudo-inverse of $J(z)$ is also a pseudo inverse of $J(\tilde{z})$. (Note that $J(z), J(\tilde{z}) \in R^{(I\times I)}$ by our condition on the variables Y_i.)

For each $k \in I$ we set $g_k(Y_i \mid j \in I) := f_k(x'_j, Y_i \mid j \in I', i \in I)$. Further, we set $b := (y_i \mid i \in I)$. We consider the system $g = (g_k)_{k\in I}$. From Proposition 9.18 we infer that $J_g(b) = J(\tilde{z})$ is a pseudo-companion of g on $b + \mathcal{M}_R^I$. By assumption, $f_k(z) = 0$ for $k \in I$. Hence, the condition $v(x_j - x'_j) > 0$ will imply that

$$vg_k(b) \;=\; vf_k(\tilde{z}) \;=\; v(f_k(\tilde{z}) - f_k(z)) \;\geq\; \min_{j\in I'} v(x_j - x'_j) > 0 .$$

Hence $vg(b) > 0$ and by Proposition 9.18 the system g has a unique zero $a = (y'_i \mid i \in I)$ on $b + \mathcal{M}_R^I$. It satisfies

$$\min_{i\in I} v(y_i - y'_i) \;=\; v(b-a) \;=\; v(g(b) - g(a)) \;=\; vg(b) \;\geq\; \min_{j\in I'} v(x_j - x'_j) .$$

□

Remark 9.20 In our theorem we needed the assumption on the variables Y_i in order to have only finitely many non-zero polynomials in each row and each column of $J(Z)$. Without this it is not automatic that the conditions $J(z) \in R^{((I\times I))}$ and $v(x_j - x'_j) > 0$ imply that $J(\tilde{z}) \in R^{((I\times I))}$. We can drop the condition on the variables if we assume instead that $J(\tilde{z}) \in R^{((I\times I))}$ and that it has a pseudo-inverse in $R^{((I\times I))}$.

9.8 Power series maps on valuation ideals

Take any field k and any ordered abelian group G. As usual, we consider $k((G))$ with its canonical valuation $v = v_t$ and denote the valuation ideal by $\mathcal{M}$. Every power series

$$f(X) = \sum_{i\in\mathbb{N}} c_i X^i \in k[[X]] \tag{9.7}$$

defines in a canonical way a map $f : \mathcal{M} \to \mathcal{M}$ (note: $0 \notin \mathbb{N}$ in our notation). This can be shown by use of Neumann's Lemma, cf. [DMM1]. We note that for every integer $r > 1$ and every $y, z \in \mathcal{M}$,

$$v(y^r - z^r) \;>\; v(y-z) . \tag{9.8}$$

Therefore, if $c_1 \neq 0$, we have that

$$v(f(y) - f(z) - c_1(y-z)) \;=\; v\sum_{i\geq 2} c_i(y^i - z^i) \;>\; v(y-z) \;=\; vc_1(y-z) \tag{9.9}$$

because $vc_i = 0$ for all i. So we see that f is pseudo-linear with slope c_1 if $c_1 \neq 0$. By Proposition **??**, we obtain:

Theorem 9.21 *If $f : \mathcal{M} \to \mathcal{M}$ is defined by the power series (9.7), then f is an isomorphism of ultrametric spaces.*

A similar result holds for power series with generalized exponents (which for instance are discussed in [DS]). Take any subgroup G of $\mathbb{R}$ and a generalized power series of the form

$$f(X) = \sum_{i\in\mathbb{N}} c_i X^{r_i} \in k[[X^G]] \tag{9.10}$$

where r_i , $i \in \mathbb{N}$, is an increasing sequence of positive real numbers in G. Suppose that the power functions $y \mapsto y^{r_i}$ are defined on $\mathcal{M}$ for all i. Then again, the generalized power series (9.10) defines a map $f : \mathcal{M} \to \mathcal{M}$. We note that (9.8) also holds for every real number $r > 1$ for which $y \mapsto y^r$ is defined on $\mathcal{M}$. Hence if $c_1 \neq 0$ and $r_1 = 1$, then (9.9) holds, with the exponent i replaced by r_i . This shows again that f is pseudo-linear with pseudo-slope c_1 . If, however, $r_1 \neq 1$, we may think of writing $f(y) = \tilde{f}(y^{r_1})$ with

$$\tilde{f}(X) \;=\; \sum_{i\in\mathbb{N}} c_i X^{r_i/r_1} \,.$$

If the power functions $y \mapsto y^{r_i/r_1}$ are defined on $\mathcal{M}$ for all i, then $\tilde{f}$ defines a pseudo-linear map from $\mathcal{M}$ to $\mathcal{M}$ with pseudo-slope c_1 . So we obtain:

Theorem 9.22 *Suppose that the power functions $y \mapsto y^{r_i}$ and $y \mapsto y^{r_i/r_1}$ are defined on $\mathcal{M}$ for all i, and that $y \mapsto y^{r_1}$ is surjective. If $f : \mathcal{M} \to \mathcal{M}$ is defined by the power series (9.10) with $c_1 \neq 0$, then f is surjective.*

9.9 Power series maps and infinite-dimensional Implicit Function Theorems

We use again the notations and assumptions from Section 9.7. We take $R[[X_j, Y_i \mid j \in I', i \in I]]$ to be the set of all formal power series in the variables X_j, Y_i in which for every $n \in \mathbb{N}$ only finitely many of the X_j, Y_i appear to a power less than n. In the previous section, our power series had well defined values because we were operating in a power series field $k((G))$. Here, we will assume throughout that R is spherically complete. But this alone does not a priori give us well defined values of the power series on $\mathcal{M}_R^{I'\cup I}$. So we will assume that we have some canonical way to determine the value of a given power series at an element of $\mathcal{M}_R^I$. This holds for instance if vR is archimedean, i.e., is a subsemigroup of an archimedean ordered abelian group.

To every power series $g \in R[[Y_i \mid i \in I]]$ we associate its **0-linear part** L_g^0, by which we mean the sum of all of its monomials of total degree 1 and with a coefficient in R of value 0. This is a polynomial, i.e., contains only finitely many of the variables Y_i. We set $Y = (Y_i \mid i \in I)$.

Theorem 9.23 *Assume that (R,v) is spherically complete. Take any index sets I and I' and a system $f=(f_k)_{k\in I}$ where $f_k\in R[[X_j,Y_i\mid j\in I',i\in I]]$. Assume that f_k, $k\in I$, admit a common zero $z=(x,y)$, $x\in\mathcal{M}_R^{I'}$, $y\in\mathcal{M}_R^I$, such that for the map $L(Y)=L^0_{f(x,Y)}(Y):\mathcal{M}_R^I\to\mathcal{M}_R^I$ the following holds: for every $a'\in\mathcal{M}_R^I\setminus\{0\}$ there is some $a\in\mathcal{M}_R^I$ such that*

$$v(a'-La)\;>\;va'\qquad\text{and}\qquad va\;=\;va'\;.$$

Take $x'=(x'_j)_{j\in I'}\in\mathcal{M}_R^{I'}$, set $\alpha=v(x-x')$ and $g(Y)=f(x',Y)$ and suppose that for all distinct $w,w'\in B_\alpha(y)$,

$$v(gw-gw'-L(w-w'))\;>\;v(gw-gw')\;.\tag{9.11}$$

Then there exists a unique $(y'_i)_{i\in I}\in\mathcal{M}_R^I$ such that $z'=(x'_j,y'_i\mid j\in I',i\in I)$ is a common zero of f_k , $k\in I$, and

$$\min_{i\in I}v(y_i-y'_i)\geq\alpha\;.$$

Proof: Note that $L_{f(x',Y)}(Y)=L_{f(x,Y)}(Y)=L(Y)$. We claim that L is a pseudo-companion of $f(x',Y):\mathcal{M}_R^I\to\mathcal{M}_R^I$ on $B_\alpha(y)$. Condition (PC2) holds by assumption. As L is a group homomorphism, our conditions together with Proposition 2.36 show that $L:\mathcal{M}_R^I\to\mathcal{M}_R^I$ is immediate; note that (IH2) holds because if $va\leq vb$ then $vLa=va\leq vb\leq vLb$. Now the assertion of our theorem follows as in earlier proofs. □

The following version of the above theorem has a similar proof:

Theorem 9.24 *Assume that (R,v) is spherically complete. Take any index sets I and I' and a system $f=(f_k)_{k\in I}$ where $f_k\in R[X_j\mid j\in I'][[Y_i\mid i\in I]]$. Assume that f_k, $k\in I$, admit a common zero $z=(x,y)$, $x\in R^{I'}$, $y\in\mathcal{M}_R^I$, such that $L(Y)=L^0_{f(x,Y)}(Y)$ satisfies the same condition as in Theorem 9.23. Take $x'=(x'_j)_{j\in I'}\in R^{I'}$ such that $\alpha=v(x-x')>0$. Suppose that (9.11) holds for $g(Y)=f(x',Y)$. Then there exists a unique $(y'_i)_{i\in I}\in\mathcal{M}_R^I$ such that $z'=(x'_j,y'_i\mid j\in I',i\in I)$ is a common zero of the polynomials f_k , $k\in I$, and $\min_{i\in I}v(y_i-y'_i)\geq\alpha$.*

Alternatively, in order to obtain maps on all of R, one can consider convergent power series. We let $R\{\{X_j,Y_i\mid j\in I',i\in I\}\}$ be the set of all formal power series in the variables X_j,Y_i in which for every $\alpha\in vR$ only finitely many monomials have coefficients of value less than α. Again we assume that R is spherically complete. Then every convergent power series defines a map from R into R. In a similar way as before, one can prove:

Theorem 9.25 *Assume that (R,v) is spherically complete. Take any index sets I and I' and a system $f=(f_k)_{k\in I}$ where $f_k\in R\{\{X_j,Y_i\mid j\in I',i\in I\}\}$. Assume that f_k, $k\in I$, admit a common zero $z=(x,y)$, $x\in R^{I'}$, $y\in R^I$, such that $L(Y)=L^0_{f(x,Y)}(Y)$ satisfies the same condition as in Theorem 9.23. Take $x'=(x'_j)_{j\in I'}\in R^{I'}$ such that $\alpha=v(x-x')>0$. Suppose that (9.11) holds for $g(Y)=f(x',Y)$. Then there exists a unique $(y'_i)_{i\in I}\in R^I$ such that $z'=(x'_j,y'_i\mid j\in I',i\in I)$ is a common zero of the polynomials f_k , $k\in I$, and $\min_{i\in I}v(y_i-y'_i)\geq\alpha$.*

9.10 Relatively algebraically closed subfields

From the already given examples of henselian fields, we obtain an abundance of other examples by the fact that every relatively separable-algebraically closed subfield of a henselian field is again henselian. This was proved in Corollary 7.38. In the following lemma, we refine this result and give a new proof. We will make use of the fact that a field is henselian if and only if it satisfies Hensel's Lemma restricted to separable polynomials. The latter is property 4″) of Theorem 9.1.

Lemma 9.26 *Assume (L,v) to be henselian and K to be relatively separable-algebraically closed in L. Then (K,v) is henselian too. Further, $\overline{K}$ is relatively separable-algebraically closed in $\overline{L}$. If in addition $\overline{K}=\overline{L}$, then the torsion subgroup of vL/vK is a p-group, with p the characteristic exponent of $\overline{K}$.*

Proof: By Theorem 9.1, a valued field is henselian if and only if it satisfies Hensel's Lemma in the version 4″). If (L,v) is henselian and $f\in K[X]$ is separable over K and satisfies the assumptions of 4″), then there exists a root $a\in L$ according to 4″). The minimal polynomial of a over K divides f and is thus separable over K. Since K is assumed to be relatively separable-algebraically closed in L, it follows that $a\in K$. Hence also (K,v) has property 4″), that is, (K,v) is henselian. (Note that a proof using the uniqueness of extensions which serves as the definition for the property "henselian" is not at all as obvious as this proof which uses Hensel's Lemma. For the proof of the fact that every algebraic extension of a henselian field is henselian it is just the opposite case.)

Now let $\overline{a}\in\overline{L}$ be separable-algebraic over $\overline{K}$. We choose a monic polynomial $g(X)\in K[X]$ whose reduction $\overline{g}(X)\in\overline{K}[X]$ modulo v is its irreducible polynomial over $\overline{K}$. Then $\overline{a}$ is a simple root of $\overline{g}$. Hence by Hensel's Lemma, there is a root $a\in L$ of g whose residue is $\overline{a}$. As all roots of $\overline{g}(X)$ are distinct, Lemma 5.7 together with part a) of Lemma 5.6 shows that also all roots of $\overline{g}(X)$ are distinct. Thus, a is separable-algebraic over K. By our assumption, it follows that $a\in K$, showing that $\overline{a}\in\overline{K}$. This proves that $\overline{K}$ is relatively separable-algebraically closed in $\overline{L}$.

Assume that $\overline{K}=\overline{L}$ and let $\alpha\in vL$ and $n\in\mathbb{N}$ not divisible by p such that $n\alpha\in vK$. Choose $a\in L$ and $b\in K$ such that $va=\alpha$ and $vb=n\alpha$. Then $v(a^n/b)=0$. Since we have $\overline{K}=\overline{L}$, there exists some $c\in K$ satisfying $vc=0$ and $\overline{c}=\overline{a^n/b}$, hence $\overline{a^n/bc}=1$. So a^n/bc is a 1-unit, and since n is not divisible by p, Corollary 9.33 (applied to (L,v)) shows that it admits an n-th root $d\in L$. We have that $(a/d)^n=bc\in K$. Since the polynomial X^n-bc has no multiple roots, it is separable and thus, a/d is separable algebraic over K. Since K was assumed to be relatively separable-algebraically closed in L, we find that $a/d\in K$. Since $vc=0$, we have the equality $nv(a/d)=vbc=vb=n\alpha$, which shows that $\alpha=v(a/d)\in vK$. □

We will show in Example **??** below that the condition that $vL=vK$ is necessary.

As an immediate consequence of Lemma 9.26, we obtain:

Corollary 9.27 *Take a henselian field (L,v) and a perfect relatively algebraically closed subfield K of L. Then $\overline{K}$ is relatively algebraically closed in $\overline{L}$. If in addition $\overline{K}=\overline{L}$, then vK is pure in vL.*

The result of Lemma 9.26 can be improved for extensions of finitely ramified fields.

Lemma 9.28 *Let (L,v) be a henselian finitely ramified field and K relatively algebraically closed in L. If $\overline{L}|\overline{K}$ is separable algebraic and if (L,v) and (K,v) have a common prime element π, then (K,v) is a henselian field, $\overline{L}=\overline{K}$ and vL/vK is torsion free.*

Proof: As in the proof of Lemma 9.26 it is shown that (K,v) is henselian and that $\overline{L}=\overline{K}$.

Let p be the residue characteristic of (L,v). It suffices to show the following: If $a\in L$ such that $pva\in vK$, then $va\in vK$. Without loss of generality, we can assume that $va\geq 0$; otherwise, we replace va by $-va$ and a by a^{-1}. Let $b=a^p\tilde{c}$, where $\tilde{c}\in K$ is chosen such that $v\tilde{c}=-va^p$, that is, $vb=0$. We claim that $mv\pi\in\Lambda^L(b,K)$ for every positive integer m. Assume the contrary, that is, that there exists a minimal $m_0\geq 0$ such that $(m_0+1)v\pi\notin\Lambda^L(b,K)$. In view of $m_0v\pi\in\Lambda^L(b,K)$, let $c_0\in K$ such that $v(b-c_0)\geq m_0v\pi$. Since $vL=\mathbb{Z}v\pi$, the inequality $v(b-c_0)>m_0v\pi$ would imply that $(m_0+1)v\pi\in\Lambda^L(b,K)$, a contradiction. Hence $v((b-c_0)\pi^{-m_0})=0$. Since $\overline{L}=\overline{K}$ by assumption, there is some $c\in K$ such that $v((b-c_0)\pi^{-m_0}-c)>0$. That is, $v(b-c_0-\pi^{m_0}c)>m_0v\pi$. As before, it follows that $(m_0+1)v\pi\in\Lambda^L(b,K)$, a contradiction. Hence, $mv\pi\in\Lambda^L(b,K)$ for every positive integer m.

By what we have shown, we can choose $c\in K$ such that $v(b-c)>2ev\pi=2vp$. Then $vc=vb=0$, and $1+\frac{b-c}{c}$ is a 1-unit of level $>2vp$. Hence, in the henselian field (L,v) there is some unit $\tilde{a}$ such that $\tilde{a}^p=1+\frac{b-c}{c}$. Consequently, $b=c\tilde{a}^p$, that is, $(a/\tilde{a})^p=c/\tilde{c}\in K$. Since K was assumed to be relatively algebraically closed in L, we find that $a/\tilde{a}\in K$. This shows that $va=v(a/\tilde{a})\in vK$. □

Corollary 9.29 *Let K be a relatively algebraically closed subfield of L and assume that (L,v) is a $\wp$-adically closed field. If both (K,v) and (L,v) have a common prime element, then also (K,v) is a $\wp$-adically closed field, and $\overline{K}=\overline{L}$.*

Proof: Since $\overline{L}$ is a finite field, also $\overline{K}$ is a finite field and the extension $\overline{L}|\overline{K}$ is separable algebraic. Hence, Lemma 9.28 shows that $\overline{L}=\overline{K}$ and that vL/vK is torsion free. From the latter, it follows that also vK is a $\mathbb{Z}$-group. Hence by the foregoing theorem, (K,v) is $\wp$-adically closed. □

Concerning the embedding of a residue field in a valued field, we can derive the following assertion:

Lemma 9.30 *Let (K,v) be a henselian field of equal characteristic. Let k be the prime field of K and $\overline{K}$. If $k'|k$ is a separably generated subextension of $\overline{K}|k$, then there is a field embedding of k' in K such that $\overline{\iota\zeta}=\zeta$ for all $\zeta\in\overline{K}$. If (K,v) is of residue characteristic 0, then $\overline{K}$ itself admits such an embedding.*

Proof: Let $\mathcal{T}$ be a transcendence basis of $k'|k$ such that $k'|k(\mathcal{T})$ is a separable algebraic extension. For every $t\in\mathcal{T}$, choose an element $a_t\in K$ such that $\overline{a_t}=t$. Then by Lemma 6.30, the elements a_t, $t\in\mathcal{T}$, are algebraically valuation independent and thus also algebraically independent in K (over k). Hence, the assignment $t\mapsto a_t$ induces a field embedding ι of $k(\mathcal{T})$ in K over k. In view of Lemma 6.35, it satisfies that $\overline{\iota\zeta}=\zeta$ for all $\zeta\in k(\mathcal{T})$. Now we take K_0 to be the relative algebraic closure of $\iota k(\mathcal{T})$ in K.

By Lemma 9.26, its residue field contains the relative separable-algebraic closure of $k(\mathcal{T})$ in $\overline{K}$. Thus, it also contains k' because $k'|k(\mathcal{T})$ is separable algebraic. But K_0 is an algebraic extension of the trivially valued field $(\iota k(\mathcal{T}), v)$ and is thus also trivially valued. Consequently, the residue map restricted to K_0 is an isomorphism. Its inverse, restricted to k', is the required embedding of k' in K.

The second assertion follows from what we have just proved and the fact that every extension of fields of characteristic 0 is separably generated. □

As a further corollary to Theorem 9.6 and Lemma 9.26, we have something like a converse to Theorem **??**. Recall that the completion of a valued field is an immediate extension. So if a valued field has rank 1, then the same is true for its completion.

Corollary 9.31 *The relative separable-algebraic closure of a valued field of rank 1 in its completion is henselian.*

9.11 n-th powers in henselian fields

Let $\mathbf{K} = (K, v)$ be a henselian field, $a \in K$ and n a natural number. We ask whether a is an n-th power in K, that is, whether there exists $b \in K$ such that $a = b^n$. A first observation is the following. If $a = b^n$ with $b \in K$, then $va = nvb \in nvK$, and if $a \in \mathcal{O}_{\mathbf{K}}^\times$, then $\overline{a} = \overline{b}^n \in (\overline{K}^\times)^n$. A partial converse of the latter implication is given by the next lemma.

Before we continue, let us note the following. If k is an arbitrary field and the natural number n is not divisible by the characteristic of k, then a polynomial $X^n - a$ with $0 \neq a \in k$ has no multiple roots. Indeed, a multiple root would also be a root of the derivative nX^{n-1} of $X^n - a$. But since $n \neq 0$ in k, the only root of nX^{n-1} is 0, which is not a root of $X^n - a$.

Lemma 9.32 *Let (K, v) be a henselian field, $a \in \mathcal{O}_{\mathbf{K}}^\times$ and n a natural number not divisible by the residue characteristic* char $\overline{K}$. *If $\overline{a} \in (\overline{K}^\times)^n$, then a is an n-th power in K.*

Proof: Consider the polynomial $X^n - a \in \mathcal{O}_{\mathbf{K}}[X]$. Its reduction is $X^n - \overline{a}$. By hypothesis, this polynomial has a root in $\overline{K}$. Since n is not divisible by the characteristic of $\overline{K}$ and $\overline{a} \neq 0$, this root is a simple root. Hence by Hensel's Lemma in the version 3) of Theorem 9.1, $X^n - a$ admits a root in K, and a is thus an n-th power in K. □

Since a 1-unit has residue 1 and the polynomial $X^n - 1$ always admits 1 as a root, we obtain the following corollary:

Corollary 9.33 *Let (K, v) be a henselian field and n a natural number not divisible by the residue characteristic* char $\overline{K}$. *Then every* 1*-unit in (K, v) is an n-th power.*

In the case of $a \in \mathcal{O}_{\mathbf{K}}^\times$, what can be said if n is divisible by the residue characteristic? In this case one can employ Newton's Lemma. The derivative of $X^n - a$ is nX^{n-1}, hence an approximative root b will satisfy the hypothesis of Newton's Lemma if $v(b^n - a) > 2vn + 2(n-1)vb$. But then, $vb^n = va = 0$ and thus, $vb = 0$. So in fact our condition is $v(b^n - a) > 2vn$. (Note that the expression "vn" denotes the value of the element $n \cdot 1$, where 1 is the multiplicative unit of K.) Now Newton's Lemma shows:

If a is an n-th power up to a summand of value $> 2vn$, then a is an n-th power.

Let us again consider the important special case of 1-units. If a is a 1-unit, then the value $v(a-1)$ is called the **level of the 1-unit** a. We have proved:

Lemma 9.34 *Let (K,v) be a henselian field and n a natural number. If a is a 1-unit of level $> 2vn$, then a is an n-th power in K.*

This lemma will be applied in the next section to improve the assertion of Lemma 9.26 for an important class of valued fields of mixed characteristic (cf. Lemma 9.28). Let us now discuss the special case of p-th roots of 1-units over henselian fields of mixed characteristic.

Throughout this section, we take C to be an element in the algebraic closure of $\mathbb{Q}$ such that $C^{p-1} = -p$, where $p > 0$ is a prime. Take a henselian field (K,v) of characteristic 0 and residue characteristic p. Then $C \in \tilde{\mathbb{Q}} \subseteq \tilde{K}$. Extend the valuation v to $\tilde{K}$. Note that

$$C^p = -pC \qquad \text{and} \qquad vC = \frac{1}{p-1} vp > 0 \,.$$

Consider the polynomial

$$X^p - (1+b) \tag{9.12}$$

with $b \in K$. Performing the transformation

$$X \;=\; CY \,+\, 1\,, \tag{9.13}$$

dividing by C^p and using that $C^p = -pC$, we obtain the polynomial

$$f(Y) \;=\; Y^p \,+\, g(Y) \,-\, Y \,-\, \frac{b}{C^p} \tag{9.14}$$

with

$$g(Y) \;=\; \sum_{i=2}^{p-1} \binom{p}{i} C^{i-p} Y^i \tag{9.15}$$

a polynomial with coefficients in $K(C)$ of value > 0.

Lemma 9.35 *Take (K,v) and C as above. Then K contains C if and only if it contains all p-th roots of unity.*

Proof: Since char $\overline{K} = p$, the restriction of v to $\mathbb{Q} \subset K$ is the p-adic valuation. Since (K,v) is henselian, it contains $(\mathbb{Q}^h, v_p)$ (Lemma **??**). Let $\eta \neq 1$ be a p-th root of unity. It suffices to show that $\mathbb{Q}^h(\eta) = \mathbb{Q}^h(C)$. Applying the transformation (9.13) to the polynomial (9.12) with $b = 0$, we obtain the polynomial $f(Y) = Y^p + g(Y) - Y \in \mathbb{Q}(C)[Y]$ which splits over $\mathbb{Q}^h$ by Hensel's Lemma because $\overline{f}(Y) = Y^p - Y$ splits over $\mathbb{F}_p$. Since the non-zero roots of f have nonzero residue and thus value zero, $v(\eta - 1) = vC = vp/(p-1)$. We find that $(v\mathbb{Q}^h(\eta) : v\mathbb{Q}^h) \geq p-1$. Consequently,

$$[\mathbb{Q}^h(\eta) : \mathbb{Q}^h] \;\geq\; p-1 \;\geq\; [\mathbb{Q}^h(C) : \mathbb{Q}^h] \;\geq\; [\mathbb{Q}^h(\eta) : \mathbb{Q}^h]\,,$$

showing that equality holds everywhere and that $\mathbb{Q}^h(\eta) = \mathbb{Q}^h(C)$. □

The following lemma supplements Lemma 9.33 in the case of p-th roots of 1-units:

Lemma 9.36 *Let (K,v) be a henselian field containing all p-th roots of unity. Then*

$$vb > \frac{p}{p-1}vp \;\Rightarrow\; 1+b \in (K^\times)^p$$

for all $b \in K$. In other words, every 1-unit of level $> \frac{p}{p-1}vp$ has a p-th root in K.

Proof: Consider the polynomial (9.14), derived from (9.12). If $vb > \frac{p}{p-1}vp = vC^p$, then $\overline{f}(Y) = Y^p - Y$, which splits over $\overline{K}$. By Hensel's Lemma, this implies that $f(Y)$ splits over K. Via the transformation (9.13), it follows that $1+b$ has a p-th root in K. □

Corollary 9.37 *Let (K,v) be a henselian field containing all p-th roots of unity. Take any 1-units $1+b$ and $1+c$ in K. Then:*

a) $1+b \in (1+b+c)\cdot(K^\times)^p$ *if* $vc > \frac{p}{p-1}vp$ *.*

b) $1+b \in (1+b+c)\cdot(K^\times)^p$ *if* $1+c \in (K^\times)^p$ *and* $vbc > \frac{p}{p-1}vp$ *.*

c) $1+c^p+pc \in (K^\times)^p$ *if* $vc^p > vp$ *.*

d) $1+b-pc \in (1+b+c^p)\cdot(K^\times)^p$ *if* $vb \geq \frac{1}{p-1}vp$ *and* $vc^p > vp$ *.*

Proof: a): $1+b \in (1+b+c)(K^\times)^p$ is true if the quotient

$$\frac{1+b+c}{1+b} \;=\; 1+\frac{c}{1+b}$$

is an element of $(K^\times)^p$. By hypothesis we have $vb > 0$ and thus $v\frac{c}{1+b} = vc$. Now our assertion follows from Lemma 9.36.

b): An application of part a) shows that

$$(1+b+c) \;\in\; (1+b)(1+c)\cdot(K^\times)^p \quad \text{if } \; vbc > \tfrac{p}{p-1}vp.$$

The assertion of b) is an immediate consequence of this.

c): If $vc^p > vp$ then for every $i = 2,\ldots,p-1$ we have

$$v\binom{p}{i}c^i \;\geq\; vp + 2vc \;>\; \frac{p+2}{p}vp \;\geq\; \frac{p}{p-1}vp\;;$$

note that the last inequality holds for every $p \geq 2$. This together with assertion a) yields

$$1+c^p+pc \;\in\; \left(1+c^p+pc+\sum_{i=2}^{p-1}\binom{p}{i}c^i\right)(K^\times)^p \;=\; (1+c)^p(K^\times)^p \;=\; (K^\times)^p\,.$$

d): In view of part c), the assertion follows from part b) where b is replaced by $b-pc$ and c is replaced by c^p+pc. Note that b) can be applied since $v(b-pc)(c^p+pc) > \frac{1}{p-1}vp+vp = \frac{p}{p-1}vp$. □

What can we say about elements of value $\neq 0$? We have seen that $va \in nvK$ if a is an n-th power in K. So we can answer: If there is some $c \in K$ such that $vac^n = 0$ and if ac^n satisfies one of the above criteria for being an n-th power, then a is an n-th power. In particular, we have:

Corollary 9.38 *Let (K,v) be a valued field and n a natural number. If K is closed under n-th roots, then vK is n-divisible, and $\overline{K}$ is closed under n-th roots. The converse holds if (K,v) is henselian and n is not divisible by the residue characteristic* $\mathrm{char}\,\overline{K}$.

(For the first assertion, cf. Lemma 6.40 and Lemma 6.41).

But this answer is not satisfactory if we want to decide for every single element whether it is an n-th power. There is an easy answer to this problem if (K,v) is equipped with a multiplicative coefficient map $\mathrm{co} : K \to \overline{K}$. In this case, we can prove the following generalization of Lemma 9.32:

Lemma 9.39 *Let (K,v) be a henselian field and n a natural number not divisible by the residue characteristic* $\mathrm{char}\,\overline{K}$. *Assume that (K,v) admits a multiplicative coefficient map. Then $a \in K$ is an n-th power in K if and only if $va \in nvK$ and* $\mathrm{co}\,a$ *is an n-th power in $\overline{K}$.*

Proof: Since co is assumed to be multiplicative, $\mathrm{co}\,a$ is an n-th power in $\overline{K}$ if $a \in K$ is an n-th power in K. The latter also implies that $va \in nvK$. Now assume that $va \in nvK$ and $\mathrm{co}\,a$ is an n-th power in $\overline{K}$. We choose $c \in K$ such that $va = nvc$. Then $\mathrm{co}\,ac^{-n} = (\mathrm{co}\,a)(\mathrm{co}\,c)^{-n}$ is again an n-th power in $\overline{K}$. Since $vac^{-n} = 0$, we have that $\overline{ac^{-n}} = \mathrm{co}\,ac^{-n}$. By Lemma 9.32 it follows that ac^{-n} is an n-th power in K and thus, the same holds for a. □

9.12 When polynomials are close to each other

The following is an application of Theorem 9.1:

Theorem 9.40 *Let (K,v) be a henselian field and $f = f_1 \cdot \ldots \cdot f_r$ where $f_1, \ldots, f_r$ are distinct monic separable irreducible polynomials over K. Then for every $\gamma \in vK$ there is some $\beta \in vK$ such that the following holds: If h is any monic polynomial over K satisfying $v(f-h) > \beta$, then $h = h_1 \cdot \ldots \cdot h_r$ where $h_1, \ldots, h_r$ are distinct monic separable irreducible polynomials over K and for each $k \in \{1, \ldots, r\}$, $\deg f_k = \deg h_k$, $v(f_k - h_k) > \gamma$, for all roots a of f_k and b of h_k, $K(a)$ and $K(b)$ are isomorphic over K, and f_k and h_k have the same splitting field.*

Proof: Let $n = \deg f$. By our hypothesis on f, it has n distinct roots $a_1, \ldots, a_n \in \tilde{K}$. Extend v to $\tilde{K}$ and choose some $\alpha \in vK$ such that

$$\alpha > \max\{v(a_i - a_j) \mid 1 \le i < j \le n\}$$

(recall that vK is cofinal in $\widetilde{vK} = v\tilde{K}$). By the Continuity of Roots (Theorem 5.11) there exists some $\beta \in vK$ such that for every monic polynomial h of degree n over K such that $v(f-h) > \beta$, there is an enumeration $b_1, \ldots, b_n$ of its roots in $\tilde{K}$ such that $v(a_i - b_i) > \alpha$ for $1 \le i \le n$. By our choice of α it follows that b_i is the unique root of h satisfying $v(a_i - b_i) > \alpha$. Consequently, h has no multiple roots.

For every $k \in \{1, \ldots, r\}$, we define $h_k = \prod (X - b_i)$ where the product is taken over all i such that a_i is a root of f_k. We have to show that h_k is an irreducible polynomial over K; since it has no multiple roots, it must be separable. Note that for every $i \in \{1, \ldots, n\}$

and every $\sigma \in \mathrm{Gal}\, K$, σb_i is a root of h. Assume that $a_i = \sigma a_j$. Since (K, v) is henselian, we have that $v(a_j - \sigma b_i) = v\sigma(a_i - b_i) > \alpha$. By our choice of α, it follows that $\sigma b_i = b_j$. Hence, b_i and b_j are conjugate over K if and only if a_i and a_j are, which in turn holds if and only if they are roots of the same f_k. By definition, this is the case if and only if b_i and b_j are roots of the same h_k. Consequently, h_k is a polynomial over K, and since all its roots are conjugate over K, it is irreducible.

Since h is separable over K, every b_i lies in K^{sep}. By our choice of α and by Krasner's Lemma (cf. Theorem 9.1), we obtain that $K(a_i) \subset K(b_i)$. But if k is such that a_i is a root of f_k, then $[K(a_i) : K] = \deg f_k = \deg h_k = [K(b_i) : K]$, showing that $K(a_i) = K(b_i)$. Therefore, f_k and h_k have the same splitting field. Moreover, if $a = a_i$ and b is an arbitrary root of h_k, then $K(b)$ is isomorphic to $K(b_i) = K(a)$ over K.

From Exercise 5.5 we infer that there is some $\delta \in vK$ such that if $v(a_i - b_i) > \delta$ for all i, then $v(f_k - h_k) > \gamma$. So if we choose $\alpha \geq \delta$, then β will have all required properties. □

Corollary 9.41 *Let (K, v) be an arbitrary field and f an irreducible polynomial over K, and let the notation be as in Lemma 7.46, where we take $K' = K^h$. Then there is some $\beta \in vK$ such that the following holds: If h is any monic polynomial over K satisfying $v(f - h) > \beta$, then $h = h_1 \cdot \ldots \cdot h_r$ where $h_1, \ldots, h_g$ are irreducible polynomials over K^h and for each $k \in \{1, \ldots, g\}$ and all roots a of f_k and b of h_k, $K(a)$ and $K(b)$ are isomorphic over K^h.*

Proof: We choose the separable polynomial $\tilde{f}$ with $f(X) = \tilde{f}(X^{p^\nu})$ as in the proof of part d) of Lemma 7.46. We observe that $v(\tilde{f} - \tilde{h}) > \beta$ if and only if $v(\tilde{f}(X^{p^\nu}) - \tilde{h}(X^{p^\nu})) > \beta$ for every polynomial $\tilde{h} \in K[X]$. Now we apply the foregoing theorem to $\tilde{f}$ and $\tilde{h}$ in the place of f and h. Setting $h(X) = \tilde{h}(X^{p^\nu})$ and $h_k(X) = \tilde{h}_k(X^{p^\nu})$, we argue further as in the proof of part d) of Lemma 7.46. □

Corollary 9.42 *Let (K, v) be an arbitrary valued field, $f \in K[X]$ an irreducible polynomial over K and $a \in \tilde{K}$ a root of f. Further, let $v_1, \ldots, v_g$ be all extensions of v from K to $K(a)$. Then there is some $\beta \in vK$ such that the following holds: If h is any monic polynomial over K satisfying $v(f - h) > \beta$, and if $b \in \tilde{K}$ is a root of h and $w_1, \ldots, w_{g'}$ are all extensions of v from K to $K(b)$, then $g = g'$ and after a suitable renumeration of the w_i, we have that*

$$\begin{aligned} d(K(a)|K, v_i) &= d(K(b)|K, w_i) \\ e(K(a)|K, v_i) &= e(K(b)|K, w_i) \\ f(K(a)|K, v_i) &= f(K(b)|K, w_i)\,. \end{aligned}$$

Choosing β large enough, we also get that h is separable if f is.

Proof: Fix an extension of the valuation v to $\tilde{K}$ and let (K^h, v) denote the corresponding henselization of (K, v). We choose β as in the foregoing lemma and assume that $v(f - h) > \beta$. Then $f = f_1 \cdot \ldots \cdot f_g$ and $h = h_1 \cdot \ldots \cdot h_g$ such that f_k and h_k are irreducible polynomials over K^h satisfying the further assertions of that corollary. From Lemma 7.47 we infer that $g = g'$. We choose automorphisms $\iota'_1, \ldots, \iota'_g$ according to that lemma for h in the place of f. Now let b be a root of h. Then after a suitable renumeration, we can assume that

$\iota_i' b$ is a root of h_i if and only if $\iota_i a$ is a root of f_i. Hence, $\iota_i K(a).K^h = K(\iota_i a).K^h$ and $\iota_i' K(b).K^h = K(\iota_i' b).K^h$ are isomorphic over K^h. Consequently, in view of Lemma 11.2, $d(K(a)|K, v_i) = d(\iota_i K(a).K^h|K^h, v) = d(\iota_i' K(b).K^h|K^h, v) = d(K(b)|K, w_i)$, and similarly for the ramification index and the inertia degree.

The last assertion follows directly from Theorem 9.40. □

With this corollary, we can also prove:

Theorem 9.43 *Let (K, v) be a valued field of characteristic $p > 0$. Then (K, v) is a separably defectless field if and only if $(K, v)^c$ is a defectless field.*

Proof: By virtue of Theorem 11.77, $(K, v)^c$ is separably defectless if and only if it is defectless. Thus it suffices to prove that $(K, v)^c$ is separably defectless if and only if (K, v) is separably defectless. Suppose first that $(K, v)^c$ is separably defectless. We wish to show that (K, v) is separably defectless. Let $(L|K, v)$ be a finite separable extension. Using Lemma 11.99 and Lemma **??**, we find that $d(L|K, v) = d(L^c|K^c, v) = 1$. Hence again by Lemma **??**, (K, v) is separably defectless.

Now suppose that (K, v) is separably defectless, and let $(L'|K^c, v)$ be a finite separable extension. Let $L' = K^c(a)$ and f the minimal polynomial of a over K^c. We choose β according to Corollary 9.42. Since (K, v) is dense in $(K, v)^c$, there exists a polynomial $h \in K[X]$ such that $v(f - h) > \beta$. Let b be a root of h. Then by Corollary 9.42, b is separable over K and $d(L'|K^c, v) = d(K^c(b)|K^c, v)$. By Lemma 11.99 and our assumption that (K, v) be separably defectless, $d(K^c(b)|K^c, v) = d(K(a)|K, v) = 1$. Hence, $d(L'|K^c, v) = 1$, showing that $(K, v)^c$ is separably defectless. □

9.13 Fields with two henselian valuations

Theorem 9.44 *Let v_1 and v_2 be two non-trivial independent henselian valuations on the field K. Then K is separable-algebraically closed.*

Proof: Let $f = X^n + c_1 X^{n-1} + \ldots + c_0 \in K[X]$ be any separable irreducible polynomial; we wish to show that it is linear. Let $n = \deg f$ and choose distinct elements $a_1, \ldots, a_n \in K$ (note that K is infinite since it admits non-trivial valuations). Set $h := \prod_{i=1}^n (X - a_i) = X^n + d_1 X^{n-1} + \ldots + d_0$. Choose any $\gamma_i \in v_i K$, $i = 1, 2$. Let $\beta_i \in v_i K$ be the values given by the foregoing theorem, depending on γ_i (for $i = 2$, we replace f by h). By the Strong Approximation Theorem, we can find elements $c_j' \in K$ such that $v_1(c_j - c_j') > \beta_1$ and $v_2(d_j - c_j') > \beta_2$ for $0 \leq j < n$. Then by the foregoing theorem, the polynomial $g := X^n + c_1' X^{n-1} + \ldots + c_0'$ is irreducible over K since f is. On the other hand, its irreducible factors are linear since the same holds for h. This proves that f is linear. □

Corollary 9.45 *Let v_1 and v_2 be incomparable henselian valuations on the field K. Let w denote their finest common coarsening (its valuation ring is the smallest ring containing $\mathcal{O}_{v_1}$ and $\mathcal{O}_{v_2}$). Then Kw is separable-algebraically closed.*

Proof: By Theorem **??**, the induced valuations v_1/w and v_2/w on Kw are henselian. Since v_1 and v_2 are incomparable by assumption, $w \neq v_1$ and $w \neq v_2$ so that both v_1/w and v_2/w are non-trivial on Kw. They are independent because otherwise, they would admit a non-trivial common coarsening w', so ww' would be a common coarsening of v_1 and v_2, finer than w. Now our assertion follows from the foregoing theorem. □

9.14 The core and the henselian part of a valuation

In [POP4], F. Pop has given the definition of the core of a valuation (it is a revised version of an earlier definition given in [POP2]). In this section, let v be a fixed valuation on the algebraic closure $\tilde{K}$ of the field K. By this, also the henselization (K^h, v) of (K, v) is fixed. Let

$$\mathcal{V} := \{v\} \cup \{w \text{ coarsening of } v \mid K^h w \text{ is separable-algebraically closed}\} .$$

Then we define the **core of** v **on** K to be the valuation $\inf \mathcal{V}$. We will denote it by v_c^K. Its valuation ring is the union of all valuation rings $\mathcal{O}_w$ for $w \in \mathcal{V}$, and its valuation ideal is the intersection of their valuation ideals $\mathcal{M}_w$.

Lemma 9.46 *Let (K, v) be henselian. If $v \neq v_c^K$, then Kv_c^K and Kv are separable-algebraically closed. In any case, $v_c^K \in \mathcal{V}$.*

Proof: Assume that $v \neq v_c^K$, then there is a coarsening w of v such that Kw is separable-algebraically closed. Hence by **??**, also the residue field Kv of $(Kw, v/w)$ is separable-algebraically closed.

By virtue of Theorem **??**, every coarsening w of v is henselian. Hence in view of **??**, Kw is separable-algebraically closed if and only if $\mathrm{Gal}\, K$ is equal to the absolute inertia group of K, i.e., $aw = (\sigma a)w$ for all $\sigma \in \mathrm{Gal}\, K$ and every $a \in \mathcal{O}_w$. Here, $\mathcal{O}_w$ can be replaced by $\mathcal{O}_v$ because Kw is the quotient field of the valuation ring $\mathcal{O}_v/\mathcal{M}_w$ of $(Kw, v/w)$. On the other hand, "$aw = (\sigma a)w$" is equivalent to "$a - \sigma a \in \mathcal{M}_w$". So we see that $a - \sigma a \in \bigcap_{w\in\mathcal{V}} \mathcal{M}_w = \mathcal{M}_{v_c^K}$ for every $a \in \mathcal{O}_v$ and every $\sigma \in \mathrm{Gal}\, K$. This proves that Kv_c^K is separable-algebraically closed. □

Suppose that v_1 and v_2 are henselian valuations with cores w_1 resp. w_2 on the field K. Since the latter are coarsenings of the former, they are henselian too. So if w_1 and w_2 were incomparable, then we could apply Corollary 9.45 to find that their finest common coarsening w has a separable-algebraically closed residue field. On the other hand, w would be coarser than w_1, which contradicts the definition of the core valuation. So we have proved:

Lemma 9.47 *Let v_1 and v_2 be two henselian valuations on a field K. Then their cores are comparable.*

Lemma 9.48 *Assume that $(L|K, v)$ is a normal extension and that (L, v) is henselian. Then v_c^K is henselian already on K.*

Proof: By **??**, all extensions of v from K to L are conjugate. Since $L|K$ is normal, $(L, v\sigma)$ and (L, v) are isomorphic for every $\sigma \in \mathrm{Gal}\, K$. Consequently, also $v\sigma$ is henselian on L. So by the foregoing lemma, v and $v\sigma$ have comparable cores on L. The reader may verify that $(v\sigma)_c^K = v_c^K\sigma$. Hence, comparability implies equality (cf. **??**). This shows that v_c^K is the only extension of its restriction to K. Since v_c^K is henselian on L (being a coarsening of v), this implies that v_c^K is henselian also on K. □

Let us introduce a second significant coarsening of a given valuation. Beforehand, we observe:

Lemma 9.49 *Let $(L|K,v)$ be an algebraic extension. Define*

$$\mathrm{Uniq}(L|K,v) := \{w \mid w \leq v \text{ and } w \text{ admits a unique extension to } L\}\ .$$

Then the valuation

$$v_{L|K} := \sup \mathrm{Uniq}(L|K,v)$$

also admits a unique extension from K to L. If $L|K$ is finite, then $v/v_{L|K}$ admits an extension from $Kv_{L|K}$ to $Lv_{L|K}$ which is independent from $v/v_{L|K}$ on $Lv_{L|K}$.

Proof: The valuation ring of $v_{L|K}$ is $\bigcap_{w\in\mathrm{Uniq}(L|K,v)} \mathcal{O}_{(K,w)}$, and its valuation ideal is $\bigcup_{w\in\mathrm{Uniq}(L|K,v)} \mathcal{M}_{(K,w)}$. For every $w \in \mathrm{Uniq}(L|K,v)$, there is a unique valuation ring $\mathcal{O}_{(L,w)}$ lying above $\mathcal{O}_{(K,w)}$. Hence if v' is an extension of $v_{L|K}$ to L, then $\mathcal{O}_{(L,w)} \supset \mathcal{O}_{v'}$ and $\mathcal{M}_{(L,w)} \subset \mathcal{M}_{v'}$. This shows that the valuation ring $\bigcap_{w\in\mathrm{Uniq}(L|K,v)} \mathcal{O}_{(L,w)}$ of L lies above the valuation ring $\mathcal{O}_{v'}$ of L, which proves that they are equal.

Now assume that $L|K$ is finite, and let $v_1 = v, v_2, \ldots, v_n$ denote all extensions of v to L. Let $w = \min\{v_1, \ldots, v_n\}$. Then w is the only extension of its restriction to K. Hence, $w \leq v_{L|K}$. On the other hand, **??** shows that there is $i \in \{1, \ldots, n\}$ such that $w = \min\{v, v_i\}$. That is, v/w and v_i/w are independent extensions of v/w from Kw to Lw. This implies that $w \geq v_{L|K}$, so $w = v_{L|K}$ and our assertions are proved. □

Now we define the **henselian part** of v on K to be the valuation $v_{K^h|K}$. By the foregoing lemma, the henselian part has a unique extension to K^h. Since this extension is a coarsening of the henselian valuation v on K^h, it is also henselian. This shows that the henselian part of v on K is itself henselian. On the other hand, it follows from the definition that it is the finest coarsening of v which is henselian on K.