Multiplizitäten

Die wohl einfachste Art und Weise, wie man die Lösungsmenge eines nichtlinearen Gleichungssystem angeben kann, ist die in einer Form gemäß dem Shape-Lemma. So eine Form läßt sich genau dann finden, wenn erstens die Lösungsmenge über dem algebraischen Abschluß des Grundkörpers endlich ist, und wenn zweitens die Polynome des Gleichungssystems ein Radikalideal erzeugen. Wenn die erste Bedingung nicht erfüllt ist, können wir nichts machen. Wenn allerdings nur die zweite nicht erfüllt ist, können wir das Ideal ersetzen durch sein Radikal, denn ein Ideal I hat die gleichen Nullstellen wie sein Radikal $\sqrt{I}$. Die Frage ist nur, wie wir dieses Radikal bestimmen können.

Betrachten wir zunächst den Fall von Polynomen einer Veränderlichen. Das von m Polynomen $f_1, \ldots, f_m \in k[X]$ erzeugte Ideal ist das vom ggT dieser Polynome erzeugte Hauptideal; wir können uns also auf Hauptideale $I = (f)$ beschränken. Ein Hauptideal (f) ist genau dann ein Radikalideal, wenn in der Darstellung von f als Produkt von Potenzen irreduzibler Polynome kein Faktor mit einer Potenz größer eins auftritt, und das ist äquivalent dazu, daß f auch in keinem algebraisch abgeschlossenen Körper K, der k enthält, mehrfache Nullstellen hat.

Um zu verstehen, wie man zum Radikal kommt, sollten wir daher zunächst wissen, wie man auch für Ideale eines Polynomrings in mehreren Veränderlichen Vielfachheiten definieren kann. Betrachten wir auch dazu zunächst wieder den Polynomring in einer Veränderlichen.

Für ein Polynom aus $\mathbb{R}[X]$ können wir Vielfachheiten über Ableitungen definieren, und um die zu berechnen, reicht es, die Funktion in einer ε-Umgebung des betrachteten Punkts z zu kennen; alles was außerhalb dieser Umgebung passiert, ist uninteressant und für die Frage nach dem Verhalten in z eher störend. Für einen beliebigen Körper k haben wir keine ε-Umgebungen, aber wir können uns trotzdem auf die Umgebung eines Punktes konzentrieren, indem wir Funktionen betrachten, die nicht unbedingt auf ganz k definiert sind, aber auf jeden Fall im Punkt z. Sind $f, g \in \mathbb{R}[X]$ zwei Polynome, so gibt es genau dann ein $\varepsilon > 0$, so daß f/g eine Funktion auf einer ε-Umgebung von z definiert, wenn $g(z)$

nicht verschwindet, und das ist eine algebraische Bedingung, die wir über jedem Körper stellen können.

Sei also k ein beliebiger Körper und $f \in k[X]$. Dann wird der Faktorring $A = k[X]/(f)$ erzeugt von den Monomen X^ℓ mit $0 \leq \ell < \deg f$.

Nun sei K ein Erweiterungskörper von k und $z \in K$ eine Nullstelle von f. Dann wird $\bar{A} = K[X]/(f)$ auch erzeugt von den Potenzen $(X - z)^\ell$ mit $0 \leq \ell < \deg f$, denn auch die bilden eine Basis des Vektorraums aller Polynome vom Grad kleiner deg f über K.

Ersetzen wir $K[X]$ durch den Ring

$$R = \left\{ \frac{g}{h} \;\middle|\; g, h \in K[X] \qquad \text{und} \qquad h(z) \neq 0 \right\} ,$$

in dem wir die Rechenoperationen nach den üblichen Regeln der Bruchrechnung definieren, erzeugt f auch dort ein Hauptideal, und wir können den Faktorring $R/(f)$ betrachten. Wir schreiben $f = (X-z)^r \cdot q$ mit einem Polynom $q \in K[X]$, das nicht in z verschwindet; die Vielfachheit der Nullstelle z von f ist also r. In R ist

$$(X - z)^r = \frac{X - z}{1} = \frac{(X - z)q}{q} = \frac{f}{q} = \frac{1}{q} \cdot f \, .$$

Somit liegt $(X - z)^r$ im vom f erzeugten Hauptideal von R, und damit natürlich auch $(X - z)^\ell$ für jedes $\ell \geq r$. Für $\ell < r$ ist $(X - z)^\ell$ kein Vielfaches von f, denn jetzt brauchen wir für den Vorfaktor im Zähler auch noch eine Potenz von $X - z$, um auf f zu kommen, und durch die können wir in R nicht dividieren. Der Faktorring $R/(f)$ hat also die r Restklassen der Polynome $(X - z)^\ell$ mit $0 \leq \ell < r$ als Basis, und seine Dimension ist gleich der Vielfachheit r der Nullstelle z von f.

Für Polynome einer Veränderlichen ist das sicherlich eine sehr umständliche Art der Betrachtung; sie hat aber den Vorteil, daß sie sich auf Polynome in mehreren Veränderlichen verallgemeinern läßt.

Als erstes müssen wir klar definieren, was oben kurz als die „Einführung von Nennern“ bezeichnet wurde:

Definition: R sei ein (kommutativer) Ring.
a) Eine Teilmenge $S \subseteq R \setminus \{0\}$ heißt *multiplikativ abgeschlossen,* wenn sie mit je zwei Elementen $f, g \in S$ auch deren Produkt enthält.
b) Die *Lokalisierung* von R nach der multiplikativ abgeschlossenen Menge S ist die Menge aller Paare $(f, g) \in R \times S$ modulo der folgenden Äquivalenzrelation:

$$(f, g) \sim (r, s) \Longleftrightarrow \exists h \in R \setminus \{0\} : h(fs - rg) = 0 \,.$$

Die Gleichung $h(fs - rg) = 0$ ist natürlich äquivalent zu $h \cdot fs = h \cdot rg$; bis auf den Faktor h entspricht sie also dem aus der Bruchrechnung bekannten Überkreuzmultiplizieren. Falls R nullteilerfrei ist, können wir auf den Faktor h verzichten, denn dann folgt aus $h(fs - rg) = 0$, daß $fs - rg = 0$ sein muß. Die Ringe $A = k[X_1, \ldots, X_n]/I$, die wir hier betrachten, sind allerdings im Allgemeinen keine Integritätsbereiche.

Die Äquivalenzklasse des Paars (f, g) bezeichnen wir mit $\frac{f}{g}$, die Menge aller Äquivalenzklassen mit $S^{-1}R$. Addition und Multiplikation definieren wir nach den üblichen Regeln der Bruchrechnung als

$$\frac{f}{g} + \frac{r}{s} = \frac{fs + rg}{gs} \qquad \text{und} \qquad \frac{f}{g} \cdot \frac{r}{s} = \frac{fr}{gs} \,,$$

und wir müssen uns überlegen, daß diese Verknüpfungen wohldefiniert sind, daß das Ergebnis also nicht von der Wahl spezieller Repräsentanten (f, g) und (r, s) abhängt: Sind $(f, g) \sim (\tilde{f}, \tilde{g})$ und $(r, s) \sim (\tilde{r}, \tilde{s})$, so ist

$$\frac{\tilde{f}}{\tilde{g}} + \frac{\tilde{r}}{\tilde{s}} = \frac{\tilde{f}\tilde{s} + \tilde{r}\tilde{g}}{\tilde{g}\tilde{s}} \qquad \text{und} \qquad \frac{\tilde{f}}{\tilde{g}} \cdot \frac{\tilde{r}}{\tilde{s}} = \frac{\tilde{f}\tilde{r}}{\tilde{g}\tilde{s}} \,.$$

Nach Definition gibt es Elemente $h, t \in R \setminus \{0\}$, so daß $h \cdot f\tilde{g} = h \cdot \tilde{f}g$ und $t \cdot r\tilde{s} = t \cdot \tilde{r}s$ ist. Dann ist

$$\begin{aligned} ht \cdot (\tilde{f}\tilde{s} + \tilde{r}\tilde{g}) \cdot gs &= t \cdot (h \cdot \tilde{f}g)s\tilde{s} + h \cdot (t \cdot \tilde{r}s)g\tilde{g} \\ &= t \cdot (h \cdot f\tilde{g})s\tilde{s} + h \cdot (t \cdot r\tilde{s})g\tilde{g} \\ &= ht \cdot (fs + rg) \cdot \tilde{q}\tilde{s} \,, \end{aligned}$$

das heißt $\left(\tilde{f}\tilde{s} + \tilde{r}\tilde{g}, \tilde{g}\tilde{s}\right) \sim (fs + rg, gs)$. Entsprechend ist

$$ht \cdot \tilde{f}\tilde{r}gs = (h \cdot \tilde{f}g)(t \cdot \tilde{r}s) = (h \cdot f\tilde{g})(t \cdot r\tilde{s}) = ht \cdot fr\tilde{g}\tilde{s} \,,$$

das heißt $\left(\tilde{f}\tilde{r}, \tilde{g}\tilde{s}\right) \sim (fr, gs)$.

Die größte multiplikativ abgeschlossene Teilmenge eines Integritätsbereichs R ist $S = R \setminus \{0\}$; in diesem Fall ist $S^{-1}R$ ein Körper, den wir bereits als den *Quotientenkörper* Quot R von R kennen.

Wir interessieren uns für Ideale $I \lhd k[X_1, \ldots, X_n]$, für die $V_K(I)$ eine endliche Menge ist; dabei bezeichnet K wie üblich einen algebraisch abgeschlossenen Körper, der k enthält. Die Elemente der Vektorräume $A = k[X_1, \ldots, X_n]/I$ und $\bar{A} = K[X_1, \ldots, X_n]/\bar{I}$ können wir als Funktionen auf $V_K(I)$ mit Werten in K interpretieren. Da sich Funktionen addieren und multiplizieren lassen, sind auch A und $\bar{A}$ Ringe, deren Multiplikation offensichtlich mit der im Polynomring kompatibel ist. Für jedes $x \in V_K(I)$ ist die Menge

$$S_x = \left\{ f \in \bar{A} \;\middle|\; f(x) \neq 0 \right\}$$

multiplikativ abgeschlossen, denn die Funktionswerte liegen ja im (nullteilerfreien) Körper K. Diese Lokalisierungen wollen wir im folgenden genauer untersuchen.

Definition: *a)* $\bar{A}_x \underset{\text{def}}{=} S_x^{-1}\bar{A}$
b) Die *Vielfachheit* oder *Multiplizität* einer Nullstelle $x \in V_K(I)$ ist die Dimension von $\bar{A}_x$ als K-Vektorraum.

Wie wir oben gesehen haben, entspricht dies für Polynome einer Veränderlichen der gewohnten Vielfachheit; wir wollen uns überlegen, daß sich die Vielfachheiten der verschiedenen Elemente von $V_K(I)$ auch im Falle von Polynomen mehrerer Veränderlichen zu $\dim_K \bar{A}$ addieren.

Dazu benötigen wir noch einen Begriff aus der Linearen Algebra:

Definition: $V_1, \ldots, V_r$ seien Vektorräume über dem Körper k. Die direkte Summe

$$\bigoplus_{i=1}^{r} V_i = V_1 \oplus \cdots \oplus V_r$$

ist als Menge gleich dem kartesischen Produkt $V_1 \times \cdots V_r$ der Vektorräume; die Vektorraumaddition ist definiert durch

$$(v_1, \ldots, v_r) + (w_1, \ldots, w_r) = (v_1 + w_1, \ldots, v_r + w_r)\,,$$

und für einen Skalar $\lambda \in k$ setzen wir

$$\lambda(v_1, \ldots, v_n) = (\lambda v_1, \ldots, \lambda v_n)\,.$$

Die Vektorräume V_i können identifiziert werden mit jenen Untervektorräumen von $\oplus_{i=1}^{r} V_i$, in denen alle Komponenten außer eventuell der i-ten gleich dem Nullvektor sind.

Wenn alle Räume V_i endliche Dimensionen haben, ist die Dimension ihrer direkten Summe einfach die Summe dieser Dimensionen: Wählen wir in jedem der Vektorräume V_i eine Basis und fassen wir die V_i auf als Untervektorräume der direkten Summe, so ist die Vereinigung der Basen der V_i eine Basis des Summenraums. Insbesondere ist jeder endlichdimensionale k-Vektorraum mit einer Basis $b_1, \ldots, b_n$ isomorph zur direkten Summe der eindimensionalen Untervektorräume kb_i.

Satz: Ist $V_K(I)$ endlich, so ist $\bar{A} \cong \bigoplus_{x \in V_K(I)} \bar{A}_x$

Beweis: Im vierten Schritt des Beweises, daß die Elementanzahl von $V_K(I)$ höchstens gleich der Dimension von A ist, haben wir gesehen, daß es es ein homogenes lineares Polynom über K gibt, das für jeden Punkt aus $V_K(I)$ einen anderen Wert annimmt. Durch einen linearen Koordinatenwechsel können wir erreichen, daß X_1 diese Eigenschaft hat. Wir bezeichnen die X_1-Koordinate eines Punktes $x \in V_K(I)$ mit x_1 und betrachten die LAGRANGE-Polynome

$$s_x = \frac{\prod_{y \in V_K(I) \setminus \{x\}} (X_1 - y)}{\prod_{y \in V_K(I) \setminus \{x\}} (x_1 - y)} \in K[X_1]\,;$$

offensichtlich ist $s_x(x) = 1$ und $s_x(y) = 0$ für alle $y \neq x$ aus $V_K(I)$. Somit verschwindet das Produkt $s_x s_y$ zweier solcher Funktionen in jedem Punkt von $V_K(I)$; nach dem HILBERTschen Nullstellensatz liegt daher eine Potenz von $s_x s_y$ im Ideal $\bar{I}$. Bezeichnet r den größten Exponenten, den wir für eines der Produkte $s_x s_y$ brauchen, haben daher die Polynome $t_x = s_x^r$ die Eigenschaft, daß $t_x t_y$ für $x \neq y$ in $\bar{I}$ liegt, und $t_x(x) = 1$ ist.

Wir betrachten nun das Ideal $J \lhd K[X_1, \ldots, X_n]$, das von I und den sämtlichen t_x erzeugt wird. Es hat offensichtlich keine gemeinsame

Nullstelle, denn die gemeinsamen Nullstellen von $\bar{I}$ sind die $x \in V_K(I)$, und für jedes dieser x ist $t_x(x) = 1$. Nach der schwachen Form des HILBERTschen Nullstellensatzes enthält J daher die Eins; es gibt also Polynome $p_x \in K[X_1, \ldots, X_n]$ und ein Polynom $p \in \bar{I}$, so daß

$$\sum_{x \in V_K(I)} p_x t_x + p = 1$$

ist. Die Restklassen $e_x \in \bar{A}$ von $p_x t_x$ modulo $\bar{I}$ erfüllen die Gleichungen

1.) $\sum_{x \in V_K(I)} e_x = 1$
2.) $e_x e_y = 0$ für $x \neq y$ aus $V_K(I)$
3.) $e_x^2 = e_x$
4.) $e_x(x) = 1$

Die erste Gleichung ist klar, denn gehen wir in der Gleichung

$$\sum_{x \in V_K(I)} p_x t_t + p = 1$$

zu Restklassen modulo $\bar{I}$ über, wird p zur Klasse der Null und $p_x t_x$ zu e_x. Für $x \neq y$ ist $t_x t_y \in \bar{I}$; modulo $\bar{I}$ verschwindet das Produkt und damit auch $e_x e_y$, was die zweite Gleichung beweist.

Die dritte Gleichung folgt aus den ersten beiden: Nach der ersten ist $1 - e_x$ gleich der Summe der übrigen e_y, also ist

$$e_x - e_x^2 = e_x(1 - e_x) = e_x \sum_{y \neq x} e_y = \sum_{y \neq x} e_x e_y = 0\,.$$

Für die vierte Gleichung schließlich beachten wir, daß für $x \neq y$ mit $s_y(x)$ auch $t_y(x)$ verschwindet, d.h.

$$\sum_{y \in V_K(I)} p_y(x) t_y(x) + p(x) = p_x(x) t_x(x) = e_x(x) = 1\,.$$

Elemente e eines Rings R mit der Eigenschaft $e^2 = e$ bezeichnet man als *Idempotente;* sie haben die Eigenschaft, daß das Ideal $(e) = Re$ selbst ein Ring ist mit e als der Eins, denn $(ae)(be) = abe^2 = abe$ für alle $a, b \in R$.

Wir wollen uns als nächstes überlegen, daß der Ring $\bar{A}e_x$ isomorph ist zur Lokalisierung von $\bar{A}$ bei x; der Isomorphismus ist gegeben durch

$$\begin{cases} \bar{A}e_x \to \bar{A}_x \\ fe_x \mapsto \dfrac{f}{1} \end{cases} .$$

Zum Nachweis der Bijektivität konstruieren wir eine Umkehrabbildung $\bar{A}_x \to \bar{A}e_x$ wie folgt: Zu jedem $g \in \bar{A}$ mit $g(x) \neq 0$ setzen wir

$$\tilde{g} \underset{\text{def}}{=} \frac{g}{g(x)} - 1 \in \bar{A}_x, \quad \text{d.h.} \quad g = g(x)(1+\tilde{g}) .$$

Da $\tilde{g}(x)$ verschwindet und $e_x(y) = 0$ für alle $\neq x$ aus $V_K(I)$, verschwindet $\tilde{g}e_x$ auf ganz $V_K(I)$. Nach dem HILBERTschen Nullstellensatz gibt es somit eine Potenz eines Repräsentanten, die in $\bar{I}$ liegt, d.h. es gibt eine natürliche Zahl N, so daß $\big(\tilde{g}e_x\big)^N = \tilde{g}^N e_x$ die Null von $\bar{A}$ ist. Dann ist

$$(1+\tilde{g})e_x \cdot \big(1 - \tilde{g} + \tilde{g}^2 - \cdots + (-1)^{N-1}\tilde{g}^{N-1}\big)e_x = (1-\tilde{g}^N)e_x = e_x ;$$

im Ring $\bar{A}_x$ hat also $1+\tilde{g}$ ein Inverses und damit auch $ge_x = g(x)(1+\tilde{g})e_x$. Wir bilden daher den Bruch $f/g \in \bar{A}_x$ ab auf

$$f \cdot \frac{1}{g(x)} \cdot \big(1 - \tilde{g} + \tilde{g}^2 - \cdots + (-1)^{N-1}\tilde{g}^{N-1}\big)e_x \in \bar{A}_x ,$$

und mit Hilfe der gerade durchgeführten Rechnung folgt leicht, daß die beiden Abbildungen zueinander invers, also Isomorphismen sind.

Zum Beweis des Satzes fehlt nun nur noch, daß $\bar{A}$ die direkte Summe der Ringe $\bar{A}e_x$ ist; das ist klar, da die Summe der e_x gleich eins ist und $e_x e_y = 0$ für $x \neq y$.

∎

Im Falle, daß alle Nullstellen einfach sind, läßt sich dieser Satz einfacher formulieren: Die Elemente von $V_K(I)$ seien die Punkte

$$x^{(j)} = (x_1(j), \ldots, x_n^{(j)}) \quad \text{für } j = 1, \ldots, r ,$$

und für jedes j betrachten wir das maximale Ideal

$$\mathfrak{m}_j = (X_1 - x_1^{(j)}, \ldots, X_n - x_n^{(j)})$$

von $\bar{R} = K[X_1, \ldots, X_n]$. Für jedes j ist $\mathfrak{m}_j$ der Kern der Abbildung

$$\begin{cases} \bar{R} \to K \\ f \mapsto f(x^{(j)}) \end{cases} .$$

Da $\mathfrak{m}_j$ als maximales Ideal insbesondere prim ist, ist $S_j = \bar{R} \setminus \mathfrak{m}_j$ eine multiplikativ abgeschlossene Menge und wir können die entsprechend definierte Abbildung für $S_j^{-1}\bar{R}$ betrachten; ihr Kern ist das von $\mathfrak{m}_j$ in $S_j^{-1}\bar{R}$ erzeugte Ideal $\mathfrak{m}_j S_j^{(-1)}\bar{R}$, und nach dem Homomorphiesatz ist

$$\bar{R}/\mathfrak{m}_j = S_j^{-1}\bar{R}/\mathfrak{m}_j S_j^{-1}\bar{R} \cong K .$$

Da die Nullstelle $x^{(j)}$ einfach ist, muß auch $S_j^{-1}\bar{R}/IS_j^{-1}\bar{R}$ ein eindimensionaler Vektorraum, also isomorph zu K sein, und da $IS_j^{-1}\bar{R}$ in $\mathfrak{m}S_j^{-1}\bar{R}$ enthalten ist, müssen die beiden Ideale übereinstimmen. Die haben somit nach dem vorigen Satz einen Isomorphismus

$$\bar{A} = \bar{R}/\bar{I} \cong \bigoplus_{j=1}^{r} \bar{R}/\mathfrak{m}_j \cong K^r ,$$

der durch die Abbildung $f \mapsto \big(f(x^{(1)}), \ldots, f(x^{(r)})\big)$ gegeben ist. Um das Ideal $\bar{I}$ mit den $\mathfrak{m}_j$ in Verbindung zu bringen, brauchen wir die ringtheoretische Version des chinesischen Restesatzes:

Satz: $I_1, \ldots, I_r$ seien Ideale eines Rings R derart, daß $I_j + \bigcap_{\ell \neq j} I_\ell = R$ ist für alle j. Dann ist $R/\bigcap_{j=1}^{r} I_j \cong \bigoplus_{j=1}^{r} R/I_j$.

Beweis: Wir betrachten die Abbildung

$$\begin{cases} R \to \bigoplus_{j=1}^{r} R/I_j \\ f \mapsto (f \bmod I_1, \ldots, f \bmod I_r) \end{cases} .$$

Ihr Kern besteht aus allen Elementen von f, die modulo jedem I_j verschwinden, die also in allen I_j liegen. Somit ist der Kern der Durchschnitt der I_j und der Satz folgt aus dem Homomorphiesatz, wenn wir zeigen können, daß die Abbildung surjektiv ist.

Dies beweisen wir durch vollständige Induktion nach r: Für $r = 1$ ist das klar; sei also $r > 1$ und $(f_1, \ldots, f_r) \in R^r$. Nach Induktionsvoraussetzung gibt es ein Element $f^* \in R$, so daß $f^* \bmod I_j = f_j \bmod I_j$ für $1 \leq j < r$, und dieses Element f^* ist eindeutig modulo $I^* = \bigcap_{j=1}^{r-1} I_j$. Nach Voraussetzung ist $I^* + I_r = R$; es gibt also Elemente $g \in I^*$ und $h \in I_r$ mit $g + h = 1$. Modulo I^* ist $g = 0$ und $h = 1$, modulo I_r ist $g = 1$ und $h = 0$. Somit ist $f = hf^* + gf_r$ modulo I^* gleich f^* und modulo I_r gleich f_r, d.h. $f \bmod I_j = f_j \bmod I_j$ für alle j, so daß f ein Urbild von $(f_1, \ldots, f_r)$ ist.

■

Da die Ideale $\mathfrak{m}_j$ allesamt maximal sind, erfüllen sie die Voraussetzung dieses Satzes, d.h.

$$\bar{R} \Big/ \bigcap_{j=1}^{r} \mathfrak{m}_j \cong \bigoplus_{j=1}^{r} \bar{R}/\mathfrak{m}_j \,.$$

Andererseits ist die rechte Seite auch isomorph zu $\bar{A} = \bar{R}/\bar{I}$, und natürlich liegt $\bar{I}$ im Durchschnitt der $\mathfrak{m}_j$, denn dieser Durchschnitt besteht gerade aus den sämtlichen Polynomen, die in den Punkten $x^{(1)}, \ldots, x^{(r)}$ verschwinden. Da die Faktorringe übereinstimmen, muß somit

$$\bar{I} = \bigcap_{j=1}^{r} \mathfrak{m}_j$$

sein. Damit ist $\bar{I}$ ein Radikalideal, denn die $\mathfrak{m}_j$ sind als maximale Ideale insbesondere Primideale. Liegt für ein $f \in \bar{R}$ eine Potenz $f^m \in \bar{I}$, so liegt sie in jedem $\mathfrak{m}_j$, und da ein Produkt nur dann in einem Primideal liegen kann, wenn mindestens einer der Faktoren dort liegt, folgt $f \in \mathfrak{m}_j$ für alle j, also $f \in \bar{I}$.