Hensels Lemma

Claudia Strauch, Andrea Staudenmayer,
Katja Setzer, Gertraud Johne

Seminar Algebra SS 2005
PD Dr. Koenigsmann

Zunächst zur Erinnerung:

Das Newton-Verfahren im Körper der reellen Zahlen

Sei φ eine gegebene Funktion in einer reellen Variablen. Wir betrachten die Iterationen von φ:

$$\begin{aligned} \varphi^2(x) &= \varphi(\varphi(x)) \\ \varphi^{n+1}(x) &= \varphi(\varphi^n(x)), \quad n \geq 0 \end{aligned}$$

Dieser Iterationsprozess ermöglicht es uns, einen Fixpunkt zu finden, d.h. $\alpha \in \mathbb{R}$ mit $\varphi(\alpha) = \alpha$. Damit die Iterationsfolge gegen einen Fixpunkt konvergiert, müssen wir annehmen, dass die Funktion φ eine Kontraktion ist, d.h. φ ist differenzierbar auf einem Intervall J und es gibt eine Zahl $\rho < 1$, so dass die Ableitung von φ die Bedingung $|\varphi'(x)| < \rho < 1 \quad \forall x \in J$ erfüllt.

1. Lemma

> Ist φ eine Kontraktion auf einem Intervall J und besitzt φ einen Fixpunkt α in J, so konvergiert die Iterationsfolge $x_n = \varphi^n(x_0), \quad n = 1, 2, ...$ für jeden Startpunkt $x_0 \in J$ gegen α. Insbesondere (Start mit $x_0 = \alpha$) ist α der einzige Fixpunkt von φ in J.

Beweis: Laut ZWS gilt $\varphi(x_0) - \varphi(\alpha) = \varphi(x_0) - \alpha = (x_0 - \alpha)\varphi'(\tilde{x}_0)$, wobei $x_0 < \tilde{x}_0 < \alpha$. Daher ist $|x_1 - \alpha| < \rho|x_0 - \alpha|$, d.h. x_1 liegt näher bei α als x_0. Indem wir dasselbe Argument wiederholt anwenden, erhalten wir $|x_n - \alpha| < \rho^n|x_0 - \alpha| \quad \forall n \geq 1$. Somit konvergiert die Folge gegen α.

Wenden wir dieses Resultat nun auf das Problem der Bestimmung einer Nullstelle einer Funktion f auf einem Intervall J an. Wir nehmen an, dass f entweder monoton steigend oder fallend auf dem gesamten Intervall ist, d.h., dass f' differenzierbar ist und keine Nullstelle in J hat. Wir können sicher sein, dass f eine Nullstelle in J hat, wenn sich das Vorzeichen am einen Endpunkt von J vom Vorzeichen am anderen Endpunkt unterscheidet. Um diese Nullstelle beliebig nah zu approximieren, betrachten wir die Funktion

$\varphi(x) = x - \frac{f(x)}{f'(x)}$ und bemerken, dass die Nullstelle von f der Fixpunkt von φ ist. Wir können dann obigen Iterationsprozess auf φ anwenden, vorausgesetzt φ ist eine Kontraktion. Die Anwendung des obigen Lemmas führt dann zu folgendem Ergebnis:

2. Newton'sches Lemma

> Die Funktion f habe eine Nullstelle im Intervall J; ihre Ableitung sei differenzierbar und nullstellenfrei auf J. Ferner existiere $\rho < 1$, so dass
>
> $$\left|\frac{f(x)f''(x)}{f'(x)^2}\right| < \rho \quad \forall x \in J$$
>
> gilt. Dann konvergiert die Folge $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, ...$ für jeden Startpunkt $x_0 \in J$ gegen die einzige Nullstelle von f in J.

Folgende Bezeichnungen sollen für alle kommenden Sätze in dieser Bedeutung verwendet werden: Sei R ein **vollständiger, diskreter Bewertungsring** mit **Bewertung** v, **Uniformisierender** π, **Restklassenkörper** k und **Quotientenkörper** K.
Wähle eine feste reelle Zahl $\gamma > 1$, eine Standardwahl ist $\gamma = e$ oder in $K = \mathbb{Q}_p \quad \gamma = p$. Definiere einen Absolutbetrag auf K durch $|x| = \gamma^{-v(x)}$ für $x \neq 0$ und setze $|0| = 0$. Diese Funktion erfüllt

$$|xy| = |x||y| \tag{1}$$

$$|x + y| \leq max\{|x|, |y|\} \quad \text{(ultrametrische Dreiecksungleichung)}, \tag{2}$$

wie unmittelbar aus der Definition der Bewertung folgt. Ein Absolutbetrag, der die ultrametrische Dreiecksungleichung erfüllt, heißt **nicht-archimedisch**.
Für $|x| \neq |y|$ gilt wegen $v(x) \neq v(y) \Rightarrow v(x + y) = min\{v(x), v(y)\}$ sogar $|x + y| = \gamma^{-min\{v(x),v(y)\}} = max\{|x|, |y|\}$. In jedem Fall ist aber $|x + y| \leq |x| + |y|$, sodass wir durch die Definition des Abstandes zwischen x und y als $d(x, y) := |x - y|$ einen metrischen Raum K erhalten.
Beachte: $x \in R \Leftrightarrow v(x) \geq 0 \Leftrightarrow |x| = \gamma^{-v(x)} \leq 1$

Zur Erinnerung: Ein metrischer Raum ist eine nichtleere Menge X zusammen mit einer Metrik d auf X. Eine Abbildung $d: X \times X \longrightarrow \mathbb{R}$ heißt Metrik auf X, falls $\forall x, y, z \in X$ gilt:

(1)	Positivität	$d(x,y) \geq 0$
(2)	Definitheit	$d(x,y) = 0 \Leftrightarrow x = y$
(3)	Symmetrie	$d(x,y) = d(y,x)$
(4)	Dreiecksungleichung	$d(x,z) \geq d(x,y) + d(y,z)$

3. Proposition

R ist genau dann ein vollständiger, diskreter Bewertungsring, wenn K ein vollständiger metrischer Raum ist.

Beweis: Sei (x_n) eine Folge von Elementen aus K. Diese Folge konvergiert gegen einen Grenzwert $x \in K$, wenn für jedes reelle $\varepsilon > 0$ eine ganze Zahl $N = N(\varepsilon)$ existiert mit $|x - x_n| < \varepsilon$, d.h. mit den Bezeichnungen der Bewertung $\gamma^{-v(x-x_n)} < \varepsilon \quad \Leftrightarrow \quad v(x-x_n) > -log_\gamma \varepsilon$. Wählt man ε genügend klein und setzt man $\nu = [-log_\gamma \varepsilon]$ (Gaußklammer), so bedeutet dies, dass die Entwicklungen von x und x_n in Potenzen von π mit derselben Potenz beginnen und bis zur $\gamma-$ten Potenz übereinstimmen. Sei (x_n) jetzt eine Cauchy-Folge von Elementen aus K. Für jedes gegebene $\varepsilon > 0$ gibt es dann eine ganze Zahl $N = N(\varepsilon)$, sodass $|x_m - x_n| < \varepsilon \quad \forall m \geq N,\ n \geq N$. Für jedes vorgegeben $\nu \in \mathbb{Z}$ haben also (setze $\varepsilon = \gamma^{-\nu}$) alle Terme der Folge vom $N-$ten Term an dieselbe Entwicklung in Potenzen von π bis hin zur $\nu-$ten Potenz.

Ist R vollständig, so stellt die derart bestimmte Entwicklungsreihe ein Element $x \in K$ dar, welches der Grenzwert dieser Folge ist, also ist K vollständig.

Umgekehrt legt jede Entwicklungsreihe in nichtnegativen Potenzen von π, die nach der n-ten Potenz abbricht $x_n = \sum_{i=1}^{n} a_i \pi^i$, ein Element $x \in R$ fest; die Folge (x_n) ist offensichtlich eine Cauchy-Folge. Wenn K vollständig ist, lässt sich zeigen, dass diese Entwicklung ein Element von R darstellt. Also ist R vollständig.

Während K dem Körper der reellen Zahlen, in welchem er ein vollständiger metrischer Raum ist, ähnelt, ist die Geometrie in K eine völlig andere - wegen der Ungleichung $|x+y| \leq max\{|x|,|y|\}$.

- Jedes Dreieck in K ist gleichschenklig.

  Wir wählen z. B. 3 Punkte $x, y, z \in K$ und nehmen an, dass $z = 0$ sei. Falls $|x| \neq |y|$, o.B.d.A. $|x| < |y|$, so folgt jetzt $|x - y| = |y|$.

- Jedes Intervall ist offen und abgeschlossen; d.h. K ist völlig unzusammenhängend.

  Wenn x durch K läuft, so bilden die Werte von $|x|$ nämlich eine diskrete Teilmenge von $\mathbb{R}$:

$$\begin{aligned} & |x-a| < \gamma^{-n} \\ \Leftrightarrow \quad & v(x-a) > n \\ \Leftrightarrow \quad & v(x-a) \geq n+1 \\ \Leftrightarrow \quad & |x-a| \geq \gamma^{-(n+1)} \end{aligned}$$

- Wenn die unendliche Reihe $\sum\limits_{n=1}^{\infty} x_n$ in K konvergiert, so folgt unmittelbar, dass $\lim\limits_{n\to\infty} x_n = 0$ gilt.

Für eine nicht-archimedische Bewertung ist dieses Kriterium sogar hinreichend für die Konvergenz:

4. Proposition

Ist K ein vollständiger Körper bzgl. eines nicht-archimedischen Absolutbetrags $|\ |$ und ist (x_n) eine Folge von Elementen aus K mit $\lim\limits_{n\to\infty} x_n = 0$, so konvergiert $\sum x_n$.

Beweis: Sei $s_n = x_1 + ... + x_n$, $s_m = x_1 + ... + x_m$, wobei $m < n$. Dann ist $|s_n - s_m| = |x_{m+1} + ... + x_n| \leq \max\limits_{m+1\leq i\leq n} |x_i| \to 0$.
Weil K vollständig ist, existiert $\lim\limits_{n\to\infty} s_n$, d.h. $\sum x_n$ konvergiert.

5. Hensel'sches Lemma

Sei f ein Polynom in einer Variablen mit Koeffizienten in R. Sei $a \in R$ so gewählt, dass $f'(a) \neq 0$ und $\left|\frac{f(a)}{f'(a)^2}\right| < 1$ gilt.
Dann konvergiert die Newton-Folge $a_{n+1} = a_n - \frac{f(a)}{f'(a)}$, $n = 0, 1, ...$, die mit $a_0 = a$ beginnt, gegen eine Nullstelle α von f in R. Diese Nullstelle erfüllt die Ungleichung $|\alpha - a| < |f'(a)|$; sie ist die einzige Nullstelle, die dieser Ungleichung genügt.

Beweis: • Es bezeichne $\delta := v(f'(a))$.

- Es gilt:

$$\begin{aligned}
\left|\frac{f(a)}{f'(a)^2}\right| &< 1 \\
\Leftrightarrow \quad \gamma^{-v(\frac{f(a)}{f'(a)^2})} &< 1 \\
\Leftrightarrow -(v(f(a)) - v(f'(a)^2)) = -v(f(a)) + 2v(f'(a)) &< 0 \\
\Leftrightarrow \quad v(f(a)) &> 2\delta \\
\Rightarrow \quad f(a) &\equiv 0 \bmod \pi^{2\delta+1} \quad (*)
\end{aligned}$$

- Weiterhin gilt:

$$\begin{aligned}
a_1 - a = \frac{f(a)}{f'(a)} &\equiv 0 \bmod \pi^{\delta+1} \\
\text{also :} \quad a_1 &\equiv a \bmod \pi^{\delta+1} \quad (**)
\end{aligned}$$

- Zeige nun induktiv:

$$\begin{aligned}
\text{(i)} \quad f(a_n) &\equiv 0 \bmod \pi^{2\delta+n+1} \\
\text{(ii)} \quad a_n &\equiv a_{n-1} \bmod \pi^{\delta+n}
\end{aligned}$$

Beachte:

$$\begin{aligned}
(i) \quad &\Leftrightarrow \quad v(f(a)) > 2\delta + n \\
&\Leftrightarrow \quad |f(a_n)| < \gamma^{-(2\delta+n)} \\
&\Rightarrow \quad |f(a_n)| \overset{n\to\infty}{\longrightarrow} 0 \\
(ii) \quad &\Leftrightarrow \quad v(a_n - a_{n-1}) > \delta + n + 1 \\
&\Leftrightarrow \quad |a_n - a_{n-1}| < \gamma^{-\delta+n-1} \\
&\Rightarrow \quad |a_n - a_{n-1}| \overset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

d.h., die Newton-Folge konvergiert gegen eine Nullstelle α von f.

Induktionsanfang: (i) $n = 0$ klar (vgl. $(*)$)

(ii) $n = 1$ klar (vgl. $(**)$)

Induktionsschritt: $n \to n + 1$

$$\begin{aligned}
(ii) \quad &\Rightarrow \quad a_n \equiv a \bmod \pi^{\delta+1} \\
&\Rightarrow \quad f'(a_n) \equiv f'(a) \bmod \pi^{\delta+1} \\
&\Rightarrow \quad v(f'(a_n)) = \delta \\
&\Rightarrow \quad \frac{f(a_n)}{f'(a_n)} = a_{n+1} - a_n \equiv 0 \bmod \pi^{\delta+n+1} \\
&\Rightarrow \quad a_{n+1} \equiv a_n \bmod \pi^{\delta+n+1}
\end{aligned}$$

(i) Laut Taylor gilt:

$$f(a_{n+1}) = f\big(a_n - \tfrac{f(a_n)}{f'(a_n)}\big) = f(a_n) - f'(a_n)\tfrac{f(a_n)}{f'(a_n)} + \left(\tfrac{f(a_n)}{f'(a_n)}\right)^2 c_n, \quad c_n \in R$$

$$\Rightarrow f(a_{n+1}) \equiv 0 \bmod \pi^{2\delta+2n+2}$$

- Zur Eindeutigkeit: Sei η Nullstelle von f mit $\eta \equiv a \bmod \pi^{\delta+1}$.

  Zeige induktiv: $\eta \equiv a_n \bmod \pi^{\delta+n+1}$ $(\Leftrightarrow |\eta - a_n| \overset{n\to\infty}{\longrightarrow} 0$, d.h. $\eta = \alpha)$

  Laut Taylor gilt wiederum:

  $0 = f(\eta) = f(a_n + \eta - a_n) = f(a_n) + f'(a_n)(\eta - a_n) + (\eta - a_n)^2 d_n$, $d_n \in R$

  $\Rightarrow -\frac{f(a_n)}{f'(a_n)} - (\eta - a_n) = \frac{(\eta - a_n)^2}{f'(a_n)} d_n \equiv 0 \bmod \pi^{\delta+2n+2}$

  $\Rightarrow \eta \equiv a_{n+1} \bmod \pi^{\delta+n+2}$

Von Hensels Lemma existieren etliche verschiedene Versionen, deshalb soll hier eine alternative Formulierung sowie die Aufsplittung in zwei Theoreme angegeben werden.

6. Hensel'sches Lemma (Version II)

Sei $f \in \mathbb{Z}_p[X]$ und sei $\bar{f}$ das Polynom, welches sich ergibt, wenn man die Koeffizienten mod p reduziert. Dann kann jede einfache Nullstelle von $\bar{f}$ in $\mathbb{Z}/p\mathbb{Z}$ zu einer Nullstelle von f in $\mathbb{Z}_p$ hochgehoben werden.

Beweis: Sei $a \in \mathbb{Z}/p\mathbb{Z}$ einfache Nullstelle von $\bar{f}$:

$$\bar{f}(a) = \bar{0} \iff f(a) \equiv 0 \bmod p \iff |f(a)|_p < 1$$
$$\bar{f}'(a) \neq \bar{0} \iff f'(a) \not\equiv 0 \bmod p \iff |f'(a)|_p = 1$$

Somit gilt: $\left|\frac{f(a)}{f'(a)^2}\right| < 1,\ f'(a) \neq 0$

Mit dem Hensel'schen Lemma folgt:

$\exists$ Nullstelle $b \in \mathbb{Z}_p$ mit $|b-a|_p < |f'(a)|_p = 1 \ \Leftrightarrow\ b \equiv a \bmod p$

Die Aussage gilt nur, wenn a eine einfache Nullstelle ist: $f(X) = X^2 + 1$ besitzt eine Nullstelle mod 2 ($a = 1$), aber keine Nullstelle in $\mathbb{Z}_2$, denn $-1 \not\equiv 1 \bmod 8$.

7. Theorem I

Sei $f(X)$ ein Polynom mit Koeffizienten in R, dem mit der Bewertung $|\ |$ assoziierten Bewertungsring im Körper K. $f(X)$ habe den Leitkoeffizienten 1. Gibt es ein $\alpha_1 \in K$ mit $|f(\alpha_1)| < 1$ und $|f'(\alpha_1)| = 1$, so konvergiert die Folge $\alpha_{n+1} = \alpha_n - \frac{f(\alpha_n)}{f'(\alpha_n)},\ n \geq 1$, gegen eine Nullstelle $\alpha \in R$ von $f(X)$.

Beweis: • Es gilt: $\alpha_1 \in R$, d.h. $|\alpha_1| \leq 1$.

Sei $f(X) = X^n + a_{n-1}X^{n-1} + ... + a_0$ mit $a_i \in R$.

$\Rightarrow |f(\alpha_1)| = |\alpha_1^n + a_{n-1}\alpha_1^{n-1} + ... + a_0| = \max\limits_{0 \leq i \leq n} |a_i \alpha_1^i| \geq |\alpha_1^n| > 1$, falls $|\alpha_1| > 1$

Widerspruch

• Laut Taylor gilt:

$$f(X+h) = f(X) + hf'(X) + h^2 g(X,h)$$

mit $g(X,h) = f_2(X) + hf_3(X) + ... + h^{n-2}f_n(X)$,

also $f(\alpha_2) = f(\alpha_1 - \frac{f(\alpha_1)}{f'(\alpha_1)}) = f(\alpha_1) - \frac{f(\alpha_1)}{f'(\alpha_1)}f'(\alpha_1) + (\frac{f(\alpha_1)}{f'(\alpha_1)})^2 g(\alpha_1, -\frac{f(\alpha_1)}{f'(\alpha_1)})$ $(*)$,

wobei $g(\alpha_1, -\frac{f(\alpha_1)}{f'(\alpha_1)}) = f_2(\alpha_1) + (\frac{-f(\alpha_1)}{f'(\alpha_1)})f_3(\alpha_1) + ...$

Es gilt:

- die Koeffizienten der f_i liegen alle in R,

- $\alpha_1 \in R$

- $|\frac{-f(\alpha_1)}{f'(\alpha_1)}| < 1$ laut Voraussetzung

$\Longrightarrow |g(\alpha_1, \frac{-f(\alpha_1)}{f'(\alpha_1)}| \leq 1$

Mit der Voraussetzung und $(*)$ folgt dann

(1) $\quad |f(\alpha_2)| \leq |f(\alpha_1)|^2 < 1.$

- Laut Taylor gilt außerdem:

$f'(X+h) = f'(X) + hF(X,h)$, also

$f'(\alpha_2) = f'(\alpha_1 - \frac{f(\alpha_1)}{f'(\alpha_1)}) = f'(\alpha_1) + (-\frac{f(\alpha_1)}{f'(\alpha_1)})F(\alpha_1, -\frac{f(\alpha_1)}{f'(\alpha_1)})$

$\Longrightarrow \quad$ (2) $\quad |f'(\alpha_2)| = |f'(\alpha_1)| = 1$

Wegen (1) und (2) erfüllt α_2 wieder die beiden an α_1 gestellten Bedingungen, so dass wir den Iterationsprozess fortsetzen können:

- Es ergibt sich:

$$\begin{aligned} |\alpha_2 - \alpha_1| &= |f(\alpha_1)| \\ |\alpha_3 - \alpha_2| &= |f(\alpha_2)| \leq |f(\alpha_1)|^2 \\ |\alpha_4 - \alpha_3| &\leq |f(\alpha_1)|^4 \\ &\vdots \\ |\alpha_n - \alpha_{n-1}| &\leq |f(\alpha_1)|^{2^{n-2}} \end{aligned}$$

- Setze $\alpha := \alpha_1 + (\alpha_2 - \alpha_1) + (\alpha_3 - \alpha_2) + ... = \lim \alpha_n$

Dann gilt:

- $|\alpha_n - \alpha_{n-1}| \to 0$ für $n \to \infty \overset{\text{4.Prop.}}{\Longrightarrow}$ Die Reihe konvergiert.

- $|f(\alpha_n)| = |\alpha_{n+1} - \alpha_n| \leq |f(\alpha_1)^{2^{n-1}}| \;\Rightarrow\; f(\alpha) = 0$

- $|\alpha - \alpha_1| = \lim |\alpha_n - \alpha_1|$ sowie $|\alpha_n - \alpha_1| \leq \max\limits_{2 \leq i \leq n} |\alpha_i - \alpha_{i-1}| < 1$

$\Rightarrow\ |\alpha - \alpha_1| = \max\{|\alpha|, |\alpha_1|\} \leq 1$, also $\alpha \in R$

Als Anwendung wollen wir feststellen, wann die Gleichung $f(X) = X^2 - a = 0$, $a \in \mathbb{Z}$, in $\mathbb{Q}_p$ gelöst werden kann. p sei dabei eine Primzahl, für die $p^2 \nmid a$ gilt.

Fallunterscheidung:

1. Fall: $p|a$, d.h. $a = mp,\ m \in \mathbb{Z}$ und o.B.d.A. $p \nmid m$.

Hat $X^2 - a$ eine Nullstelle $\sqrt{a} \in \mathbb{Q}_p$, so ist $\sqrt{|a|_p} = |\sqrt{a}|_p$, denn

$|a|_p = |\sqrt{a}\ \sqrt{a}|_p = |\sqrt{a}|_p |\sqrt{a}|_p$.

$\Rightarrow |\sqrt{a}|_p = |m^{\frac{1}{2}} p^{\frac{1}{2}}|_p = |p|_p^{\frac{1}{2} \notin \mathbb{Z}}$ (da $p \nmid m$, also $|m| = |m^{\frac{1}{2}}| = 1$), im Widerspruch zu

$|\mathbb{Q}_p|_p = \{|p|_p^n \quad | \ n \in \mathbb{Z}\}$

$\Rightarrow$ also ist $\sqrt{a} \notin \mathbb{Q}_p$

2. Fall: $p \nmid a,\ p \neq 2$

Bestimme $\alpha \in \mathbb{Z}$ mit

$|f(\alpha)|_p = |\alpha^2 - a|_p < 1 \quad \overset{v(\alpha^2 - a) > 0}{\Longleftrightarrow} \quad p|\alpha^2 - a \quad \Leftrightarrow \quad \alpha^2 \equiv a \bmod p$ und

$|f'(\alpha)|_p = |2\alpha|_p = 1 \quad \Leftrightarrow \quad p \nmid 2\alpha$

(a) Gelte $\alpha^2 \equiv a \bmod p$

$\Rightarrow p \mid \alpha^2 - a$, aber (laut Vorauss.) $p \nmid a$

$\Rightarrow p \nmid \alpha \overset{p \neq 2}{\Rightarrow} p \nmid 2\alpha$

(b) Gelte $b^2 \not\equiv a \bmod p \quad \forall b \in \mathbb{Z}$ und $\alpha^2 - a = 0$ für ein $\alpha \in \mathbb{Q}_p$.

$\Rightarrow \alpha + J_1 = b + J_1$ für ein $b \in \mathbb{Z}$, wobei $J_1 = \langle p \rangle$ das von p erzeugte Primideal ist

$\Rightarrow a + J_1 = \alpha^2 + J_1 = b + J_1$

$\Rightarrow b^2 \equiv a \bmod p$ Widerspruch

Zusammenfassend erhalten wir also:

8. Korollar

Ist p eine ungerade Primzahl, so hat die Gleichung $X^2 - a = 0$, wobei $a \in \mathbb{Z}$ und $p^2 \nmid a$,

(1) keine Lösung in $\mathbb{Q}_p$, falls $p|a$,

(2) zwei (a ist quadratischer Rest modulo p) oder keine Lösung (a ist quadratischer Nichtrest modulo p, d.h. $b^2 \not\equiv a \bmod p \ \forall b \in \mathbb{Z}$) in $\mathbb{Q}_p$, falls $p \nmid a$.

Übungsaufgabe: Sei p eine ungerade Primzahl. Hat die Funktion $f(X) := X^2 - b$, $b \in \mathbb{Z}$, eine Nullstelle in $\mathbb{Z}/p\mathbb{Z}$, so besitzt sie auch eine Nullstelle in $\mathbb{Z}/p^m\mathbb{Z}$, $m \geq 1$.

- Gelte also $a \equiv b^2 \bmod p$ für ein $a \in \mathbb{Z}$.
- Sei $r \in \mathbb{Z}$ Primitivwurzel modulo p^m, d.h. $\langle r \bmod p^m \rangle = (\mathbb{Z}/p^m\mathbb{Z})^\times$.

  $\Rightarrow b \equiv r^s \bmod p^m$, $s \in \mathbb{Z}$

  Z. z.: *s ist gerade*
- Es gilt: $b \equiv r^s \bmod p$, und laut Vorauss. $b \equiv (\underbrace{r^t}_{=a})^2 \bmod p$

  $\Rightarrow b \equiv r^{s-2t} \bmod p$

  $\Rightarrow s - 2t \mid p - 1 = |(\mathbb{Z}/p^m\mathbb{Z})^\times|$

  $\Rightarrow s$ ist gerade.

9. Theorem II

Sei $f(X)$ ein Polynom mit Koeffizienten in R, dem mit der Bewertung $|\ |$ assoziierten Ring im vollständigen Körper K. Der Leitkoeffizient von $f(X)$ sei 1. Gibt es ein $\alpha_1 \in K$ mit $|f(\alpha_1)| < 1$, $|f(\alpha_1)| \neq 0$, $|f'(\alpha_1)| \leq 1$ und für $d_1 := \frac{f(\alpha_1)}{f'(\alpha_1)^2}$, $|d_1| < 1$, so konvergiert die Newton-Folge gegen eine Nullstelle $\alpha \in R$ von $f(X)$.

Beweis:

- Wie im Beweis von Theorem I folgt: $\alpha_1 \in R$.
- Offensichtlich ist $|\alpha_2 - \alpha_1| = |d_1||f'(\alpha_1)|$.
- Mit Taylor gilt:

  $f(\alpha_2) = f(\alpha_1 - \frac{f(\alpha_1)}{f'(\alpha_1)}) = f(\alpha_1) - \frac{f(\alpha_1)}{f'(\alpha_1)} f'(\alpha_1) + \left(\frac{f(\alpha_1)}{f'(\alpha_1)}\right)^2 \beta \qquad$ **(1)**,

wobei $\beta \in R$, da $-\frac{f(\alpha_1)}{f'(\alpha_1)} = -d_1 f'(\alpha_1) \in R$.
$\Rightarrow |f(\alpha_2)| \leq |f(\alpha_1)|\ |d_1|$

- Zudem ist $f'(\alpha_2) = f'(\alpha_1 - \frac{f(\alpha_1)}{f'(\alpha_1)}) = f'(\alpha_1) - \frac{f(\alpha_1)}{f'(\alpha_1)}\gamma$, wobei $\gamma \in R$.
  $\Rightarrow f'(\alpha_2) = f'(\alpha_1)(1 - d_1\gamma)$
  $\Rightarrow |f'(\alpha_2)| = |f'(\alpha_1)|\ \max\{|1|, |-d_1\gamma|\} = |f'(\alpha_1)|$
  $\Rightarrow f'(\alpha_2) = f'(\alpha_1)\varepsilon$, wobei ε eine Einheit ist, d.h. $|\varepsilon| = 1$. **(2)**

- Mit (1) und (2) gilt dann:
  $d_2 := \frac{f(\alpha_2)}{f'(\alpha_2)} = \frac{f(\alpha_1)^2}{f'(\alpha_1)^2}\ \beta\ \frac{1}{f'(\alpha_1)^2\varepsilon^2} = \frac{f(\alpha_1)^2}{f'(\alpha_1)^4}\ \delta$, wobei $\delta \in R$ ($|\delta| = \left|\frac{\beta}{\varepsilon^2}\right| = |\beta| \leq 1$).
  $\Rightarrow |d_2| \leq \left|\frac{f(\alpha_1)^2}{f'(\alpha_1)^4}\right| = |d_1|^2$

- Indem wir induktiv vorgehen, erhalten wir dann
  $|\alpha_{n+1} - \alpha_n| = |d_n|\ |f'(\alpha_n)| = |d_n|\ |f'(\alpha_1)| \leq |d_1|^{2^{n-1}}|f'(\alpha_1)|$.

- Setze wie zuvor $\alpha = \alpha_1 + (\alpha_2 - \alpha_1) + (\alpha_3 - \alpha_2) + ... = \lim \alpha_n$ und folgere, dass die Reihe konvergiert, $f(\alpha) = 0$ gilt und $\alpha \in R$.

Zurück zur Gleichung $X^2 - a = 0$. Wir betrachten nun den Fall $p = 2$, d.h. wir suchen eine Lösung $\sqrt{a} \in \mathbb{Q}_2$. Zusätzlich setzen wir $2 \not| a$ und $4 \not| a,\ a \in \mathbb{Z}$, voraus.
Es ist klar, dass man kein $\alpha \in \mathbb{Z}$ finden kann, das die Voraussetzung von Theorem I erfüllt:

$$|f'(\alpha)|_2 = |2\alpha|_2 = 1 \ \Leftrightarrow\ 2 \not| 2\alpha$$

Daher suchen wir stattdessen ein $\alpha \in \mathbb{Z}$ mit

- $|f(\alpha)|_2 = |\alpha^2 - a|_2 < 1 \quad \Leftrightarrow \quad 2|\alpha^2 - a$ und
  (Beachte: Mit $2 \not| a$ folgt hieraus $2 \not| \alpha^2$)

- $\left|\frac{f(\alpha)}{f'(\alpha)^2}\right|_2 = \left|\frac{\alpha^2 - a}{4\alpha^2}\right|_2 = |\alpha^2 - a|_2 \cdot \underbrace{|\frac{1}{4}|_2}_{=4} < 1$,
  d.h. $2^{-v(\alpha^2-a)} < \frac{1}{4}\ \Leftrightarrow\ v(\alpha^2 - a) > 2\ \Leftrightarrow\ 8|\alpha^2 - a$

Somit genügt es, $8|\alpha^2 - a$ zu fordern, um beide Bedingungen zu erfüllen, d.h. wir müssen die Kongruenz $X^2 \equiv a \bmod 8$ mit einem ungeraden α lösen. Für jede bel. ungerade Zahl $\alpha \in \mathbb{Z}$ gilt aber $\alpha^2 \equiv 1 \bmod 8$. Wenn also $a \equiv 1 \bmod 8$ gilt, so erfüllt jede bel. ungerade

Zahl die Voraussetzungen des Henselschen Lemmas.

Sei α umgekehrt eine Lösung von $X^2 - a = 0$ in $\mathbb{Q}_2$.

$\alpha + J_3 = b + J_3$ für ein $b \in \mathbb{Z}$, $J_3 := \langle p^3 \rangle$

$\Rightarrow \alpha^2 + J_3 = a + J_3 = b^2 + J_3$

$\Rightarrow a \equiv b^2 \equiv 1 \bmod 8$, da b ungerade sein muss.

In diesem Fall erhalten wir also:

10. Proposition

Die Gleichung $X^2 - a = 0$ hat für $a \in \mathbb{Z}$, $4 \nmid a$

(1) keine Lösung in $\mathbb{Q}_2$, falls $2|a$,

(2) zwei ($a \equiv 1 \bmod 8$) oder keine Lösung ($a \not\equiv 1 \bmod 8$) in $\mathbb{Q}_2$, falls $2 \nmid a$.

Übungsaufgabe: Zeige mit Hensels Lemma, dass $\sqrt{-1}$ in $\mathbb{Z}_{13}$ existiert.

Oder allgemeiner: Wann existiert $\sqrt{-1}$ in $\mathbb{Z}_p$?

Betrachte das Polynom $f(X) = X^2 + 1$ über dem Ring $\mathbb{Z}_p$, wobei $p > 2$ ungerade ist. Dann gilt für jedes X mit $f(X) = X^2 + 1 \equiv 0 \bmod p$, dass $f'(X) = 2X \not\equiv 0 \bmod p$ ist.

Hensels Lemma besagt dann, dass $\sqrt{-1}$ in $\mathbb{Z}_p$ genau dann existiert, wenn -1 ein quadratischer Rest mod p ist:

$\exists X \in (\mathbb{Z}/\mathbb{Z}p)^\times : \; X^2 \equiv -1 \bmod p$

$\Leftrightarrow (-1)^{\frac{p-1}{2}} \equiv 1 \bmod p$

$\Leftrightarrow p \equiv 1 \bmod 4$

Nun zur eigentlichen Übungsaufgabe:

Dann folgt: $13 = 1 + 12 \equiv 1 \bmod 4 \Rightarrow \sqrt{-1}$ existiert in $\mathbb{Z}_{13}$

$a^2 + 1 \equiv 0 \bmod 13, \; a \in \{0, ..., 12\}$

$5^2 = 25 \equiv 12 \bmod 13$ und

$12 + 1 = 13 \equiv 0 \bmod 13 \Rightarrow 5^2 \equiv -1 \bmod 13$

Man wähle also $a = 5$ und erhält $2a = 10 \not\equiv 0 \bmod 13$. Mit Hilfe der Newton-Folge mit Startwert $a = 5$ erhält man die 13-adische Entwicklung von $\sqrt{-1}$.

Berechnung der ersten zwei Terme:
$a_1 = a - \frac{f(a)}{f'(a)} = 5 - \frac{5^2+1}{2\cdot 5} = 5 - \frac{26}{10} = 5 - \frac{2}{10} \cdot 13$, d.h. $X \cdot 10 \equiv -2 \bmod 13 \Rightarrow X = 5$
Man erhält also

$$a_1 \equiv 5 + 5 \cdot 13 \pmod{13^2}$$

Der zweite Term berechnet sich mit $a_1 = 70$:
$a_2 = a_1 - \frac{f(a_1)}{f'(a_1)} = 70 - \frac{4901}{140} = 70 - \frac{29}{140} \cdot 13^2$,
d.h. $X \cdot 140 \equiv -29 \bmod 13 \Leftrightarrow X \cdot 10 \equiv -3 \bmod 13 \Rightarrow X = 1$
Man erhält

$$a_2 \equiv 5 + 5 \cdot 13 + 1 \cdot 13^2 \pmod{13^3}$$

Man sieht also, wenn man einmal die Kongruenz

$$X^2 \equiv -1 \bmod p^n$$

lösen kann, bekommt man mit dem Newton-Verfahren die Lösungen der Kongruenzen

$$X^2 \equiv -1 \pmod{p^n} \quad \text{für alle } n \in \mathbb{N}.$$

Weitere interessante Beispiele:
(1) Sei p eine Primzahl. Betrachte das Polynom $f(X) = X^{p-1} - 1$ über $\mathbb{Z}_p$.
Für jedes $a \in \mathbb{Z}_p$, das nicht durch p teilbar ist, bekommt man

$$f'(a) = (p-1)a^{p-2} \not\equiv 0 \pmod{p}$$

$$f(a) = a^{p-1} - 1 \equiv 0 \bmod p, \text{ falls } a \text{ teilerfremd zu } p$$

Mit Hensels Lemma folgt somit, dass es eine eindeutige $(p-1)$ste Einheitswurzel in $\mathbb{Z}_p$ gibt, die kongruent zu a mod p ist. Dies bestimmt eine Bijektion von der multiplikativen Gruppe $(\mathbb{Z}/\mathbb{Z}p)^\times$ in die Gruppe aller $(p-1)$sten Einheitswurzeln, d.h. $X^{p-1} - 1$ zerfällt in Linearfaktoren über $\mathbb{Z}_p$. Diese Gruppe zusammen mit der 0, nennt man die Menge der multiplikativen Repräsentanten des Restklassenkörpers $\mathbb{Z}/p$.
(2) Sei $R = k[[t]]$ und $f(X) = a_n(t)X^n + a_{n-1}(t)X^{n-1} + ... + a_1(t)X + a_0$ ein Polynom in einer Variablen X mit Koeffizienten in $k[[t]]$. Man kann $f(X)$ als Funktion $g(X,t)$ in zwei

Variablen auffassen. Die Substitution $t = 0$ liefert ein Polynom $g(X, 0) \in k[X]$ (sozusagen $f(X)$ mod t gerechnet).
Gibt es nun eine einfache Nullstelle α des Polynoms $g(X, 0)$
(d.h. $g(\alpha, 0) = a_n(0)\alpha^n + ... + a_1(0)\alpha + a_0(0) = 0$ und $\frac{\partial g}{\partial X}(\alpha, 0) = na_n(0)\alpha^{n-1} + ... + a_1(0) \neq 0$),
dann folgt mit Hensels Lemma, dass es eine eindeutige Potenzreihe $\xi(T) \in R$ gibt, so dass gilt:

$$0 = f(\xi(t)) = a_n(t)\xi(t)^n + ... + a_1(t)\xi(t) + a_0(t)$$

und

$$\alpha = \xi(0)$$

Somit haben wir für formale Potenzreihen ein Theorem über implizite Funktionen erhalten.

Zum Vergleich nochmals der Satz über implizite Funktionen aus der Analysis:

11. Satz über implizite Funktionen

Seien die Mengen $G \subset \mathbb{R}^n$, $H \subset \mathbb{R}^m$ nichtleer und offen und sei $F : G \times H \to \mathbb{R}^m$ stetig differenzierbar. Gelte für zwei Punkte $\xi \in G$ und $\eta \in H$

$$F(\xi, \eta) = 0 \text{ und } \det \frac{\partial F}{\partial y}(\xi, \eta) \neq 0.$$

Dann gibt es eine δ-Umgebung $U \subset G$ von ξ und eine ε-Umgebung $V \subset H$ von η sowie genau eine stetige Funktion

$$f : U \to V \text{ mit } f(\xi) = \eta \text{ und } F(x, f(x)) = 0 \text{ für alle } x \in U.$$

Bemerkung: Betrachtet man $k = \mathbb{C}$ und R als den Unterring von $\mathbb{C}[[t]]$, der die Potenzreihen enthält, die in einer Umgebung von 0 konvergieren, dann erhält man dasselbe Resultat. Denn obwohl R nicht vollständig ist, lässt sich zeigen, dass R die Voraussetzungen von Hensels Lemma erfüllt. Das heisst, dass das Theorem über implizite Funktionen auch für analytische Funktionen gilt.

Spezialfall: $f(X) = X^m - a(t),\ m \in \mathbb{N}$, m nicht von der Charakteristik von R teilbar

$\exists a \in k : \alpha^m = a(0) \neq 0$, $b(t) \in k[[t]]$ mit $b(t)^m = a(t),\ b(0) = \alpha$

$b(t)$ lässt sich sogar explizit angeben: Setze $a(t) = a(0) \cdot (1 + c(t))$ mit $v(c(t)) \geq 1$

(d.h. $c(t)$ ohne konstanten Term)

$\Rightarrow b(t) = \alpha \cdot (1 + c(t))^{\frac{1}{m}}$, wobei $(1 + c(t))^{\frac{1}{m}} = \sum_{n=0}^{\infty} \binom{\frac{1}{m}}{n} c(t)^n$ ist.

Dabei ist $\binom{\frac{1}{m}}{n} := \frac{\frac{1}{m}(\frac{1}{m}-1)(\frac{1}{m}-2)\cdots(\frac{1}{m}-n+1)}{n!}$.

12. Proposition

Sei k der Restklassenkörper eines vollständigen diskreten Bewertungsrings (DBR) R mit Charakteristik $k = 0$. Dann enthält R einen zu k isomorphen Unterkörper.

Beweis (indirekt mit Zorns Lemma): Sei $\phi : R \to k$ der kanonische Epimorphismus. Für jede ganze Zahl $n \neq 0$ ist $\phi(n) \neq 0$, da k Charakteristik 0 hat. Im Bewertungsring existiert genau ein maximales Ideal I_1, d.h. $n \notin I_1 \Rightarrow n \in R^\times\ \forall n \in \mathbb{Z} \setminus \{0\}$. Es sind also die ganzen Zahlen in den Einheiten von R eingebettet und R enthält somit den Körper der rationalen Zahlen $\mathbb{Q}$. Betrachte die nicht-leere Menge aller Unterkörper von R und ordne die Elemente dieser Menge bzgl. Inklusion. Dies ergibt eine partielle, induktive Ordnung.

Wendet man darauf das Zorn'sche Lemma an, erhält man einen maximalen Unterkörper k_0 von R und $\phi(k_0)$ ist dann ein Unterkörper von k, welcher isomorph ist zu k_0 (Homomorphismen zwischen Körpern sind injektiv).

Nun muss man folgende zwei Fälle betrachten:

1. Fall: k ist transzendente Erweiterung von $\phi(k_0)$.

Das bedeutet, es gibt ein Element $\tau \in k$, welches keine polynomiale Gleichung mit Koeffizienten in $\phi(k_0)$ erfüllt. Also wählen wir ein $t \in R$, sodass $\phi(t) = \tau$ gilt. Betrachte $k_0[t] \subset R$; es existiert für alle $b \in k_0[t] \setminus \{0\}$ eine Darstellung $b = \sum_{i=0}^{n} a_i t^i$, $a_i \in k_0$. Also $\phi(b) = \sum_{i=0}^{n} \phi(a_i)(\phi(t))^i = \sum_{i=0}^{n} \underbrace{\phi(a_i)}_{\in \phi(k_0)} \tau^i$.

Angenommen es gelte $b \notin R^\times$
$\Rightarrow b \in I_1, \phi(b) = 0$, d.h. $\sum_{i=0}^{n} \phi(a_i)\tau^i = 0$, was ein Widerspruch zur Transzendenz der Erweiterung ist.

$\Rightarrow b \in R^\times \ \forall b \in k_0 \setminus \{0\}$
$\Rightarrow k_0(t)$ ist ein Unterkörper von R.

Somit ist der Körper $k_0(t)$ in R enthalten und es gilt $k_0 \subset k_0(t)$. Dies ist ein Widerspruch zur Maximalität von k_0.

2. Fall: k ist algebraische Erweiterung von $\phi(k_0)$.
Annahme: $\exists \alpha \in k$ mit $\alpha \notin \phi(k_0)$

Nun sei

$$\bar{f}(X) = X^m + \phi(b_{m-1})X^{m-1} + ... + \phi(b_0)$$

das Minimalpolynom von α in $\phi(k_0)$, $b_i \in k_0$, $i = 0, ..., m$.

Weil k Charakteristik Null hat, gilt

$$0 \neq \bar{f}'(\alpha) = m\alpha^{m-1} + (m-1)\phi(b_{m-1})\alpha^{m-2} + ... + \phi(b_1) \qquad (*)$$

Wir betrachten nun das Polynom

$$f(X) = X^m + b_{m-1}X^{m-1} + ... + b_0 \in k_0[X]$$

Mit Hensels Lemma folgt, dass dieses Polynom eine Nullstelle $a \in R$ hat, sodass $\phi(a) = \alpha$ gilt:
$\exists c \in R : \phi(c) = \alpha, \ \phi(f(c)) = \bar{f}(\alpha) = 0$
$\Rightarrow f(c) \in I_1 \ \Rightarrow \ f(c) \equiv 0 \bmod I_1, \ f'(c) \notin I_1$
$\Rightarrow \exists a \in R : f(a) = 0, \ a \equiv c \bmod I_1$
$\Rightarrow \phi(a) = \phi(c) = 0$
Da $\bar{f}$ irreduzibel ist, ist auch f irreduzibel. Somit ist $k_0(a)$ ein Unterkörper von R. Dies ist erneut ein Widerspruch zur Maximalität von k_0, da $a \notin k_0$ (wegen $\alpha \notin \phi(k_0)$).

Also ist $\phi(k_0) = k$.

Bemerkung: Falls k perfekt ist, verläuft der Beweis analog. Da dann das Minimalpolynom von α in $\phi(k_0)$ separabel ist, d.h. α ist einfache Nullstelle, gilt $(*)$.

Man könnte sogar zeigen, dass die Folgerung der Proposition für einen beliebigen vollständigen DBR gilt, dessen Restklassenkörper dieselbe Charakteristik wie der Ring hat.

13. Korollar:

Wenn der Restklassenkörper k eines vollständigen DBRs R dieselbe Charakteristik wie R hat, dann ist R isomorph zum Ring der formalen Potenzreihen über k (d.h. zu $k[[t]]$).

Beweis: Man kann annehmen $k \subset R$ (aus Prop. und Bemerkung), $\pi \in R$ sei Uniformisierende. Dann ist $k[\pi] \subset R$ und für

$$f(\pi) := a_n\pi^n + ... + a_1\pi + a_0,\ a_i \in k,\ a_n \neq 0$$

ist

$$v(f(\pi)) \leq n,\ f(\pi) \neq 0.$$

Damit ist $k[\pi] \cong k[t]$ auf kanonische Art und Weise:

$$k[t] \longrightarrow k[\pi]$$

$$\sum_{i=0}^{n} a_i t^i \longmapsto \sum_{i=0}^{n} a_i \pi^i$$

Dieser Isomorphismus läßt sich zu einem Isomorphismus $k[[t]] \longrightarrow k[[\pi]] \cong R$ fortsetzen, denn jedes $r \in R$ läßt sich als Limes einer Potenzreihe mit den Repräsentanten des Restklassenkörpers darstellen und R ist vollständig.

Als nächstes wollen wir Hensels Lemma verallgemeinern, von dem Fall eines Polynoms in einer Variablen zum Fall mehrerer Polynome in mehr als einer Variablen. Zunächst wollen wir den „quadratischen“ Fall von n Polynomen in n Variablen betrachten.

Schreibweise:

$\underline{f} = (f_1, ..., f_n)$ mit $f_1, ..., f_n \in R[X_1, ..., X_n]$

$x = (x_1, ..., x_n) \in R^n$ mit $x_1, ..., x_n \in R$

$\underline{f}' = (\frac{\partial f_i}{\partial x_j})$ (Jacobi-Matrix), $\det \underline{f}' = \det(\frac{\partial f_i}{\partial x_j})$

Wir werden auch Vektoren $\underline{f}$ benötigen, deren Einträge formale Potenzreihen sind. In diesem Fall macht es keinen Sinn, $\underline{f}$, $x \in R^n$ zu betrachten, wenn $f_i(x)$ nicht konvergiert. Diese Potenzreihen konvergieren, falls jede Komponente von x durch die Uniformisierende π teilbar ist.

$f_i(x) = \sum_{i=0}^{n} a_i x^i$,

$|S_{n+1}(x) - S_n(x)| = |s_{n+1}x^{n+1}| = |\pi^{n+1}a_{n+1}y^{n+1}| = \gamma^{-v(\pi^{n+1}a_{n+1}y^{n+1})}$

$= \gamma^{-(v(\pi^{n+1})+v(a_{n+1})+v(y^{n+1}))} \leq \gamma^{-(n+1)} \overset{n\to\infty}{\longrightarrow} 0$

Wir benötigen außerdem die Komposition zweier solcher Vektoren $\underline{f}, \underline{g}$, deren Komponenten jeweils formale Potenzreihen sind.

Schreibweise: $\underline{f} \circ \underline{g} = \underline{f}(\underline{g})$,

wobei $\underline{f}(\underline{g})$ bedeutet, dass die Variable X_i von $\underline{f}$ durch $\underline{g}$ ersetzt wird. Ein sich ergebendes Problem ist dann die Bestimmung der $\underline{f}$, die unter der Komposition invertierbar sind, d.h. für welche gilt:

$\exists\, \underline{g}$ mit $\underline{f} \circ \underline{g} = X = \underline{g} \circ \underline{f}$, wobei $X = (X_1, ..., X_n)$

14. Proposition

> Sei $\underline{f}$ ein System von n formalen Porenzreihen in n Variablen ohne konstanten Term und $\det(\underline{f}'(0)) \in R^\times$.
>
> $\Rightarrow \exists!\, \underline{g}$ mit $\underline{f} \circ \underline{g} = X = \underline{g} \circ \underline{f}$

Beweis: Betrachte $Y = (Y_1, ..., Y_n)$, $X_j = \sum_{j=1}^{n} a_{ij}Y_j + \sum_{\nu} a_{i\nu}Y^\nu \quad (i \leq 1 \leq n)$,

wobei $Y^\nu = Y_1^{\nu_1}Y_2^{\nu_2} \cdot ... \cdot Y_n^{\nu_n}$, $\nu_1 + ... + \nu_n \geq 2$ und a_{ij} der Koeffizient von X_j der Reihe f_i ist.

Da $\det(\underline{f}'(0)) \in R^\times$ ist, existiert ein zu $A = (a_{ij})$ inverse Matrix $B = (b_{ij})$ mit Koeffizienten in R.

Nun schreiben wir das Gleichungssystem um in eine Vektorgleichung $X = \underline{f}(Y)$ und erhalten

$$Z = B \cdot \underline{f}(Y) \text{ mit } Z = B \cdot X$$

In dieser Gleichung ist der Koeffizient von Y_j in der i-ten Komponente δ_{ij} $\forall i, j$. Daher können wir $a_{ij} = \delta_{ij}$ $\forall i, j$ annehmen. Somit müssen wir ein System $\underline{g}$ in den Variablen X finden, so dass

$$X_i = g_i + \sum_{\nu} a_{i\nu} g_1^{\nu_1} ... g_n^{\nu_n} \qquad (1 \leq i \leq n)$$

$$\Leftrightarrow g_i = X_i - \sum_{k=2}^{\infty} \sum_{|\nu|=k} a_{i\nu} g_1^{\nu_1} ... g_n^{\nu_n}$$

$$\Rightarrow h_{i2} = X_1, \text{ wobei } h_{id} := \text{Terme von } g_i \text{ mit Grad } \leq d-1$$

Sei g_{id} := der Term von g_i, der genau Grad d hat.

Wir wollen nun induktiv die restlichen Terme von g_i bestimmen.

Induktionshypothese: h_{id} bekannt für alle i

Dann ist der Teil g_{id} vom Grad d in g_i eindeutig bestimmt als der homogene Teil der Reihe

$$-\sum_{\nu} a_{i\nu} h_{1d}^{\nu_1} ... h_{nd}^{\nu_n},$$

da die $h_{jd},\ i \leq j \leq n$ keine konstanten Terme besitzen.

$\Rightarrow \underline{f}$ hat eindeutig bestimmtes Rechtsinverses $\underline{g}$

Da $\det(\underline{g}'(0)) = \det(\underline{f}'(0))^{-1} \Rightarrow \underline{g}$ hat Rechtsinverses $\underline{e}$

Es gilt $\underline{e} = \underline{f}$, denn

$\underline{e} = X \circ \underline{e} = (\underline{f} \circ \underline{g}) \circ \underline{e} = \underline{f} \circ (\underline{g} \circ \underline{e}) = \underline{f} \circ X = \underline{f}$.

Nun können wir den „quadratischen Fall“ von Hensels Lemma zeigen, d.h. für ein System von n Polynomen in n Variablen:

15. Proposition

Sei $\underline{f} = (f_1, ..., f_n)$ mit $f_i \in R[X_1, ..., X_n]$, wobei R vollständiger DBR ist. Sei $a \in R^n$ so gewählt, dass

$$\underline{f}(a) \equiv 0 \pmod{\pi^{2\delta+1}}$$

mit $\delta = v(\det(\underline{f}'(a))) < \infty$.

Dann gibt es eine eindeutige Nullstelle $b \in R^n$ von $\underline{f}$, so dass

$$b \equiv a \pmod{\pi^{\delta+1}}.$$

Beweis: Sei N die adjungierte Matrix von $\underline{f}'(a)$.

$\Rightarrow \underline{f}'(a) \cdot N = \det(\underline{f}'(a)) \cdot I$, I sei die $n \times n$-Einheitsmatrix

Da $\delta < \infty$, ist $\det(\underline{f}'(a)) = u \cdot \pi^\delta$, $u \in R^\times$.

Die Taylorentwicklung angewandt auf jedes Polynom in $\underline{f}$ ergibt

$$\underline{f}(a + \pi^\delta X) = \underline{f}(a) + \underline{f}'(a)\pi^\delta X + \pi^{2\delta}\underline{r}(X),$$

wobei $\text{Grad}(\underline{r}(X)) \geq 2$.

Mit $\tilde{N} := u^{-1}N$ kann diese Formel umgeschrieben werden als

$$\underline{f}(a + \pi^\delta X) = \underline{f}(a) + \underline{f}'(a)\ \pi^\delta\ X + \underline{f}'(a)\ \tilde{N}\ \pi^\delta\ \underline{r}(X)$$

Setze $\underline{g}(X) := X + \tilde{N}\ \underline{r}(X)$, dann erhalten wir

$$\underline{f}(a + \pi^\delta X) = \underline{f}(a) + \underline{f}'(a)\ \pi^\delta \underline{g}(X)$$

Nun wenden wir Proposition 1 auf $\underline{g}$ an und erhalten damit ein eindeutig bestimmtes zu $\underline{g}$ inverses System $\underline{h}$ von formalen Potenzreihen ohne konstanten Term, d.h. $\underline{g} \circ \underline{h} = X$. Substituiert man X durch $\underline{h}(X)$, so erhält man

$$\underline{f}(a + \pi^\delta \underline{h}(X)) = \underline{f}(a) + \underline{f}'(a)\ \pi^\delta X$$

Sei $x \in R^n$ ein beliebiger Vektor mit $x \equiv 0 \bmod \pi$, da die Reihe dann konvergiert.

Da nach Voraussetzung $\underline{f}(a) \equiv 0 \bmod \pi^{2\delta+1}$

$\Rightarrow \underline{f}(a) = \pi^{2\delta} \cdot c$ mit einem $c \equiv 0 \bmod \pi$

Nun müssen wir ein $x \equiv 0 \bmod \pi$ finden, so dass

$$0 = \pi^{2\delta} c + \underline{f}'(a)\, \pi^{\delta}\, x$$

$$\Leftrightarrow 0 = \pi^{\delta}\, \underline{f}'(a)(\tilde{N}c + x)$$

Da $\underline{f}'(a)$ reglär ist, hat dieses Gleichungssystem eine eindeutige Lösung

$$x = -\tilde{N}c \equiv 0 \bmod \pi.$$

Also ist $b = \pi^{\delta}\underline{h}(-\tilde{N}c) + a$ der eindeutige Vektor, den wir gesucht haben.

Nun verallgemeinern wir dieses Resultat. Wir betrachten nun ein System von $n - r$ Polynomen in n Variablen $(0 \leq r < n)$.

16. Allgemeines Hensel'sches Lemma

Sei $\underline{f} = (f_{r+1}, ..., f_n)$ mit $f_i \in R[X_1, ..., X_n]$, $(r + 1 \leq i \leq n)$ und R vollständiger DBR. Sei $a \in R^n$ und $\delta \in \mathbb{N}$, so dass $\underline{f}(a) \equiv 0 \bmod \pi^{2\delta+1}$ und $\underline{f}'(a) \bmod \pi^{\delta+1}$ maximalen Rang hat.

Dann existiert ein $b \in R^n$ mit $\underline{f}(b) = 0$ und $b \equiv a \bmod \pi^{\delta+1}$.

Beweis: Wir können annehmen, dass die „Teildeterminante"

$$D = \det\left(\frac{\partial f_i}{\partial X_j}\right),\ r + 1 \leq i, j \leq n$$

nicht kongruent 0 mod $\pi^{\delta+1}$ ist. Wir erweitern das System zu einem System von n Polynomen, indem wir

$$f_i(X) = X_i - a_i$$

setzen für $1 \leq i < r$, erhalten also ein quadratisches System.

Dann ist die Jakobideterminante des erweiterten Systems gleich D. Darauf können wir die quadratische Form von Hensels Lemma anwenden.

Bemerkung: Für $r > 0$ ist die Lösung b im allgemeinen Hensel'schen Lemma nicht eindeutig. Wir können jedoch die Menge der Lösungen kongruent zu a mod $\pi^{\delta+1}$ parametrisieren.

Es ist nämlich klar (s. Beweis des quadratischen Hensel'schen Lemmas, Bezeichnungen wie dort), dass

$$g_i(X) = X_i = h_i(X),\ 1 \leq i \leq r$$

nach Definition der f_i für $1 \leq i \leq r$.

Außerdem hat die Matrix $\tilde{N}$ nun die Form $\begin{pmatrix} \pi^\delta I_{r\times r} & 0 \\ * & * \end{pmatrix}$, so dass für einen Vektor x die ersten r Koordinaten von $\tilde{N}x$ $\quad \pi^\delta x_1, ..., \pi^\delta x_r$ sind.

Die Gleichung

$$\underline{f}(a + \pi^\delta \underline{h}(X)) = \underline{f}(a) + \underline{f}'(a)\ \pi^\delta X$$

substituieren wir X durch $\tilde{N}x$, wobei $x \equiv 0$ mod π ist, so dass wir

$$f_i(a_1+\pi^{2\delta}x_1, ..., a_r+\pi^{2\delta}x_r, a_{r+1}+\pi^\delta h_{r+1}(\tilde{N}x), ..., a_n+\pi^\delta h_n(\tilde{N}x)) = f_i(a_1, ..., a_n)+\pi^{2\delta}x_i$$

erhalten für $i = 1, ..., n$.

Nach Voraussetzung ist $f_j(a) = \pi^{2\delta}c_j$, $c_j \equiv 0$ mod π für $r+1 \leq j \leq n$.

Übungsaufgabe: Gegeben seien die Gleichungen

$$\begin{aligned} X^2 + Y^2 &= 1 \\ X^3 + Y^2 &= 2 \end{aligned}$$

Bestimme die Nullstellen mod 17^4 von $f(X,Y) = \begin{pmatrix} X^2+Y^2-1 \\ X^3+Y^2-2 \end{pmatrix}$.

Startwert: $\quad X_0 = 3,\ Y_0 = 3$

denn: $\quad 3^2 + 3^2 - 1 = 18 - 1 = 17 \equiv 0$ mod 17

$\quad 3^3 + 3^2 - 2 = 36 - 2 = 34 \equiv 0$ mod 17

$$f'(X,Y) = \begin{pmatrix} 2X & 2Y \\ 3X^2 & 2Y \end{pmatrix}, \quad \det f'(X,Y) = 4XY - 6X^2Y$$

$$\Rightarrow (f'(X,Y))^{-1} = \tfrac{1}{4XY-6X^2Y} \begin{pmatrix} 2X & -2Y \\ -3X^2 & 2Y \end{pmatrix}$$

$\det f'(3,3) = -126$

$17 \nmid -126 \;\Rightarrow\; v(\det f'(3,3)) = 0 \quad (< \infty)$

$$\Rightarrow (f'(3,3))^{-1} = \tfrac{1}{-126} \begin{pmatrix} 6 & -6 \\ -27 & 6 \end{pmatrix}$$

$$\Rightarrow a_1 = a_0 - (f'(a_0))^{-1} \cdot f(a_0) = \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \tfrac{1}{126} \begin{pmatrix} 6 & -6 \\ -27 & 6 \end{pmatrix} \begin{pmatrix} 17 \\ 34 \end{pmatrix}$$
$$= \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \tfrac{17}{126} \begin{pmatrix} 6 & -6 \\ -27 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} - \tfrac{17}{126} \begin{pmatrix} 6 \\ 15 \end{pmatrix}$$

Gesucht: $y \in \{0, ..., 16\}$ mit $y \cdot 126 \equiv -1 \bmod 17$

$126 \equiv 7 \bmod 17$, also $y \cdot 7 \equiv -1 \bmod 17 \Rightarrow y \equiv 12 \bmod 17$

$$a_1 \equiv \begin{pmatrix} 3 \\ 3 \end{pmatrix} + 12 \cdot \begin{pmatrix} 6 \\ 15 \end{pmatrix} \cdot 17 \bmod 17^2 \equiv \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 10 \end{pmatrix} \cdot 17 \bmod 17^2$$

$$a_1 = \begin{pmatrix} 71 \\ 173 \end{pmatrix}$$

$$a_2 = a_1 - (f'(a_1))^{-1} \cdot f(a_1) = \begin{pmatrix} 71 \\ 173 \end{pmatrix} + \tfrac{1}{5183426} \begin{pmatrix} 346 & -346 \\ -15123 & 142 \end{pmatrix} \begin{pmatrix} 34969 \\ 387838 \end{pmatrix}$$
$$= \begin{pmatrix} 71 \\ 173 \end{pmatrix} + \underbrace{\frac{1}{5183426}}_{\equiv 5 \bmod 17} \begin{pmatrix} -422466 \\ -1639319 \end{pmatrix} \cdot 17^2$$

$$\begin{aligned} 422466 &\equiv 16 \bmod 17 \\ \Rightarrow -422466 &\equiv 1 \bmod 17 \\ 1639319 &\equiv 9 \bmod 17 \\ \Rightarrow -1639319 &\equiv 8 \bmod 17 \end{aligned}$$

$$\Rightarrow a_2 \equiv \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 10 \end{pmatrix} \cdot 17 + \begin{pmatrix} 5 \\ 6 \end{pmatrix} \cdot 17^2 \bmod 17^3$$

$$a_2 = \begin{pmatrix} 1516 \\ 1907 \end{pmatrix}, \ f(a_2) \equiv 0 \bmod 17^3$$

$$a_3 = a_2 - (f'(a_2))^{-1} \cdot f(a_2)$$

$$\vdots$$

$$a_3 \equiv \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 10 \end{pmatrix} \cdot 17 + \begin{pmatrix} 5 \\ 6 \end{pmatrix} \cdot 17^2 + \begin{pmatrix} 11 \\ 3 \end{pmatrix} \cdot 17^3 \bmod 17^4$$

$$a_3 = \begin{pmatrix} 55559 \\ 16646 \end{pmatrix}, \ f(a_3) \equiv 0 \bmod 17^4$$

Die Berechnung mod 17^4 lässt sich um einen Schritt abkürzen, wenn man die Kongruenzen bei a_2 bereits mod 17^2 berechnet:

$$a_2 = \begin{pmatrix} 71 \\ 173 \end{pmatrix} + \tfrac{1}{5183426} \begin{pmatrix} 346 & -346 \\ -15123 & 142 \end{pmatrix} \begin{pmatrix} 34969 \\ 387838 \end{pmatrix}$$

$$= \begin{pmatrix} 71 \\ 173 \end{pmatrix} + \underbrace{\frac{1}{5183426}}_{\equiv 226 \bmod 17^2} \begin{pmatrix} 346 & -346 \\ -15123 & 142 \end{pmatrix} \begin{pmatrix} 121 \\ 1342 \end{pmatrix} \cdot 17^2$$

$$= \begin{pmatrix} 71 \\ 173 \end{pmatrix} + \begin{pmatrix} 78196 & -78196 \\ -3417798 & 32092 \end{pmatrix} \begin{pmatrix} 121 \\ 1342 \end{pmatrix} \cdot 17^2$$

$$\begin{aligned} 78196 &\equiv 166 \bmod 17^2 \\ \Rightarrow -78196 &\equiv 123 \bmod 17^2 \\ -3417798 &\equiv 205 \bmod 17^2 \\ 32092 &\equiv 13 \bmod 17^2 \end{aligned}$$

$$\Rightarrow a_2 = \begin{pmatrix} 71 \\ 173 \end{pmatrix} + \begin{pmatrix} 166 & 123 \\ 205 & 13 \end{pmatrix} \begin{pmatrix} 121 \\ 1342 \end{pmatrix} \cdot 17^2 = \begin{pmatrix} 71 \\ 173 \end{pmatrix} + \begin{pmatrix} 185152 \\ 42251 \end{pmatrix} \cdot 17^2$$

$$\equiv \begin{pmatrix} 71 \\ 173 \end{pmatrix} + \begin{pmatrix} 192 \\ 57 \end{pmatrix} \cdot 17^2 \bmod 17^4 = \begin{pmatrix} 55559 \\ 16646 \end{pmatrix} \bmod 17^4 \text{ (s.o.)}$$

Literatur

Bachman, George: Introduction to p-adic numbers and valuation theory, New York, Academic Press, 1964.

Cohn, Paul M.: Algebra, London, Wiley, 1972.

Endler, Otto: Valuation theory, Berlin, Heidelberg [u.a.], Springer, 1972.

Greenberg, Marvin: Lectures on Forms in Many Variables, New York, Benjamin, 1969.

Scheja, Guenter; Storch, Uwe: Lehrbuch der Algebra 2, Stuttgart, Teubner, 1988.