Berechnung von Multiplizitäten und Radikalen

Bislang haben wir Multiplizitäten und Radikale hauptsächlich abstrakt betrachtet; heute soll es darum gehen, daraus Algorithmen zur ihrer konkreten Berechnung abzuleiten.

Wir betrachten weiterhin nur Gleichungssysteme mit endlicher Lösungsmenge über einem Körper k, den wir in einen (überabzählbaren) algebraisch abgeschlossenen Körper K einbetten. Dann ist der Restklassenring $A = k[X_1, \ldots, X_n]/I$ ein endlichdimensionaler k-Vektorraum, und die Standardmonome bezüglich irgendeiner GRÖBNER-Basis von I bilden eine Basis dieses Vektorraums und auch des entsprechenden Vektorraum $\bar{A}$ über K. Für jedes Element $f \in A$ betrachten wir die k-lineare Abbildung

$$L_f: \begin{cases} A \to A \\ g \mapsto fg \end{cases}$$

Satz von Stickelberger: L_f induziert für jedes $x \in V_K(I)$ eine lineare Abbildung $\bar{A}_x \to \bar{A}_x$. Diese hat $f(x)$ als ihren einzigen Eigenwert; dessen Vielfachheit ist also die Vielfachheit $\mu(x)$ der Nullstelle x.

Beweis: Wie wir aus der letzten Vorlesung wissen, enthält $\bar{A}$ für jedes $x \in V_K(I)$ ein idempotentes Element e_x derart, daß $\bar{A}_x \cong \bar{A}e_x$. Für $ge_x \in \bar{A}e_x$ ist $f \cdot ge_x = (fg)e_x \in \bar{A}e_x$, so daß auch $L_f(g)$ in $\bar{A}_x$ liegt.

$\big(f - f(x)\big)e_x$ verschwindet auf ganz $V_K(I)$, da e_x für alle $y \neq x$ aus der Lösungsmenge verschwindet. Nach der starken Form des HILBERTschen Nullstellensatzes liegt daher eine Potenz, etwa die m-te, eines Repräsentanten dieses Elements im Ideal $\bar{I}$. Im Restklassenring ist daher

$$\Big(\big(f - f(x)\big)e_x\Big)^m = \big(f - f(x)\big)^m e_x^m = \big(f - f(x)\big)^m e_x = 0\,,$$

d.h. $L^m_{f-f(x)}$ ist die Nullabbildung auf $\bar{A}e_x \cong \bar{A}_x$, so daß $f(x)$ der einzige Eigenwert ist. Seine (algebraische) Vielfachheit ist daher gleich der Dimension von $\bar{A}_x$, von der wir aus der letzten Vorlesung wissen, daß sie gleich der Vielfachheit der Nullstelle ist.

■

Aus diesem Satz können wir sofort Aussagen über die Spur, die Determinante und das charakteristische Polynom $\chi(L_f)$ von L_f ableiten:

Korollar: Für $f \in A$ gilt
a) $\operatorname{Sp} L_f = \sum_{x \in V_K(I)} \mu(x) f(x)$
b) $\det L_f = \prod_{x \in V_K(I)} f(X)^{\mu(x)}$
c) $\chi(L_f) = \det(L_f - T \cdot \mathrm{id}) = \prod_{x \in V_K(I)} \big(T - f(x)\big)^{\mu(x)}$

■

Vor allem die Spuren der linearen Abbildungen L_f werden für uns im folgenden wichtig sein. Man kann sie einfach berechnen, sobald man *irgendeine* GRÖBNER-Basis G des Ideal I kennt: Die Standardmonome bezüglich G bilden eine Vektorraumbasis $\omega_1, \ldots, \omega_r$ des Restklassenrings A über k. Zur Berechnung der Spur von L_f müssen wir die Elemente $L_f(\omega_i)$ in dieser Basis darstellen. Dazu müssen wir ein Polynom $F \in k[X_1, \ldots, X_n]$ wählen mit Restklasse f; meist wird f ohnehin in dieser Weise gegeben sein. Wenn wir einfach die Monome ω_i mit F multiplizieren, erhalten wir im Allgemeinen kein Polynom, das sich als Linearkombination der Monome ω_j schreiben läßt, denn wir multiplizieren ja im Polynomring, nicht im Restklassenring. Daher müssen wir die Produkte $F\omega_i$ mit dem Divisionsalgorithmus modulo der GRÖBNER-Basis G reduzieren; der Divisionsrest ist eine Linearkombination der Standardmonome. In dieser müssen wir den Koeffizienten von ω_i bestimmen, und diese Koeffizienten müssen wir über alle i aufsummieren.

Bei älteren Computeralgebrasystemen kann es hier zu Problemen kommen: Für den BUCHBERGER-Algorithmus ist es egal, ob wir den exakten Divisionsrest des S-Polynoms verwenden oder ein skalares Vielfaches davon. Da der exakte Rest im Fall $k = \mathbb{Q}$ oft große Nenner hat, wird die weitere Berechnung effizienter, wenn man ihn mit einer ganzen Zahl multipliziert derart, daß das Ergebnis nur noch ganzzahlige Koeffizienten hat. Auf diese Weise arbeitet beispielsweise das Kommando `poly_normal_form` für den Divisionsalgorithmus in Maxima (und wahrscheinlich auch in vielen anderen Versionen von macsyma). Für

Spurberechnungen ist ein solcher modifizierter Divisionsrest natürlich unbrauchbar; wir brauchen den exakten Rest.

Zu dessen Bestimmung, auch mit einem solchen System, gibt es mehrere Möglichkeiten. Am einfachsten ist die Situation, wenn man bereits ein Element von $V_K(I)$ kennt, für das das zu reduzierende Polynom P nicht verschwindet. Da die Differenz von P und dem (korrekten) Divisionsrest im Ideal I liegt, nehmen beide in diesem Punkt denselben Wert an. Wertet man daher sowohl P als auch das vom Computeralgebrasystem berechnete Vielfache des Rests in diesem Punkt aus, findet man den Multiplikator, mit dem man auf den korrekten Rest kommt.

Leider kennt man im Voraus nur selten ein Element von $V_K(I)$. Der folgende Ansatz funktioniert allgemein: Angenommen, wie kennen für ein Polynom P ein Vielfaches Q des korrekten Divisionsrests bezüglich der GRÖBNER-Basis G. Im Falle $Q = 0$ ist auch der korrekte Divisionsrest gleich Null. Andernfalls gibt es genau eine Zahl $\lambda \in k$, für die $P - \lambda Q$ in I liegt: Gäbe es nämlich ein $\mu \neq \lambda$ mit $P - \mu Q \in I$, so läge auch die Differenz $(\mu - \lambda)Q$ und damit auch Q in I, d.h. auch P müßte in I liegen, so daß wir bei der Division durch die Elemente einer GRÖBNER-Basis von I Rest Null erhalten müßten.

Wir betrachten nun λ als eine neue Variable und rechnen im Polynomring über $k(\lambda)$. Auch hier können wir den BUCHBERGER-Algorithmus durchführen und erhalten als Rest ein Polynom mit Koeffizienten aus $k(\lambda)$. Wir wir uns bereits überlegt haben, gibt es genau ein $\lambda \in k$, für das alle diese Koeffizienten verschwinden, und für dieses ist λQ der korrekte Divisionsrest.

Wir betrachten nun für jedes $h \in A$ die Bilinearform

$$\mathrm{SpB}_h\colon \begin{cases} A \times A \to K \\ (f, g) \mapsto \mathrm{Sp}\, L_{fgh} \end{cases}$$

und die dazugehörige quadratische Form, die sogenannte HERMITE-Form

$$Q_h\colon \begin{cases} A \to K \\ f \mapsto \mathrm{Sp}\, L_{f^2 h} \end{cases}$$

Im Falle $h = 1$ werden wir den Index h in beiden Fällen meist weglassen.

Für die Beweise der folgenden Sätze benötigen wir die VANDERMONDEsche Determinante. Für Leser, die sie nicht aus der Linearen Algebra oder Numerik kennen, sei sie hier kurz definiert und berechnet.

Definition: Für $a_1, \ldots, a_n \in K$ ist die VANDERMONDEsche Determinante

$$V(a_1, \ldots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \ldots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \ldots & a_n^{n-1} \end{vmatrix}.$$

Zur Berechnung dieser Determinante nach dem LAPLACEschen Entwicklungssatz subtrahieren wir zunächst die erste Zeile von jeder der folgenden; da sich der Wert der Determinanten dadurch nicht ändert, ist

$$V(a_1, \ldots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^{n-1} \\ 0 & a_2 - a_1 & a_2^2 - a_1^2 & \ldots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_n - a_1 & a_n^2 - a_1^2 & \ldots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}.$$

Wenn wir hier nach der ersten Spalte entwickeln, muß nur eine einzige $(n-1) \times (n-1)$-Determinante berücksichtigt werden, alle anderen haben den Vorfaktor Null. Also ist

$$V(a_1, \ldots, a_n) = \begin{vmatrix} a_2 - a_1 & a_2^2 - a_1^2 & \ldots & a_2^{n-1} - a_1^{n-1} \\ a_3 - a_1 & a_3^2 - a_1^2 & \ldots & a_3^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n - a_1 & a_n^2 - a_1^2 & \ldots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}$$

gleich jeder Determinanten, die durch Streichung der ersten Spalte und der ersten Zeile entsteht.

Hier können wir in jeder Zeile die jeweils vorne stehende Differenz ausklammern, denn genau wie

$$x^k - 1 = (x-1)(x^{k-1} + x^{k-2} + \cdots + x + 1)$$

durch $(x-1)$ teilbar ist, ist auch

$$a_i^k - a_1^k = (a_i - a_1)(a_i^{k-1} + a_i^{k-2}a_1 + a_i^{k-3}a_1^2 + \cdots + a_i a_1^{k-2} + a_1^{k-1})$$

durch $(a_i - a_1)$ teilbar; den Quotienten schreiben wir kurz als $q_{i,k-1}$:

$$q_{i,k-1} \underset{\text{def}}{=} a_i^{k-1} + a_i^{k-2}a_1 + a_i^{k-3}a_1^2 + \cdots + a_i a_1^{k-2} + a_1^{k-1} .$$

Wegen der Linearität der Determinante können wir jeden Faktor, den wir aus einer Zeile (oder Spalte) ausklammern, vor die Determinante ziehen und erhalten für $V(a_1, \ldots, a_n)$ somit den Wert

$$(a_2 - a_1)(a_3 - a_1)\cdots(a_n - a_1) \begin{vmatrix} 1 & q_{21} & q_{22} & \cdots & q_{2,n-2} \\ 1 & q_{31} & q_{32} & \cdots & q_{3,n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & q_{n1} & q_{n2} & \cdots & q_{n,n-2} \end{vmatrix} .$$

Die Nützlichkeit dieser Formel steht und fällt damit, daß wir die q_{ij} gut miteinander in Verbindung bringen können. Für verschiedene Indizes i haben die entsprechenden Ausdrücke offensichtlich wenig miteinander zu tun; sie enthalten nicht einmal dieselben Variablen. Schreiben wir allerdings

$$\begin{aligned} q_{ij} &= a_i^j + a_i^{j-1}a_1 + a_i^{j-2}a_1^2 + \cdots a_i a_1^{j-1} + a_1^j \\ &= a_i^j + a_1(a_i^{j-1} + a_i^{j-2}a_1 + \cdots a_i a_1^{j-2} + a_1^{j-1}) , \end{aligned}$$

so sehen wir, daß

$$q_{ij} = a_i^j + a_1 q_{i,j-1} \quad \text{oder} \quad q_{ij} - a_1 q_{i,j-1} = a_i^j$$

ist. Subtrahieren wir also zuerst a_1 mal die vorletzte Spalte von der letzten, so werden die Einträge der letzten Spalte zu a_i^{n-2}. Entsprechend subtrahieren wir a_1 mal die $(n-2)$-te Zeile von der $(n-1)$-ten und erhalten lauter Einträge a_i^{n-3} und so weiter, bis schließlich die Subtraktion des a_1-fachen der ersten Spalte von der zweiten die Einträge der letzteren zu

$$q_{i1} - a_1 = (a_i + a_1) - a_1 = a_i$$

macht. Somit ist $V(a_1, \ldots . a_n)$ gleich

$$(a_2 - a_1)(a_3 - a_1)\cdots(a_n - a_1) \begin{vmatrix} 1 & a_2 & a_2^2 & \ldots & a_2^{n-2} \\ 1 & a_3 & a_3^2 & \ldots & a_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \ldots & a_n^{n-2} \end{vmatrix} .$$

Die Determinante rechts ist offensichtlich wieder eine VANDERMONDEsche Determinante, allerdings mit um eins verminderter Zeilen- und Spaltenzahl und mit einer Variablen weniger.

Damit haben wir die Rekursionsformel

$$V(a_1,\ldots,a_n) = (a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)\,V(a_2,\ldots,a_n)\,,$$

die es erlaubt, die Berechnung von $V(a_1,\ldots,a_n)$ auf eine einzige VANDERMONDEsche Determinante der Größe $(n-1)\times(n-1)$ zurückzuführen.

Zur vollständigen Berechnung von $V(a_1,\ldots,a_n)$ fehlt uns jetzt nur noch ein Induktionsanfang. Direktes Nachrechnen zeigt sofort, daß

$$V(a_n) = 1 \qquad \text{und} \qquad V(a_{n-1},a_n) = \begin{vmatrix} 1 & a_{n-1} \\ 1 & a_n \end{vmatrix} = a_n - a_{n-1}$$

ist. Daher folgt induktiv

$$V(a_1,\ldots,a_n) = \prod_{j<i}(a_i-a_j)\,.$$

Satz: Ein Polynom $f \in k[X_1,\ldots,X_n]$ liegt genau dann im Radikal von I, wenn $f \bmod I$ in

$$\operatorname{Kern} Q \underset{\text{def}}{=} \left\{ f \in A \;\middle|\; \operatorname{SpB}(f,g) = 0 \quad \forall g \in A \right\}$$

liegt.

Beweis: Für $f \in \sqrt{I}$ verschwindet f für alle $x \in V_K(I)$; nach obigem Korollar ist daher für alle $g \in A$

$$\operatorname{SpB}(f,g) = \operatorname{Sp} L_{fg} = \sum_{x\in V_K(I)} \mu(x) f(x) g(x) = 0\,.$$

Umgekehrt sei $\operatorname{SpB}(f,g) = 0$ für alle $g \in A$, d.h.

$$\sum_{x\in V_K(I)} \mu(x) f(x) g(x) = 0 \quad \forall g \in A\,.$$

Sei $V_K(I) = \{x^{(1)},\ldots,x^{(s)}\}$. Wie wir aus §3 wissen, gibt es eine separierende Linearform $u \in A$, d.h. eine Funktion, die für jedes $x^{(j)}$

einen anderen Wert annimmt. Wir betrachten dazu das Produkt der Matrix

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ u(x^{(1)}) & u(x^{(2)}) & \dots & u(x^{(s)}) \\ u^2(x^{(1)}) & u^2(x^{(2)}) & \dots & u^2(x^{(s)}) \\ \vdots & \vdots & \ddots & \vdots \\ u^{s-1}(x^{(1)}) & u^{s-q}(x^{(2)}) & \dots & u^{s-1}(x^{(s)}) \end{pmatrix}$$

mit dem Vektor

$$v = \begin{pmatrix} \mu(x^{(1)})f(x^{(1)}) \\ \mu(x^{(2)})f(x^{(2)}) \\ \mu(x^{(s)})f(x^{(s)}) \end{pmatrix} .$$

Der ℓ-te Eintrag von Mv ist

$$\sum_{j=1}^{s} u(x^{(j)})^{\ell-1} \cdot \mu(x^{(j)})f(x^{(j)}) = \mathrm{SpB}(u^{\ell-1}, f) = 0 ,$$

da $f \in \operatorname{Kern} Q$.

Die Determinante von M (oder genauer seiner Transponierten) ist eine VANDERMONDEsche Determinante. Da u separierend ist, sind alle $u(x^{(j)})$ verschieden, so daß $\det M$ nach obiger Formel nicht verschwindet. Das homogene lineare Gleichungssystem $Mv = 0$ hat daher nur die triviale Lösung $v = 0$. Damit verschwinden alle $\mu(x)f(x)$, also alle $f(x)$, d.h. f verschwindet auf $V_K(I)$ und muß daher im Radikal von I liegen. ∎

Satz: Für alle $h \in A$ ist $\operatorname{Rang} Q_h = \#\{x \in V_K(I) \mid h(x) \neq 0\}$. Insbesondere ist $\operatorname{Rang} Q$ gleich der Anzahl der Lösungen in $V_K(I)$.

Beweis: Wie im vorigen Beweis sei $u \in A$ eine separierende Linearform, und $V_K(I) = \{x^{(1)}, \dots, x^{(s)}\}$. Dann sind $1, u, \dots u^{s-1}$ linear unabhängig, denn ist

$$\lambda_0 + \lambda_1 u + \cdots + \lambda_s u^{s-1} = 0 ,$$

so ist insbesondere

$$\lambda_0 + \lambda_1 u(x^{(j)}) + \cdots + \lambda_s u^{s-1}(x^{(j)}) = 0$$

für $j = 1, \dots, s$. Dies ist ein homogenes lineares Gleichungssystem für $\lambda_0, \dots, \lambda_{s-1}$, dessen Matrix eine VANDERMONDEsche Determinante

hat, die nicht verschwindet, da u separierend ist. Somit gibt es nur die triviale Lösung $\lambda_0 = \cdots = \lambda_{s-1} = 0$.

Falls $s < r = \dim_k A$ ist, können wir daher das System der Vektoren $\omega_1 = 1$, $\omega_2 = u, \ldots, \omega_s = u^{s-1}$ ergänzen zu einer Basis $\omega_1, \ldots, \omega_r$ von A.

Für jedes Element $g = g_1\omega_1 + \cdots + g_r\omega_r \in A$ (mit $g_\ell \in k$) ist nach obigem Korollar

$$Q_h(g) = \operatorname{Sp} L_{g^2h} = \sum_{j=1}^{s} \mu(x^{(j)}) \left(\sum_{\ell=1}^{r} g_\ell \omega_\ell(x^{(j)}) \right)^2 \cdot h(x^{(j)}) \,.$$

$v \in K^s$ sei der Vektor mit Komponenten $v_j = \sum_{\ell=1}^{r} g_\ell \omega_\ell(x^{(j)})$, und Δ sei die Diagonalmatrix mit Einträgen $\mu(x^{(j)})h(x^{(j)})$. Dann ist

$$Q_h(g) = v^T \Delta v \,.$$

Mit

$$\Gamma = \begin{pmatrix} 1 & \omega_1(x^{(1)}) & \ldots & \omega_s(x^{(1)}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_1(x^{(s)}) & \ldots & \omega_s(x^{(s)}) \end{pmatrix} \in K^{s\times r} \quad \text{ist} \quad v = \Gamma \begin{pmatrix} g_1 \\ \vdots \\ g_r \end{pmatrix}$$

und damit

$$Q_h(q) = (g_1, \ldots, g_r)\Gamma^T \Delta \Gamma \begin{pmatrix} g_1 \\ \vdots \\ g_r \end{pmatrix} \,.$$

Γ hat Rang s, da die ersten s Spalten eine VANDERMONDEsche Determinante haben, und der Rang von Δ ist gleich der Anzahl der $x \in V_K(I)$ mit $h(x) \neq 0$, also eventuell kleiner als s. Somit ist der Rang von $Q_h = \Gamma^T\Delta\Gamma$ gleich letzterer Anzahl, wie behauptet. ∎

Ist $f = f_1\omega_1 + \cdots + f_r\omega_r$ mit $j_i \in k$ ein weiteres Element von A, so ist

$$fgh = \sum_{i=1}^{r} \sum_{j=1}^{r} h f_i g_j \omega_i \omega_j \,;$$

da für zwei Matrizen $A, B \in k^{r\times r}$ und zwei Skalare λ, μ aus k gilt $\mathrm{Sp}(\lambda A + \mu B) = \lambda \,\mathrm{Sp}\, A + \mu \,\mathrm{Sp}\, B$, ist also

$$\mathrm{Sp}\, L_{fgh} = h \sum_{i=1}^{r} \sum_{j=1}^{r} f_i g_j \,\mathrm{Sp}(\omega_i \omega_j)\,.$$

Mit der Matrix

$$\mathrm{SpM} \underset{\text{def}}{=} \begin{pmatrix} \mathrm{Sp}(\omega_1\omega_1) & \dots & \mathrm{Sp}(\omega_1\omega_r) \\ \vdots & \ddots & \vdots \\ \mathrm{Sp}(\omega_r\omega_1) & \dots & \mathrm{Sp}(\omega_1\omega_r) \end{pmatrix}$$

ist daher

$$\mathrm{Sp}(L_{fgh}) = (f_1, \dots, f_r)\,\mathrm{SpM} \begin{pmatrix} g_1 \\ \vdots \\ g_r \end{pmatrix}\,.$$

Ein Polynom $f \in k[X_1, \dots, X_n]$ liegt nach dem vorletzten Satz genau dann in $\sqrt{I}$, wenn seine Restklasse $f + I$ im Kern von Q liegt, wenn also für $f + I = f_1\omega_1 + \cdots + f_r\omega_r$ gilt

$$(f_1, \dots, f_r)\,\mathrm{SpM}\, g = 0 \quad \forall g \in k^r\,.$$

Dies ist genau dann der Fall, wenn

$$(f_1, \dots, f_r)\,\mathrm{SpM} = (0, \dots, 0)$$

ist. Transposition auf beiden Seiten macht daraus das homogene lineare Gleichungssystem

$$\mathrm{SpM}(f_1, \dots, f_r)^T = 0\,,$$

das wir explizit aufstellen und lösen können. Die Vektoren $p^{(1)}, \dots, p^{(t)}$ aus k^r seien eine Basis des Lösungsraums. Dann sind die Polynome $P^{(j)} = \sum_{i=1}^{r} p_i^{(j)} \omega_i$ nach dem vorletzten Satz Elemente von $\sqrt{I}$, und $\sqrt{I}$ wird erzeugt von diesen Polynomen und den Erzeugenden von I.