New light on Hensel's lemma

David Brink

(To appear in *Expositiones Mathematicae*)

Abstract: The historical development of Hensel's lemma is briefly discussed (section 1). Using Newton polygons, a simple proof of a general Hensel's lemma for separable polynomials over Henselian fields is given (section 3). For polynomials over algebraically closed, valued fields, best possible results on continuity of roots (section 4) and continuity of factors (section 6) are demonstrated. Using this and a general Krasner's lemma (section 7), we give a short proof of a general Hensel's lemma and show that it is, in a certain sense, best possible (section 8). All valuations here are non-archimedean and of arbitrary rank. The article is practically self-contained.

MSC 2000: primary 12J25; secondary 12J20

Keywords: Valued fields, Hensel's lemma, Krasner's lemma, Newton polygons, continuity of roots

1 Introduction and historical remarks

The p-adic numbers were introduced in 1904 by Hensel in *Neue Grundlagen der Arithmetik*. In the same article, Hensel showed that if a monic polynomial f with integral p-adic coefficients has an *approximate* factorisation $f \approx gh$, meaning that the coefficients of the difference $f - gh$ are p-adically smaller than the discriminant of f, then there exists an *exact* factorisation $f = g^*h^*$. Four years later, in 1908, Hensel gave a somewhat more general result in his book *Theorie der algebraischen Zahlen*, where f is no longer assumed monic, and the discriminant of f is replaced by the squared resultant of g and h.

Since then, many variations and generalisations of Hensel's result have been found, some of which bear only little resemblance to the original. Confusingly, all these theorems are known today as "Hensel's lemma". We mention here the most important. Kürschak (1913) introduced real valuations on the abstract fields recently defined by Steinitz and indicated that Hensel's arguments would carry over

to complete, non-archimedean valued fields. Rychlík (1919) undertook these generalisations explicitly. Krull (1932) introduced general valuations, gave a new concept of completeness, and showed that a weak Hensel's lemma (g and h are assumed relatively prime modulo the valuation ideal) holds for such fields. Nagata (1953) showed that if a weak Hensel's lemma holds in some field with a valuation v, then the original Hensel's lemma holds too under the extra assumption that $v(f-gh)-2v(\mathrm{Res}(g,h))$ is not contained in the maximal convex subgroup of the value group not containing $v(\mathrm{Res}(g,h))$. Rim (1957) and Rayner (1958) proved that the unique extension property implies weak forms of Hensel's lemma. Ribenboim (1985) showed the logical equivalence between these and other "Hensel's lemmas". The reader is referred to the very interesting paper of Roquette (2002) regarding the history of Hensel's lemma and valuation theory in general.

In the present paper, a new proof of Hensel's lemma is presented that generalises the original in another direction, namely with respect to the accuracy of the approximate factorisation. It will be seen that the discriminant and the resultant disappear completely. They are replaced by two new polynomial invariants, here called the *separant* and the *bipartitionant*. The core of the proof is an analysis of the continuous behaviour of the roots of a polynomial as functions of the coefficients. These arguments, in contrast to earlier proofs, work equally well for arbitrary as for real valuations and make Nagata's extra assumption superfluous. The only thing we need is that the valuation has the *unique extension property*.

After proving his lemma in Hensel (1908), Hensel demonstrated the following: *If the p-adic polynomial F of degree ν has an approximate root ξ_0 satifying*

$$\rho > \max\left\{ \frac{i\rho' - \rho^{(i)}}{i-1} \Big|\ i = 2,3,\ldots,\nu \right\} \tag{1}$$

where ρ is the value of $F(\xi_0)$, and $\rho^{(i)}$ is the value of $F^{(i)}(\xi_0)/i!$, then Newton approximation gives an exact root ξ, provided that the values $\rho', \rho'', \ldots, \rho^{(\nu)}$ remain unchanged during the approximation process. In a short note from 1924, Rella showed the last condition to follow from (1). Our general Hensel's lemma will be seen to cover this *Hensel-Rella criterion*.

As noted by Rella in 1927, the existence of ξ is an almost immediate consequence of the *Newton polygon method*, a ubiquitous theme of this article. The p-adic Newton polygon was introduced by Dumas already in 1906 and later studied by Kürschák, Rella, and Ostrowski, but surprisingly never mentioned by Hensel.

2 Valuations, Newton polygons, and the unique extension property

Consider a field K. By a **valuation** on K we understand a map v from K into a totally ordered, additively written abelian group with infinity $\Gamma \cup \{\infty\}$ satisfying $v(0) = \infty$, $v(x) \in \Gamma$ if $x \neq 0$, $v(xy) = v(x) + v(y)$, and the **strong triangle inequality** $v(x+y) \geq \min\{v(x), v(y)\}$. In this situation, the pair (K, v) is called a **valued field**, $v(x)$ is called the **value** of $x \in K$, and x is called **integral** if $v(x) \geq 0$. If Γ is order-isomorphic to a subgroup of $\mathbb{R}^+$, the valuation is called **real** (the term "rank 1" is also standard). Sometimes we will use that Γ has division from $\mathbb{N}$. This may indeed be assumed without loss of generality, for we can always embed Γ into some larger group Γ' having that property. For a polynomial $f = a_0X^n + a_1X^{n-1} + \cdots + a_n$ with coefficients in K, we define $v(f) := \min\{v(a_0), \ldots, v(a_n)\}$.

The **Newton polygon** is a simple, yet powerful tool in valuation theory. It seems to have been always restricted to the case of real valuations, so we give here a definition for arbitrary valuations in the above sense. Consider a polynomial $f = a_0X^n + a_1X^{n-1} + \cdots + a_n$ of degree $n > 0$ with coefficients and roots in a valued field (K, v). Usually, it is difficult to compute the roots by means of the coefficients, but in contrast to this, it is easy to compute the values of the roots by means of the values of the coefficients. Define f's Newton polygon as the maximal convex map $\mathsf{NP} : \{0, 1, \ldots, n\} \to \Gamma \cup \{\infty\}$ satisfying $\mathsf{NP}(i) \leq v(a_i)$ for all i. By "convex" is understood the obvious, i.e. that $2 \cdot \mathsf{NP}(i) \leq \mathsf{NP}(i-1) + \mathsf{NP}(i+1)$ for all $i \neq 0, n$. The differences $\mathsf{NP}(i) - \mathsf{NP}(i-1)$, with the convention $\infty - \infty = \infty$, are the **slopes** of NP. They form an increasing sequence. Now write $f = a_0 \cdot \prod_{i=1}^{n}(X - \alpha_i)$ such that $v(\alpha_1) \leq \cdots \leq v(\alpha_n)$. Then $v(\alpha_i) = \mathsf{NP}(i) - \mathsf{NP}(i-1)$ for all $i = 1, \ldots, n$. In words, *the values of the roots of a polynomial equal the slopes of its Newton polygon.* The conceptually easy, but notationally cumbersome proof expresses the a_i as elementary symmetric functions in the α_i whereupon the $v(a_i)$ are computed from the $v(\alpha_i)$ using the strong triangle inequality.

We call a valued field (K, v) **Henselian** if it has the **unique extension property**, i.e. if v has a unique extension (also denoted v) to the algebraic closure $\tilde{K}$ of K. Note that the existence of a valuation extension is automatic with this definition. The unique extension property is, as a matter of fact, equivalent to many (maybe all) variants of Hensel's lemma, see for instance Ribenboim (1985). We actually only use a certain consequence of the unique extension property, namely this: *any K-automorphism σ of $\tilde{K}$ is isometric with respect to v* (since otherwise $v \circ \sigma$ would be an extension different from v). The slopes of the Newton polygon of an irreducible polynomial over a Henselian field are thus all the same, an observation due to Ostrowski (1935).

3 The separant and the "separable Hensel's lemma"

For a monic polynomial $f = \prod_{k=1}^{n}(X - \alpha_k)$ of degree $n > 1$ with roots in a valued field (K, v), we define the polynomial invariant

$$\mathscr{S} = \max\{v(f'(\alpha_k)) + v(\alpha_k - \alpha_l) \mid k \neq l\}$$

and call it f's **separant**. Note $f'(\alpha_k) = \prod_{l\neq k}(\alpha_k - \alpha_l)$ and that $\mathscr{S} < \infty$ iff f is separable (i.e. f has no multiple roots). A monic polynomial with integral coefficients has integral roots. So if f has integral coefficients, $\mathscr{S}$ is less than or equal to the value of f's discriminant $\operatorname{disc}(f) = \prod_{k<l}(\alpha_k - \alpha_l)^2$. Therefore, the following "separable Hensel's lemma" generalises the Hensel's lemma of 1904.

THEOREM 1 (separable Hensel's lemma). *Let f and f^* be monic polynomials of common degree $n > 1$ with integral coefficients in a Henselian field (K, v). Assume $v(f - f^*) > \mathscr{S}$ where $\mathscr{S}$ is the separant of f. Then f and f^* are both separable, and we may write $f = \prod_{k=1}^{n}(X - \alpha_k)$ and $f^* = \prod_{k=1}^{n}(X - \alpha_k^*)$ such that $K(\alpha_k) = K(\alpha_k^*)$ for all k.*

Proof. Since $\mathscr{S}$ is finite, f is separable. Write $f = \prod_{k=1}^{n}(X - \alpha_k)$ and fix a k. The Newton polygon NP of $f(X + \alpha_k)$ has $\mathsf{NP}(n) = \infty$ and $\mathsf{NP}(n-1) = v(f'(\alpha_k))$. The root α_k is integral, and therefore $v(f(X + \alpha_k) - f^*(X + \alpha_k)) = v(f - f^*)$. Consequently, the assumption $v(f - f^*) > \mathscr{S}$ implies that the Newton polygon NP^* of $f^*(X + \alpha_k)$ satisfies

$$\mathsf{NP}^*(i) = \begin{cases} \mathsf{NP}(i) & \text{for } i < n \\ v(f^*(\alpha_k)) > \mathscr{S} & \text{for } i = n \end{cases}$$

Hence, f^* has a root α_k^* with $v(\alpha_k^* - \alpha_k) = \mathsf{NP}^*(n) - \mathsf{NP}^*(n-1) > \mathscr{S} - v(f'(\alpha_k)) \geq v(\alpha_k - \alpha_l)$ for all l different from k.

This way we get n distinct roots $\alpha_1^*, \ldots, \alpha_n^*$ of f^* such that $v(\alpha_k^* - \alpha_k) > v(\alpha_k - \alpha_l)$ for all distinct k and l. Now Krasner's lemma (see section 7) gives $K(\alpha_k) = K(\alpha_k^*)$ for all k. Naturally, $f^* = \prod_{k=1}^{n}(X - \alpha_k^*)$. □

So if a polynomial f is separable, then any other polynomial f^* having coefficients sufficiently close to those of f has the same factorisation as f. This fails to be true if f has multiple roots. Over the field of dyadic numbers $\mathbb{Q}_2$, for instance, $f = X^2$ is reducible, but $f^* = X^2 + 2^\nu$ is irreducible for any ν.

EXAMPLE. Consider the polynomial $f = X(X-2)(X-4) = X^3 - 6X^2 + 8X$ over $\mathbb{Q}_2$. It has separant $\mathscr{S} = 5$ (whereas the value of the discriminant is 8). Hence, the polynomial $f^* = f + 2^\nu$ has 3 distinct roots in $\mathbb{Q}_2$ for all $\nu > 5$. For $\nu = 5$,

however, f^* has an irreducible quadratic factor over $\mathbb{Q}_2$, showing that the bound $v(f - f^*) > \mathscr{S}$ is best possible.

4 Error functions and continuity of roots

Consider two monic polynomials f and f^* of common degree $n > 1$ with coefficients in an algebraically closed, valued field (K, v). Since the coefficients of a polynomial can be expressed as elementary symmetric functions of the roots, the coefficients depend continuously on the roots. More precisely, if we write $f = \prod_{k=1}^{n}(X - \alpha_k)$ and $f^* = \prod_{k=1}^{n}(X - \alpha_k^*)$ in any way, then $v(f - f^*) \geq \min\{v(\alpha_1 - \alpha_1^*), \ldots, v(\alpha_n - \alpha_n^*)\}$.

The opposite, that the roots depend continuously on the coefficients, is less evident – it is not even clear what is to be understood by a such statement. The known results in this direction, for example Proposition 7 (page 191) of Ribenboim (1968), are of a qualitative nature and do not work well for polynomials with multiple roots.

Define the **error function** of the root α of f as the map $\Phi : \Gamma \cup \{\infty\} \to \Gamma \cup \{\infty\}$ given by

$$\Phi(x) = \sum_{l=1}^{n} \min\{x, v(\alpha - \alpha_l)\} \ .$$

It is a strictly increasing, piecewise linear (i.e. piecewise of the form $x \mapsto \nu x + \gamma$), bijective (since Γ is assumed to have division from $\mathbb{N}$) map with decreasing slopes ν from the set $\{1, \ldots n\}$. If Ψ is the error function of the root β of f, the strong triangle inequality gives

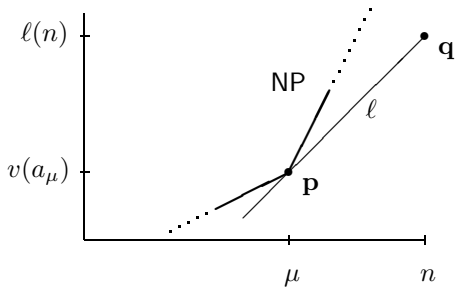
$$\Phi(x) = \Psi(x) \text{ for all } x \leq v(\alpha - \beta) \ . \tag{2}$$

Using error functions, we can now bound the error on the roots of a polynomial caused by an error on the coefficients.

THEOREM 2 (continuity of roots). *Let f and f^* be monic polynomials of common degree $n > 1$ with integral coefficients in an algebraically closed, valued field (K, v). We may then write $f = \prod_{k=1}^{n}(X - \alpha_k)$ and $f^* = \prod_{k=1}^{n}(X - \alpha_k^*)$ such that $v(\alpha_k - \alpha_k^*) \geq \Phi_k^{-1}(v(f - f^*))$ for each k. Here Φ_k denotes the error function of the root α_k of f.*

Proof. Write $f = \prod_{k=1}^{n}(X - \alpha_k)$ and put $\rho_k := \Phi_k^{-1}(v(f - f^*))$ for each k. We may assume $0 < v(f - f^*) < \infty$, and hence $0 < \rho_k < \infty$ for each k, since otherwise the claim is trivial. We show, for each k, that f and f^* have the same number of roots (counted with multiplicity) in the ball $\{x \in K \mid v(x - \alpha_k) \geq \rho_k\}$. It will then follow (for instance by assuming $\rho_1 \geq \rho_2 \geq \ldots$ and then choosing $\alpha_1^*, \alpha_2^*, \ldots$, in that order, such that, for each k, α_k^* is a root of $f^* / \prod_{l=1}^{k-1}(X - \alpha_l^*)$ and has $v(\alpha_k^* - \alpha_k) \geq \rho_k$) that we can write $f^* = \prod_{k=1}^{n}(X - \alpha_k^*)$ such that $v(\alpha_k^* - \alpha_k) \geq \rho_k$ for each k.

So fix a k. Let μ be the number of indices l with $v(\alpha_l - \alpha_k) < \rho_k$. We must show that the number of indices l with $v(\alpha_l^* - \alpha_k) < \rho_k$ is also μ. Consider the Newton polygon NP of $f(X + \alpha_k) = X^n + a_1 X^{n-1} + \cdots + a_n$. The slopes of NP are $v(\alpha_1 - \alpha_k), \ldots, v(\alpha_n - \alpha_k)$, in increasing order. So $\mathsf{NP}(i) - \mathsf{NP}(i-1) < \rho_k$ for $i \leq \mu$, and $\mathsf{NP}(i) - \mathsf{NP}(i-1) \geq \rho_k$ for $i > \mu$. Let ℓ be the "line through the point $\mathbf{p} = (\mu, v(a_\mu))$ with slope ρ_k", i.e. the map $\{0, \ldots, n\} \to \Gamma$ given by $\ell(i) = (i-\mu)\rho_k + v(a_\mu)$. Then $\mathsf{NP}(i) > \ell(i)$ for $i < \mu$, $\mathsf{NP}(\mu) = \ell(\mu)$, and $\mathsf{NP}(i) \geq \ell(i)$ for $i > \mu$ (see figure).



If we can show the same for the Newton polygon NP^* of $f^*(X + \alpha_k)$, we are done. Consider to this end the point $\mathbf{q} = (n, \ell(n))$ on ℓ and compute $\ell(n)$:

$$\begin{aligned} \ell(n) &= (n-\mu)\rho_k + v(a_\mu) = \sum_{l \in \{l \mid v(\alpha_l) \geq \rho_k\}} \rho_k + \sum_{l \in \{l \mid v(\alpha_l) < \rho_k\}} v(\alpha_l) \\ &= \sum_{l=1}^{n} \min\{\rho_k, v(\alpha_l)\} = \Phi_k(\rho_k) = v(f - f^*) \end{aligned}$$

Since α_k is integral, $v(f(X + \alpha_k) - f^*(X + \alpha_k)) = v(f - f^*)$. It follows that $\mathsf{NP}^*(i) = \mathsf{NP}(i)$ for $i \leq \mu$, and $\mathsf{NP}^*(i) \geq \ell(i)$ for $i > \mu$. This finishes the proof. $\square$

Heuristically, Theorem 2 says, *if a root α of f is far away from the other roots, then an error on the coefficients of f causes an error on α of equal or smaller magnitude; however the proximity of other roots makes α more sensitive to errors on the coefficients.* Let us note a consequence of Theorem 2 illustrating this. Fix a k, and let μ be the root multiplicity of α_k in f modulo the valuation ideal. This means that $v(\alpha_k - \alpha_l)$ is 0 for all but μ values of l. Hence $\Phi_k(v(\alpha_k - \alpha_k^*)) = \sum_{l=1}^{n} \min\{v(\alpha_k - \alpha_k^*), v(\alpha_k - \alpha_l)\} \leq \mu \cdot v(\alpha_k - \alpha_k^*)$ and thus

$$v(\alpha_k - \alpha_k^*) \geq v(f - f^*)/\mu \; . \tag{3}$$

In particular, $v(\alpha_k - \alpha_k^*) \geq v(f - f^*)/n$ holds for all k. In light of (3), we might say that the root α_k, as a function of f's coefficients, satisfies a Lipschitz condition of order $1/\mu$.

We conclude the section with a typical example where the bound given by Theorem 2 is best possible.

EXAMPLE. Consider again the polynomial $f = X(X-2)(X-4)$ over the field of dyadic numbers $\mathbb{Q}_2$. The roots $\alpha_1 = 0$ and $\alpha_3 = 4$ have the same error function $\Phi_1 = \Phi_3 : \gamma \mapsto \gamma + \min\{\gamma, 2\} + \min\{\gamma, 1\}$. The root $\alpha_2 = 2$ has error function $\Phi_2 : \gamma \mapsto \gamma + 2 \cdot \min\{\gamma, 1\}$. They are shown in Figure 1 and 2.

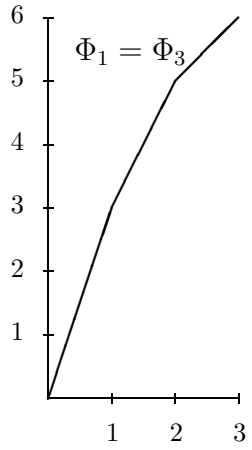


Figure 1

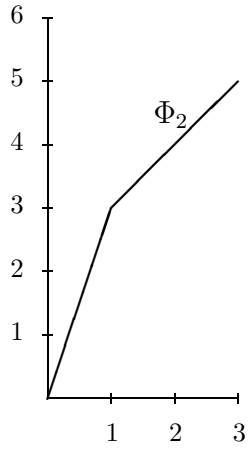


Figure 2

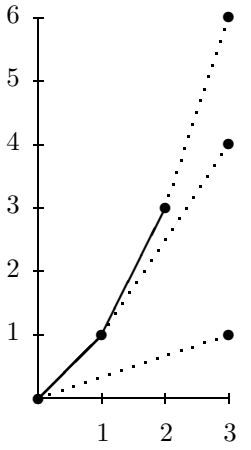


Figure 3

Now put $f^* = f + 2^\nu$ with some $\nu \geq 0$. By Theorem 2, we may write $f^* = (X - \alpha_1^*)(X - \alpha_2^*)(X - \alpha_3^*)$ such that $v(\alpha_k - \alpha_k^*) \geq \Phi_k^{-1}(\nu)$ for $k = 1, 2, 3$.

If α^* is a root of f^* maximally close to $\alpha_1 = 0$, then $v(\alpha^*)$ is the maximal slope of the Newton polygon NP^* of f^*. Figure 3 shows the Newton polygon NP of f (solid line) and NP^* for some values of ν (dotted lines). It is seen that $v(\alpha^*)$ equals $\Phi_1^{-1}(\nu)$, and hence $v(\alpha_1^*) = \Phi_1^{-1}(\nu)$. Similarly, one sees $v(\alpha_k - \alpha_k^*) = \Phi_k^{-1}(\nu)$ for $k = 2, 3$. So Theorem 2 gives in fact an optimal bound.

Finally note that, for $\nu > 5$, each root α_k^* of f^* is closer to α_k than to either of the two other roots of f. This agrees with Theorem 1 and the fact that f has separant $\mathscr{S} = 5$. □

5 The bipartitionant and the induced factorisation

Consider a monic polynomial f of degree $n > 1$ with coefficients in an algebraically closed, valued field (K, v) and write $f = \prod_{k=1}^{n}(X - \alpha_k)$. Let I and J be disjoint, nonempty sets with union $\{1, 2, \ldots, n\}$ and put $g = \prod_{i \in I}(X - \alpha_i)$ and $h = \prod_{j \in J}(X - \alpha_j)$ so that $f = gh$. Define the **bipartitionant** of the polynomials g and h as

$$\mathscr{B} := \max\{\Phi_i(v(\alpha_i - \alpha_j)) \mid i \in I \ , \ j \in J\}$$

where Φ_i is the error function of the root α_i of f. Clearly, $\mathscr{B} < \infty$ iff g and h are relatively prime. Equation (2) implies

$$\mathscr{B} = \max\{\Phi_j(v(\alpha_i - \alpha_j)) \mid i \in I \ , \ j \in J\} \ ,$$

showing that the definition is symmetric in g and h. The crucial property of the bipartitionant is this:

LEMMA 3. *Suppose the coefficients of f are integral. Let f^* be another monic polynomial of degree n with integral coefficients in K, and assume $v(f - f^*) > \mathscr{B}$. Then we may write $f^* = \prod_{k=1}^n (X - \alpha_k^*)$ such that $v(\alpha_i - \alpha_i^*)$, $v(\alpha_j - \alpha_j^*) > v(\alpha_i - \alpha_j)$ and thereby $v(\alpha_i - \alpha_j) = v(\alpha_i^* - \alpha_j^*)$ for all $i \in I$ and all $j \in J$.*

Proof. Write $f^* = \prod_{k=1}^n (X - \alpha_k^*)$ as in Theorem 2. Then $v(\alpha_i - \alpha_i^*) \geq \Phi_i^{-1}(v(f - f^*)) > \Phi_i^{-1}(\mathscr{B}) \geq v(\alpha_i - \alpha_j)$ and $v(\alpha_j - \alpha_j^*) \geq \Phi_j^{-1}(v(f - f^*)) > \Phi_j^{-1}(\mathscr{B}) \geq v(\alpha_i - \alpha_j)$ for all $i \in I$ and $j \in J$. The strong triangle inequality gives $v(\alpha_i - \alpha_j) = v(\alpha_i^* - \alpha_j^*)$. □

So in the situation of Lemma 3, the roots of f^* may be "bipartitioned" into two sets $\{\alpha_i^* \mid i \in I\}$ and $\{\alpha_j^* \mid j \in J\}$. This bipartitioning only depends on the factorisation $f = gh$ and not on the representation $f^* = \prod_{k=1}^n (X - \alpha_k^*)$ from Theorem 2. We say that the factorisation $f^* = g^* h^*$ where $g^* := \prod_{i \in I} (X - \alpha_i^*)$ and $h^* := \prod_{j \in J} (X - \alpha_j^*)$ is **induced** by the factorisation $f = gh$.

How does one compute $\mathscr{B}$? If $i_0 \in I$ and $j_0 \in J$ are such that $\mathscr{B} = \Phi_{i_0}(v(\alpha_{i_0} - \alpha_{j_0}))$, then

$$v(\alpha_{i_0} - \alpha_{j_0}) = \max\{v(\alpha_i - \alpha_{j_0}) \mid i \in I\} = \max\{v(\alpha_{i_0} - \alpha_j) \mid j \in J\} \qquad (4)$$

since the Φ's are strictly increasing. If, in turn, $i_0 \in I$ and $j_0 \in J$ satisfy (4), then

$$\begin{aligned}
\Phi_{i_0}(v(\alpha_{i_0} - \alpha_{j_0})) &= \sum_{k=1}^{n} \min\{v(\alpha_{i_0} - \alpha_{j_0}), v(\alpha_{i_0} - \alpha_k)\} \\
&= \sum_{i \in I} \min\{v(\alpha_{i_0} - \alpha_{j_0}), v(\alpha_{i_0} - \alpha_i)\} + \\
&\quad \sum_{j \in J} \min\{v(\alpha_{i_0} - \alpha_{j_0}), v(\alpha_{i_0} - \alpha_j)\} \\
&= \sum_{i \in I} v(\alpha_i - \alpha_{j_0}) + \sum_{j \in J} v(\alpha_{i_0} - \alpha_j) \\
&= v(g(\alpha_{j_0})) + v(h(\alpha_{i_0}))
\end{aligned}$$

where the third equality requires (4), the strong triangle inequality, and some consideration. Now conclude

$$\mathscr{B} = \max\{v(g(\alpha_{j_0})) + v(h(\alpha_{i_0}))\} \mid i_0 \in I \text{ and } j_0 \in J \text{ satisfy (4)}\} \ . \qquad (5)$$

Since the bipartitionant replaces twice the value of the resultant $\mathrm{Res}(g,h) = \prod_{i,j}(\alpha_i - \alpha_j)$ in our Hensel's lemma (Theorem 8), it is of interest to compare these two invariants, and from (5) follows immediately

$$\mathscr{B} \leq \sum_{j\in J} v(g(\alpha_j)) + \sum_{i\in I} v(h(\alpha_i)) = 2v(\mathrm{Res}(g,h))$$

when f has integral coefficients.

6 Continuity of factors

There is a remarkable analogue to the continuity of roots that could be called *continuity of factors*. In words it says, *if there is a factorisation $f = gh$ such that the roots of g are far away from the roots of h (but possibly close to each other), then an error on the coefficients of f causes an error on the coefficients of g which is in general smaller than the error caused on the roots of g individually.* It should be noted that the main part of Hensel's lemma is proved in the next section without the results of this section.

Consider two monic polynomials f, f^* of common degree $n > 1$ with integral coefficients in an algebraically closed, valued field (K, v), and write $f = \prod_{k=1}^{n}(X - \alpha_k)$. Let I and J be disjoint, non-empty sets with union $\{1, 2, \ldots, n\}$ and put $g = \prod_{i\in I}(X - \alpha_i)$ and $h = \prod_{j\in J}(X - \alpha_j)$. Let us call g an **isolated factor** of f if

$$\forall i, i' \in I \forall j \in J : v(\alpha_i - \alpha_{i'}) > v(\alpha_i - \alpha_j)$$

i.e. if there is a ball in K containing all roots of g and no roots of h.

Lemma 4 (continuity of isolated factors). *Assume $v(f - f^*) > \mathscr{B}$ where $\mathscr{B}$ is the bipartitionant of g and h, and consider the induced factorisation $f^* = g^*h^*$. If g is an isolated factor of f, then $v(g - g^*) \geq v(f - f^*) - \mathscr{B} + \max\{v(\alpha_i - \alpha_j) \mid i \in I, j \in J\}$.*

Proof. The idea is to use a general form of Newton approximation to come from g to g^*. We may assume $g(0) = 0$ by a change of variable. Put $\nu = \deg(g)$ and $\mu = \deg(h)$. We may then further assume $g = \prod_{i=1}^{\nu}(X - \alpha_i)$, $h = \prod_{j=\nu+1}^{n}(X - \alpha_j)$, and

$$\infty = v(\alpha_1) \geq v(\alpha_2) \geq \cdots \geq v(\alpha_\nu) > v(\alpha_{\nu+1}) \geq \cdots \geq v(\alpha_n)$$

since g is isolated. Thus, $u := \max\{v(\alpha_i - \alpha_j) \mid i \in I, j \in J\}$ equals $v(\alpha_{\nu+1})$, and $\mathscr{B}$ equals $\nu \cdot u + v(h(0))$ by (5).

Define three polynomial sequences $(g_m)_{m\in\mathbb{N}}$, $(h_m)_{m\in\mathbb{N}}$, and $(r_m)_{m\in\mathbb{N}}$ recursively like this: *Put $g_1 := g$. Given g_m, define h_m and r_m such that $f^* = g_m h_m + r_m$ and $\deg(r_m) < \nu$. Given g_m, h_m, and r_m, define $g_{m+1} := g_m + r_m/h_m(0)$.*

The difficulty of the proof lies in finding the right thing to prove. For a fixed m and for $i = 1, \ldots, \nu$, let a_i and c_i be the values of the coefficients to the terms of degree $\nu - i$ in g_m and r_m, respectively. We claim:

(A) $a_i \geq iu + \Delta$ where $\Delta := \min\{v(\alpha_\nu) - u, v(f - f^*) - \mathscr{B}\}$.
(B) The Newton polygon of h_m equals the Newton polygon NP of h.
(C) $c_i \geq v(h(0)) + iu + k_i\Delta$ where $k_i := \max\{k \in \mathbb{N} \mid k < (m + i + \nu - 1)/\nu\}$.

The claims are shown by induction after m. Assume $m = 1$ for the induction start. All roots of $g_1 = g$ have value at least $v(\alpha_\nu)$, and hence

$$a_i \geq i \cdot v(\alpha_\nu) \geq i(u + \Delta) \geq iu + \Delta \ .$$

This shows (A). Write $f^* - f = gh' + r'$ with $\deg(r') < \nu$. Then $f^* = g(h + h') + r'$ and thus $h_1 = h + h'$ and $r_1 = r'$. Also, $v(h'), v(r') \geq v(f - f^*)$. Adding h' to h does not change the Newton polygon since $v(h') > \mathscr{B} \geq v(h(0)) = \mathsf{NP}(\mu)$. This shows (B). Finally,

$$c_i \geq v(r') \geq v(f - f^*) \geq \mathscr{B} + \Delta \geq v(h(0)) + iu + k_i\Delta$$

since $k_i = 1$ for $m = 1$, showing (C).

For the induction step, assume (A), (B), and (C) hold for some m, and let (A'), (B'), and (C') be the statements corresponding to $m + 1$. (A') follows immediately from (A) and (C). Note $f^* = g_{m+1}h_m - (h_m/h_m(0) - 1)r_m$ and hence $h_{m+1} = h_m + h'$ and $r_{m+1} = r'$ if we write

$$-(h_m/h_m(0) - 1)r_m = g_{m+1}h' + r' \tag{6}$$

with $\deg(r') < \nu$. Let d_i be the value of the coefficient to the term of degree $n - i$ in the left hand side of (6). Using (A), (B), and (C) gives

$$\begin{aligned}
d_1 &\geq \mathsf{NP}(0) + u + k_1\Delta \\
d_2 &\geq \mathsf{NP}(1) + u + k_1\Delta \\
&\vdots \\
d_\mu &\geq \mathsf{NP}(\mu - 1) + u + k_1\Delta \\
d_{\mu+1} &\geq \mathsf{NP}(\mu - 1) + 2u + k_2\Delta \\
&\vdots \\
d_{n-1} &\geq \mathsf{NP}(\mu - 1) + \nu u + k_\nu\Delta \\
\infty = d_n &\geq \mathsf{NP}(\mu - 1) + (\nu + 1)u + k_{\nu+1}\Delta
\end{aligned}$$

The algorithm of polynomial division resulting in the expression (6) consists of a number of steps in each of which a monomial times g_{m+1} is subtracted from

$-(h_m/h_m(0)-1)r_m$. The key observation is that, in each step, the values of the coefficients of the remainder satify the same inequalities as the d_i. Let b'_i be the value of the coefficient to the term of degree $\mu - i$ in h'. Then

$$\begin{aligned} b'_1 &\geq \mathsf{NP}(0) + u + k_1\Delta > \mathsf{NP}(0) + u \geq \mathsf{NP}(1) \\ &\vdots \\ b'_\mu &\geq \mathsf{NP}(\mu-1) + u + k_1\Delta > \mathsf{NP}(\mu-1) + u \geq \mathsf{NP}(\mu) \end{aligned}$$

Hence $h_{m+1} = h_m + h'$ has NP as its Newton polygon, showing (B'). Let c'_i be the value of the coefficient to the term of degree $\nu - i$ in r'. Then

$$\begin{aligned} c'_1 &\geq \mathsf{NP}(\mu-1) + 2u + k_2\Delta = v(h(0)) + u + k_2\Delta \\ &\vdots \\ c'_\nu &\geq \mathsf{NP}(\mu-1) + (\nu+1)u + k_{\nu+1}\Delta = v(h(0)) + \nu u + k_{\nu+1}\Delta \end{aligned}$$

This shows (C') and finishes the induction step.

By (C), $v(r_m) \to \infty$ and hence $g_m h_m \to f^*$. By the continuity of roots, the roots of $g_m h_m$ converge to the roots of f^* (in a multiplicity-respecting way). By assumption, the roots of g have values $> u$, whereas the roots of h have values $\leq u$. Lemma 3 then gives that the roots of g^* have values $> u$, whereas the roots of h^* have values $\leq u$. By (A), the roots of g_m have values $> u$. It follows that the roots of g_m converge to the roots of g^*, and thereby the coefficients converge too: $g_m \to g^*$. Finally, $g^* = g + \sum_{m=1}^{\infty} r_m/h_m(0)$ and therefore by (C),

$$\begin{aligned} v(g - g^*) &\geq \min\{v(r_m) - v(h(0)) \mid m \in \mathbb{N}\} \\ &\geq u + \Delta \\ &\geq v(f - f^*) - \mathscr{B} + \max\{v(\alpha_i - \alpha_j) \mid i \in I, j \in J\} \ . \quad \square \end{aligned}$$

Let us show that Lemma 4 coincides with the Hensel-Rella criterion when g is linear. Given is a polynomial F with an approximate root ξ_0. Put $g = X - \xi_0$ and $h = (F - F(\xi_0))/(X - \xi_0)$. Then the left hand side of (1) is the value of $F(\xi_0) = F - gh$, and it can be seen that the right hand side of (1) equals the bipartitionant of g and h. Hence, the g_m converge to a polynomial $g^* = X - \xi$ dividing F. In the proof of Lemma 4, we could as well have defined g_{m+1} as $g_m + r_m/h_m(\xi_m)$ where ξ_m is any root of g_m (or any other element sufficiently close to 0). With this definition and with linear g, the approximation process becomes identical with usual Newton approximation.

THEOREM 5 (continuity of factors). *Let f and f^* be monic polynomials of common degree $n > 1$ with integral coefficients in an algebraically closed, valued field (K, v). Consider a monic factorisation $f = gh$, and let $\mathscr{B}$ be the bipartitionant of g and*

h. Assume $v(f-f^*) > \mathscr{B}$*, and let* $f^* = g^*h^*$ *be the induced factorisation. Then* $v(g-g^*), v(h-h^*) \geq v(f-f^*) - \mathscr{B}$*.*

Proof. Write $g = g_1 \ldots g_r$ such that each g_l is a maximal (with respect to divisibility) monic factor of g which is an isolated factor of f. The bipartitionant of g_l and $\tilde{g}_l := f/g_l$ is

$$\begin{aligned} \mathscr{B}_l &:= \max\{\Phi_i(v(\alpha_i - \alpha_j)) \mid g_l(\alpha_i) = \tilde{g}_l(\alpha_j) = 0\} \\ &= \max\{\Phi_i(v(\alpha_i - \alpha_j)) \mid g_l(\alpha_i) = 0, j \in J\} \end{aligned}$$

(last equality follows from the maximality of g_l), implying

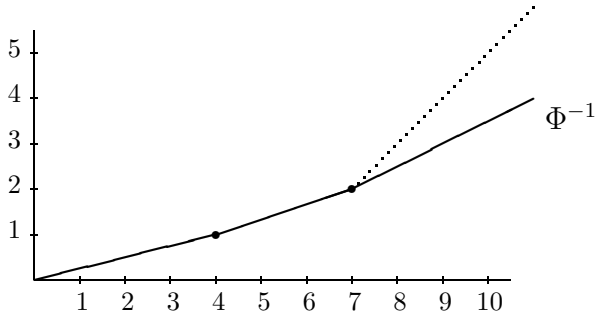
$$\begin{aligned} \mathscr{B} &= \max\{\Phi_i(v(\alpha_i - \alpha_j)) \mid i \in I, j \in J\} \\ &= \max\{\mathscr{B}_1, \ldots, \mathscr{B}_r\} \ . \end{aligned}$$

Lemma 4 gives

$$\begin{aligned} v(g-g^*) &\geq \min\{v(f-f^*) - \mathscr{B}_l + \Phi_i^{-1}(\mathscr{B}_l) \mid l = 1, \ldots, r \ , \ g_l(\alpha_i) = 0\} \\ &\geq \min\{v(f-f^*) - \mathscr{B}_l \mid l = 1, \ldots, l\} \\ &= v(f-f^*) - \mathscr{B} \ . \end{aligned}$$

The inequality for $v(h-h^*)$ can be proved the same way, but also follows directly by dividing f^* by g^*. □

Example. Consider the polynomial $f = X^2(X-2)(X-4) = X^4 - 6X^3 + 8X^2$ over the field of dyadic numbers $\mathbb{Q}_2$. The error function of the double root $\alpha_1 = \alpha_2 = 0$ is $\Phi(\gamma) = 2 \cdot \gamma + \min\{\gamma, 1\} + \min\{\gamma, 2\}$. The bipartitionant of the factors $g = X^2$ and $h = (X-2)(X-4)$ is $\mathscr{B} = v(g(4)) + v(h(0)) = 7$. Let $f^* = f + 2^\nu$ with $\nu > 7$ and consider the induced factorisation $f^* = g^*h^*$. By Lemma 4, $v(g-g^*) \geq \nu - 5$. The figure shows the inverse error function Φ^{-1} and the line $\nu \mapsto \nu - 5$ (dotted):



Let us compute $v(g-g^*)$ precisely. The Newton polygon of f^* shows that the roots $\alpha_1^*, \ldots, \alpha_4^*$ of f^* have values $v(\alpha_1^*) = v(\alpha_2^*) = (\nu-3)/2$, $v(\alpha_3^*) = 1$, and $v(\alpha_4^*) = 2$. We have

$$g^* = (X - \alpha_1^*)(X - \alpha_2^*) = X^2 - (\alpha_1^* + \alpha_2^*)X + \alpha_1^*\alpha_2^* \ .$$

From the above follows $v(\alpha_1^*\alpha_2^*) = \nu - 3$. It is more tricky to compute $v(\alpha_1^* + \alpha_2^*)$. To this end, consider the polynomial $f^*(X - \alpha_1^*)$. It has roots $2\alpha_1^*$, $\alpha_1^* + \alpha_2^*$, $\alpha_1^* + \alpha_3^*$, $\alpha_1^* + \alpha_4^*$ and constant term $f^*(-\alpha_1^*) = f(\alpha_1^*) + 12(\alpha_1^*)^3 = 12(\alpha_1^*)^3$. Thus,

$$\begin{aligned} v(\alpha_1^* + \alpha_2^*) &= v(f^*(-\alpha_1^*)) - v(2\alpha_1^*) - v(\alpha_1^* + \alpha_3^*) - v(\alpha_1^* + \alpha_4^*) \\ &= (3\nu - 5)/2 - (\nu - 1)/2 - 1 - 2 \\ &= \nu - 5 \ . \end{aligned}$$

Conclude $v(g - g^*) = \min\{\nu - 5, \nu - 3\} = \nu - 5$.

The moral of the story is that the bound on the coefficients of g^* given by Lemma 4 is best possible (contrary to that of Theorem 5) and better than the bound on the roots of g^* given by Theorem 2. □

One may wonder if there is also "continuity of factors" when $v(f - f^*) \leq \mathscr{B}$, i.e. if there is a bound on the error on the coefficients of g better than the bound on the error on the roots of g. That is not likely to be the case. For when $v(f - f^*) \leq \mathscr{B}$, it is no longer possible to bipartition the roots of f^* as in Lemma 3. In other words, the factorisation $f = gh$ no longer gives rise to a natural factorisation $f^* = g^*h^*$. This view is supported by the observation that, in the limit $v(f - f^*) = \mathscr{B}$, the bound on the error on g in the example above coincides with the bound on the error on the roots of g.

7 Krasner's lemma

The well-known Krasner's lemma (see Corollaire 1, page 190 of Ribenboim (1968), for instance) was in fact found by Ostrowski already in 1917. We give here a generalisation that will be used in the next section.

THEOREM 6 (lemma à la Krasner). *Consider a monic polynomial $f^* = \prod_{k=1}^n (X - \alpha_k^*)$ of degree $n > 1$ with coefficients in a Henselian field (K, v) and roots in the algebraic closure $\tilde{K}$. Let I and J be two disjoint, non-empty sets with union $\{1, \ldots, n\}$. Moreover, consider a polynomial $g = \prod_{i \in I}(X - \alpha_i)$ with coefficients and roots in $\tilde{K}$. Assume*

$$\forall i \in I \forall j \in J : v(\alpha_i - \alpha_i^*) > v(\alpha_i^* - \alpha_j^*) \ . \tag{7}$$

Then the coefficients of the polynomials $g^* := \prod_{i \in I}(X - \alpha_i^*)$ and $h^* := \prod_{j \in J}(X - \alpha_j^*)$ are contained in the field extension of K generated by the coefficients of g.

Proof. Part A. First some preliminary observations. From (7) follows at once that g^* and h^* are relatively prime. Since $f^* = g^*h^*$, the coefficients of g^* generate the same extension of K as the coefficients of h^*. We may assume without loss of generality

– and will do so – that g has coefficients in K. What is left to prove is that g^* has coefficients in K.

Now let K_{sep} be the separable algebraic closure of K. Since K_{sep} is a separably closed field, every irreducible polynomial over K_{sep} has only one (possibly multiple) root. Since g^*h^* has coefficients in K_{sep}, and g^* and h^* are relatively prime, it follows that g^* and h^* have coefficients in K_{sep}.

We show in part B that every K-automorphism σ on $\tilde{K}$ permutes the roots of g^*. Hence, every such σ fixes the coefficients of g^*. The coefficients of g^* are therefore purely inseparable over K.

Since the coefficients of g^* are both separable and purely inseparable over K, they do in fact belong to K.

Part B. Let σ be a K-automorphism on $\tilde{K}$. Consider the sets $A = \{\alpha_i \mid i \in I\}$, $A^* = \{\alpha_i^* \mid i \in I\}$, and $A^{**} = \{\alpha_j^* \mid j \in J\}$. Note that $A \cup A^*$ and A^{**} are disjoint by (7). Since g and f^* have coefficients in K, σ is a "multiplicity-preserving" permutation on both A and $A^* \cup A^{**}$. Since K is Henselian, σ is isometric. We show that (7) implies that σ permutes A^* and A^{**} individually. This is really a lemma on finite ultra-metric spaces.

For $\alpha \in A$, let $\mathcal{B}(\alpha)$ be the maximal ball in the finite ultra-metric space $A \cup A^* \cup A^{**}$ containing α and being contained in $A \cup A^*$. Then (7) implies

$$\forall i \in I : \alpha_i \in \mathcal{B}(\alpha) \;\Leftrightarrow\; \alpha_i^* \in \mathcal{B}(\alpha)\ . \tag{8}$$

Every $\alpha^* \in A^*$ is thereby contained in some $\mathcal{B}(\alpha)$, so we are done if we can show $\sigma(\mathcal{B}(\alpha)) \subsetneqq A \cup A^*$.

For any $\alpha_i \in A \cap \sigma(\mathcal{B}(\alpha))$, the balls $\sigma(\mathcal{B}(\alpha))$ and $\mathcal{B}(\alpha_i)$ have non-empty intersection (both contain α_i), hence one is contained in the other. If there is an $\alpha_i \in A \cap \sigma(\mathcal{B}(\alpha))$ such that $\sigma(\mathcal{B}(\alpha)) \subsetneqq \mathcal{B}(\alpha_i)$, then $\sigma(\mathcal{B}(\alpha)) \subsetneqq A \cup A^*$ and we are done. So assume from now on $\sigma(\mathcal{B}(\alpha)) \supset \mathcal{B}(\alpha_i)$ for all $\alpha_i \in A \cap \sigma(\mathcal{B}(\alpha))$.

For a subset X of $A \cup A^* \cup A^{**}$, let $\#X$ denote X's cardinality "counted with multiplicity", i.e.

$$\#X := |\{i \in I \mid \alpha_i \in X\}| + |\{k \in I \cup J \mid \alpha_k^* \in X\}|\ .$$

We then have

$$\#\mathcal{B}(\alpha) = 2 \cdot |\{i \in I \mid \alpha_i \in \mathcal{B}(\alpha)\}|$$

by (8). Since σ preserves multiplicity and permutes A,

$$\#\mathcal{B}(\alpha) = \#\sigma(\mathcal{B}(\alpha)) \;\text{ and }\; |\{i \in I \mid \alpha_i \in \mathcal{B}(\alpha)\}| = |\{i \in I \mid \alpha_i \in \sigma(\mathcal{B}(\alpha))\}|$$

hold. For $i \in I$ with $\alpha_i \in \sigma(\mathcal{B}(\alpha))$, (8) implies $\alpha_i^* \in \mathcal{B}(\alpha_i) \subset \sigma(\mathcal{B}(\alpha))$ and hence

$$|\{i \in I \mid \alpha_i^* \in \sigma(\mathcal{B}(\alpha))\}| \geq |\{i \in I \mid \alpha_i \in \sigma(\mathcal{B}(\alpha))\}|\ .$$

Putting everything together gives

$$\begin{aligned}
\#\sigma(\mathcal{B}(\alpha)) &= |\{i \in I \mid \alpha_i \in \sigma(\mathcal{B}(\alpha))\}| + |\{k \in I \cup J \mid \alpha_k^* \in \sigma(\mathcal{B}(\alpha))\}| \\
&\geq 2 \cdot |\{i \in I \mid \alpha_i \in \sigma(\mathcal{B}(\alpha))\}| + |\{j \in J \mid \alpha_j^* \in \sigma(\mathcal{B}(\alpha))\}| \\
&= 2 \cdot |\{i \in I \mid \alpha_i \in \mathcal{B}(\alpha)\}| + |\{j \in J \mid \alpha_j^* \in \sigma(\mathcal{B}(\alpha))\}| \\
&= \#\mathcal{B}(\alpha) + |\{j \in J \mid \alpha_j^* \in \sigma(\mathcal{B}(\alpha))\}| \\
&= \#\sigma(\mathcal{B}(\alpha)) + |\{j \in J \mid \alpha_j^* \in \sigma(\mathcal{B}(\alpha))\}| \ .
\end{aligned}$$

Finally, conclude $|\{j \in J \mid \alpha_j^* \in \sigma(\mathcal{B}(\alpha))\}| = 0$, i.e. $\sigma(\mathcal{B}(\alpha)) \subsetneqq A \cup A^*$. □

Theorem 6 has an immediate corollary which itself reduces to the usual Krasner's lemma when the element a is separable over K:

Corollary 7. *Consider a Henselian field K and let a and b be elements in the algebraic closure $\tilde{K}$. Assume b is closer to a than to any of a's conjugates. Then $K(b)$ contains the coefficients of the polynomial $(X-a)^\mu$ where μ is the root multiplicity of a in its minimal polynomial over K.* □

Remark. In the application of Theorem 6 in the proof of Theorem 8 below, we also have a polynomial $h = \prod_{j \in J}(X - \alpha_j)$ satisfying

$$\forall i \in I \forall j \in J : v(\alpha_j - \alpha_j^*) > v(\alpha_i^* - \alpha_j^*) \tag{9}$$

and such that gh has coefficients in K. In this situation, part B of the proof of Theorem 6 can be replaced by the following simpler argument: Assume for a contradiction that there are $i \in I$ and $j \in J$ such that $\sigma(\alpha_i^*) = \alpha_j^*$. By symmetry, we may assume $v(\alpha_i - \alpha_i^*) \geq v(\alpha_j - \alpha_j^*)$ Then $\sigma(\alpha_i) = \alpha_{i'}$ for some $i' \in I$. Since σ is isometric, $v(\alpha_{i'} - \alpha_j^*) = v(\alpha_i - \alpha_i^*) \geq v(\alpha_j - \alpha_j^*)$. But now $v(\alpha_{i'} - \alpha_j) \geq v(\alpha_j - \alpha_j^*)$, in contradiction with (9).

8 Hensel's lemma

We can now state and prove the promised general Hensel's lemma.

Theorem 8 (monic Hensel's lemma). *Consider two monic polynomials f and f^* of common degree $n > 1$ with integral coefficients in a Henselian field (K, v). Let there be given a factorisation $f = gh$ with monic g and h. Assume $v(f - f^*) > \mathscr{B}$ where $\mathscr{B}$ is the bipartitionant of g and h. Then there is a factorisation $f^* = g^*h^*$ where g^* and h^* are monic and have integral coefficients, $\deg(g^*) = \deg(g)$, $\deg(h^*) = \deg(h)$, and $v(g - g^*), v(h - h^*) \geq v(f - f^*) - \mathscr{B}$.*

Proof. Consider the induced factorisation $f^* = g^*h^*$. The factors g^* and h^* have coefficients in K by Lemma 3 and Theorem 6. The bound on $v(g-g^*)$ and $v(h-h^*)$ follows from Theorem 5. □

Example. Consider the polynomial $f^* = X^8(X+2)^8 + 2^\nu$ with $\nu \geq 0$ over the field of dyadic numbers $\mathbb{Q}_2$. The bipartitionant of $g = X^8$ and $h = (X+2)^8$ is $\mathscr{B} = 16$. By Theorem 8, f^* is reducible for all $\nu > 16$. More precisely, there is in this case a monic factorisation $f^* = g^*h^*$ with $v(g-g^*), v(h-h^*) \geq \nu - 16$ (using Lemma 4 instead of Theorem 5 gives in fact $v(g-g^*), v(h-h^*) \geq \nu - 15$). It can be shown that f^* is irreducible for $\nu = 0, 1, 3, 4, 5, 7, 8, 9, 11, 12, 13, 15, 16$, implying that the bound $v(f-f^*) > \mathscr{B}$ is best possible. The dyadic value of the resultant of g and h is 64, so the Hensel's lemma of 1908 gives a factorisation $f^* = g^*h^*$ with $v(g-g^*), v(h-h^*) \geq \nu - 64$ for $\nu > 128$. □

To make life as easy as possible, we have so far solely studied monic polynomials having integral coefficients. This is indeed the situation in almost all applications of Hensel's lemma. Also, when a given non-monic polynomial F^* has an approximate factorisation satisfying the conditions of the non-monic Hensel's lemma, the reducibility of F^* follows immediately from the observation that the Newton polygon of F^* is not a straight line.

Nevertheless, we now turn our attention to the non-monic case. The proof of the following theorem is entirely analogous to that of the monic Hensel's lemma, but the presence of non-monic polynomials forces us to reexamine the proofs of earlier theorems.

Theorem 9 (Hensel's lemma, final form). *Consider two polynomials F and F^* of common degree $n > 1$ with integral coefficients in a Henselian field (K, v) and with the same leading coefficient c. Let there be given a factorisation $F = gH$ where g is monic and has integral coefficients, and H is primitive, i.e. $v(H) = 0$. Assume $v(F-F^*) > \max\{0, \mathscr{B} + v(c)\}$ where $\mathscr{B}$ is the bipartitionant of g and $c^{-1}H$. Then there is a factorisation $F^* = g^*H^*$ where g^* is monic and has integral coefficients, H^* is primitive, $\deg(g^*) = \deg(g)$, $\deg(H^*) = \deg(H)$, and $v(g-g^*), v(H-H^*) \geq v(F-F^*) - \max\{0, \mathscr{B} + v(c)\}$.*

Proof. First introduce monic polynomials $f := c^{-1}F$, $f^* := c^{-1}F^*$, and $h := c^{-1}H$. Note $f = gh$, $v(f-f^*) = v(F-F^*) - v(c)$, and thus $v(f-f^*) > \max\{-v(c), \mathscr{B}\}$.

Write $f = \prod_{k=1}^n (X - \alpha_k)$ and let I and J be the sets with $g = \prod_{i\in I}(X-\alpha_i)$ and $h = \prod_{j\in J}(X-\alpha_j)$. Put $\rho_i := \Phi_i^{-1}(v(f-f^*))$ for each $i \in I$. Note $\Phi_i(0) = \sum_{l=1}^n \min\{0, v(\alpha_i - \alpha_l)\} = -v(c) < v(f-f^*)$ and hence $0 < \rho_i$.

The proof of Theorem 2, word for word, shows that f and f^* have the same

number of roots (counted with multiplicity) in the ball $\{x \in \tilde{K} \mid v(x-\alpha_i) \geq \rho_i\}$ for any $i \in I$. It follows that we can write $f^* = \prod_{k=1}^{n}(X-\alpha_k^*)$ such that $v(\alpha_i - \alpha_i^*) \geq \rho_i$ for each $i \in I$. We have $v(\alpha_i - \alpha_j) \leq \Phi_i^{-1}(\mathscr{B}) < \rho_i$ for $i \in I$ and $j \in J$, and therefore $v(\alpha_i^* - \alpha_j^*) < \rho_i$ for $i \in I$ and $j \in J$. Conclude $v(\alpha_i - \alpha_i^*) > v(\alpha_i^* - \alpha_j^*)$ for all $i \in I$ and $j \in J$.

By Theorem 6, $g^* := \prod_{i\in I}(X - \alpha_i^*)$ and $h^* := \prod_{j\in J}(X - \alpha_j^*)$ have coefficients in K. Reexamination of the proofs of Lemma 4 and Theorem 5 shows $v(g - g^*) \geq v(f - f^*) - \max\{-v(c), \mathscr{B}\}$. Now put $H^* := ch^*$. □

Notice that the resultant of g and H has value

$$\begin{aligned} v(\mathrm{Res}(g,H)) &= \deg(g)\cdot v(c) + v(\mathrm{Res}(g,h)) \\ &= \deg(g)\cdot v(c) + \sum_{i\in I, j\in J} v(\alpha_i - \alpha_j) \\ &= \sum_{i\in I, j\in J} \max\{0, v(\alpha_i - \alpha_j)\} \end{aligned}$$

By (5), the bipartitionant of g and h is $\mathscr{B} = \sum_{i\in I} v(\alpha_i - \alpha_{j_0}) + \sum_{j\in J} v(\alpha_{i_0} - \alpha_j)$ for suitable $i_0 \in I$ and $j_0 \in J$. There follows $\max\{0, \mathscr{B} + v(c)\} \leq 2v(\mathrm{Res}(g,H))$. Hence, Theorem 9 generalises the Hensel's lemma of 1908 as well as its in section 1 mentioned later reincarnations.

References

[1] G. Dumas, *Sur quelques cas d'irréductibilité des polynomes à coefficients rationnels*, J. Math. Pures Appl. **61** (1906), 191–258.

[2] K. Hensel, *Neue Grundlagen der Arithmetik*, J. Reine Angew. Math. **127** (1904), 51–84.

[3] K. Hensel, *Theorie der algebraischen Zahlen*, Teubner, Leipzig, 1908.

[4] W. Krull, *Allgemeine Bewertungstheorie*, J. Reine Angew. Math. **167** (1932), 160–196.

[5] J. Kürschák, *Über Limesbildung und allgemeine Körpertheorie*, J. Reine Angew. Math. **142** (1913), 211–253.

[6] M. Nagata, *On the Theory of Henselian Rings*, Nagoya Math. J. **5** (1953), 45–57.

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

[8] F. J. Rayner, *Relatively Complete Fields*, Proc. Edinburgh Math. Soc. **11** (1958), 131–133.

[9] T. Rella, *Zur Newtonschen Approximationsmethode in der Theorie der p-adischen Gleichungswurzeln*, J. Reine Angew. Math. **153** (1924), 111–112.

[10] T. Rella, *Ordnungsbestimmungen in Integritätsbereichen und Newtonsche Polygone*, J. Reine Angew. Math. **158** (1927), 33–48.

[11] P. Ribenboim, *Théorie des valuations*, Les presses de l'Université de Montréal, Montreal, 1968.

[12] P. Ribenboim, *Equivalent forms of Hensel's lemma*, Expo. Math. **3** (1985), 3–24.

[13] D. S. Rim, *Relatively complete fields*, Duke Math. J. **24** (1957), 197–200.

[14] P. Roquette, *History of Valuation Theory. Part I.* In: F. V. Kuhlmann, S. Kuhlmann, M. Marshall (ed.), *Valuation Theory and its applications, vol. 1*, Fields Inst. Commun. **32** (2002), 291–355.

[15] K. Rychlík, *Zur Bewertungstheorie der algebraischen Körper*, J. Reine Angew. Math. **153** (1924), 94–107.

David Brink
Department of Mathematics
Universitetsparken 5
2100 Copenhagen
Denmark
`brink@math.ku.dk`