Gleichungssysteme mit endlicher Lösungsmenge

Auch hier gehen wir wieder aus von einem beliebigen Körper k sowie einem algebraisch abgeschlossenen Erweiterungskörper K mit überabzählbar vielen Elementen. Letztere Bedingung ist nur notwendig, weil wir sie im Beweis des HILBERTschen Nullstellensatzes verwendet haben; wie bereits dort erwähnt, gibt es auch Beweise für den Fall, daß K ein beliebiger algebraisch abgeschlossener Körper ist, so daß alle Sätze dieses Paragraphen tatsächlich auch ohne die Voraussetzung der Überabzählbarkeit von K gelten.

Wie wir gesehen haben, kann die Lösungsmenge eines nichtlinearen Gleichungssystems im Allgemeinen nur dann explizit bestimmt werden, wenn sie endlich ist. Daher ist es nützlich, ein Kriterium zu haben, mit dem man dem System möglichst einfach ansehen kann, wann das der Fall ist. Abstrakt algebraisch haben wir so etwas:

Satz: I sei ein Ideal im Polynomring $k[X_1, \ldots, X_n]$ über dem Körper k, und K sei ein überabzählbarer algebraisch abgeschlossener Körper, in dem k enthalten sei. Dann gilt: $V_K(I)$ ist genau dann endlich, wenn der Faktorring $A = k[X_1, \ldots, X_n]/I$ ein endlichdimensionaler k-Vektorraum ist. In diesem Fall ist die Dimension von A eine obere Schranke für die Elementanzahl von $V_K(I)$.

Den recht umfangreichen *Beweis* führen wir in mehreren Schritten:

1. Schritt: Wenn der Vektorraum A endliche Dimension hat, ist $V_K(I)$ endlich.

Bezeichnet nämlich d die Dimension von A, so sind für jedes i die X_i-Potenzen $1, X_i, \ldots, X_i^d$ linear abhängig; es gibt also ein nichtverschwindendes Polynom aus $k[X_i]$ vom Grad höchstens d, das modulo I zum Nullpolynom wird und somit in I liegt. Für jeden Punkt aus $V_K(I)$ muß daher die i-te Koordinate eine der höchstens d Nullstellen dieses Polynoms sein. Damit kann die i-te Koordinate nur endlich viele Werte annehmen, und da dies für alle i gilt, ist $V_K(I)$ endlich.

2. Schritt: $\bar{I}$ sei das von I in $K[X_1, \ldots, X_n]$ erzeugte Ideal. Wenn $V_K(I)$ endlich ist, hat der K-Vektorraum $\bar{A} = K[X_1, \ldots, X_n]/\bar{I}$ endliche Dimension.

Besteht $V_K(I)$ nur aus endlich vielen Punkten, so nimmt jede der Koordinatenfunktionen $X_1, \ldots, X_n$ auf $V_K(I)$ nur endlich viele Werte an. Es gibt daher für jedes i ein Polynom aus $K[X_i]$, das auf ganz $V_K(I)$ verschwindet. Nach dem HILBERTschen Nullstellensatz muß eine Potenz dieses Polynoms in $\bar{I}$ liegen. Daher gibt es für jedes i ein Polynom aus $K[X_i]$ in $\bar{I}$; sein Grad sei d_i. Für jedes $e \geq d_i$ ist dann X_i^e modulo $\bar{I}$ linear abhängig von den X_i-Potenzen $1, X_i, \ldots, X_i^{d_i-1}$. Somit läßt sich jedes Monom aus $K[X_1, \ldots, X_n]$ modulo $\bar{I}$ als K-Linearkombination von Monomen ausdrücken, bei denen jede Variable X_i höchstens mit Exponent $d_i - 1$ auftritt. Da es nur endlich viele solche Monome gibt, ist $K[X_1, \ldots, X_n]/\bar{I}$ ein endlichdimensionaler K-Vektorraum.

3. Schritt: A ist genau dann endlichdimensional, wenn $\bar{A}$ endlichdimensional ist; in diesem Fall haben beide dieselbe Dimension.

Ist A endlichdimensional, so wählen wir eine Basis $\{b_1, \ldots, b_r\}$ und zu jedem Basiselement b_i ein Polynom $B_i \in k[X_1, \ldots, X_n]$, das modulo I gleich b_i ist. Zusammen mit einer Basis von I als k-Vektorraum bilden die B_i dann eine k-Vektorraumbasis von $k[X_1, \ldots, X_n]$. Über K wird die Basis von I zu einer K-Vektorraumbasis von $\bar{I}$, da sich jedes Element von $\bar{I}$ als eine K-Linearkombination von Elementen aus I schreiben läßt. Zusammen mit den B_i, die wir auch als Elemente von $K[X_1, \ldots, X_n]$ auffassen können, erhalten wir sowohl über k als auch über K eine Basis des ganzen jeweiligen Polynomrings, und damit ist klar, daß die Restklassen der B_i modulo $\bar{I}$ den Faktorring $\bar{A}$ erzeugen. Somit ist dieser als K-Vektorraum endlichdimensional.

Die Gleichheit von $\dim_k A$ und $\dim_K \bar{A}$ folgt, falls wir zeigen können, daß die Restklassen der B_i modulo $\bar{I}$ linear unabhängig sind.

Dazu zeigen wir die folgende, etwas allgemeinere Aussage: Sind $B_1, \ldots, B_r$ Polynome aus $k[X_1, \ldots, X_n]$ mit Restklassen $b_1, \ldots, b_r$ modulo I und Restklassen $\bar{b}_1, \ldots, \bar{b}_r$ modulo $\bar{I}$, so sind die b_i genau dann linear abhängig, wenn es die $\bar{b}_i$ sind.

Die eine Richtung ist einfach: Falls die b_i linear abhängig sind, gibt es Skalare $\lambda_i \in k$, die nicht alle verschwinden, so daß $\lambda_1 b_1 + \cdots + \lambda_r b_r$ der Nullvektor aus A ist. $\lambda_1 B_1 + \cdots + \lambda_r B_r$ liegt daher in I, also erst recht in $\bar{I}$, so daß auch $\lambda_1 \bar{b}_1 + \cdots + \lambda_r \bar{b}_r$ der Nullvektor aus $\bar{A}$ ist.

Wenn die $\bar{b}_i$ linear abhängig sind, gibt es $\lambda_i \in K$, so daß $\lambda_1\bar{b}_1+\cdots+\lambda_r\bar{b}_r$ der Nullvektor aus $\bar{A}$ ist, d.h. $\lambda_1 B_1 + \cdots + \lambda_r B_r$ liegt in $\bar{I}$. Da die λ_i nicht in k liegen müssen, nützt und das noch nichts, um etwas über die b_i auszusagen.

Um trotzdem deren lineare Abhängigkeit zu beweisen, wählen wir ein endliches Erzeugendensystem $f_1, \ldots, f_m$ des Ideals I. Wir wissen dann, daß es Polynome $g_1, \ldots, g_m$ aus $K[X_1, \ldots, X_n]$ gibt mit

$$\lambda_1 B_1 + \cdots + \lambda_r B_r = g_1 f_1 + \cdots + g_m f_m \,.$$

Die Polynome g_j sind K-Linearkombinationen von Monomen $M_{j\ell}$ in den Variablen X_i. Die obige Gleichung ist also äquivalent zu einer Gleichung der Form

$$\lambda_1 B_1 + \cdots + \lambda_r B_r - \sum_{j=1}^{m}\sum_{\ell=1}^{r_j} \mu_{j\ell} M_{j\ell} f_j = 0$$

mit Elementen $\mu_{j\ell} \in K$, die von den g_j abhängen. Sortieren wir diese Gleichung nach Monomen, können wir dies so interpretieren, daß ein (recht großes) lineares Gleichungssystem in den Variablen λ_i und $\mu_{j\ell}$ eine nichttriviale Lösung hat. Da die B_i und die f_j Polynome mit Koeffizienten aus k sind, ist dies ein homogenes lineares Gleichungssystem mit Koeffizienten aus k. Seine Lösungsmenge über k ist ein k-Vektorraum, für den uns der GAUSS-Algorithmus eine Basis liefert. Da der GAUSS-Algorithmus nirgends aus dem Körper hinausführt, in dem die Koeffizienten liegen, ist dies auch eine Basis des Lösungsraums über K; die beiden Vektorräume haben also dieselbe Dimension. Da wir wissen, daß es über K eine nichttriviale Lösung gibt, muß es daher auch über k eine geben.

Es gibt somit Elemente $\lambda'_i \in k$ und $\mu'_{j\ell} \in k$, die das Gleichungssystem lösen. Damit ist dann

$$\lambda'_1 B_1 + \cdots + \lambda'_r B_r = g'_1 f_1 + \cdots + g'_m f_m$$

mit Polynomen $g'_j \in k[X_1, \ldots, X_n]$, die linke Seite liegt also im Ideal I. Somit ist $\lambda'_1 b_1 + \cdots + \lambda'_r b_r$ der Nullvektor in A. Die λ'_i können nicht allesamt verschwinden, denn ansonsten müßte mindestens ein $\mu_{j\ell} \neq 0$ sein, Null wäre also gleich einer nichttrivialen Linearkombination von Monomen, was absurd ist. Also sind auch die b_i linear abhängig.

Bleibt noch zu zeigen, daß A endlichdimensional ist, wenn $\bar{A}$ endlichdimensional ist. Das folgt sofort aus der gerade gezeigten Äquivalenz der linearen Abhängigkeit über k und über K: Hat $\bar{A}$ die endliche Dimension d, so ist jede Teilmenge von $\bar{A}$ mit mehr als d Elementen linear abhängig. Damit ist, wie wir gerade gesehen haben, auch jede Teilmenge von mehr als d Elementen aus A linear abhängig über k, also ist A endlichdimensional.

Im nächsten Schritt wollen wir das Zählen der Lösungen zurückführen auf das Zählen von Nullstellen eines Polynoms einer Veränderlichen.

Definition: Ein Polynom $u \in K[X_1, \ldots, X_n]$ heißt *separierend,* wenn es für keine zwei Elemente von $V_K(I)$ denselben Wert annimmt.

4. Schritt: Falls $V_K(I)$ endlich ist, gibt es ein separierendes homogenes lineares Polynom $u = c_1X_1 + \cdots c_nX_n$. Wir können dabei für u eines der speziellen Polynome

$$u_a = X_1 + aX_2 + a^2X_3 + \cdots + a^{n-1}X_n$$

wählen, wobei a in einer beliebig vorgebbaren Teilmenge von K mit mehr als $(n-1)\binom{s}{2} = \frac{1}{2}s(s-1)(n-1)$ Elementen liegt.

Für zwei verschiedene Punkte $z = (z_1, \ldots, z_n)$ und $w = (w_1, \ldots, w_n)$ aus $V_K(I)$ ist $u_a(z) = u_a(w)$ genau dann, wenn

$$(z_1 - w_1) + (z_2 - w_2)a + (z_3 - w_3)a^2 + \cdots + (z_n - w_n)a^{n-1}$$

verschwindet. Die Koordinaten z_i, w_i von z und w sind Elemente von K; die Elemente $a \in K$, für die $u_a(z) = u_a(w)$ ist, sind daher die Nullstellen eines Polynoms in einer Veränderlichen über K vom Grad höchstens $n-1$. Daher gibt es höchstens $n-1$ Werte $a \in K$, für die $u_a(z) = u_a(w)$ ist. Ist $s = \#V_K(I)$ endlich, so gibt es $\binom{s}{2}$ Paare (z, w) aus voneinander verschiedenen Elementen von $V_K(I)$; somit gibt es höchstens $(n-1)\binom{s}{2}$ Elemente $a \in K$, für die $u_a(z)$ und $u_a(w)$ für *irgendwelche* voneinander verschiedene Elemente von $V_K(I)$ übereinstimmen. Da K als algebraisch abgeschlossener Körper unendlich ist, folgt sogar ohne die Voraussetzung der Überabzählbarkeit, daß es Elemente $a \in K$ gibt, für die $u_a(z) \neq u_a(w)$ ist für jedes Paar (z, w) aus zwei verschiedenen Elementen von $V_K(I)$, d.h. u_a ist separierend.

(Falls bereits k unendlich ist, etwa $k = \mathbb{Q}$, können wir sogar entsprechende Werte $a \in k$ finden. Für Körper, die $\mathbb{Q}$ enthalten, gibt es sogar ganzzahlige a mit dieser Eigenschaft, konkret sogar ein $a \in \mathbb{Z}$, das beispielsweise der Ungleichung $0 \leq a \leq (n-1)\binom{s}{2}$ genügt.)

5. Schritt: Die Elementanzahl s von $V_K(I)$ ist höchstens gleich der Dimension von A.

Da wir im 3. Schritt gesehen haben, daß $\dim_k A = \dim_K \bar{A}$ ist, können wir auch mit dieser Dimension argumentieren. Aus dem 4. Schritt wissen wir, daß es ein Polynom $u \in K[X_1, \ldots, X_n]$ gibt, das für jedes Element von $V_K(I)$ einen anderen Wert annimmt. Wir ersetzen u durch seine Restklasse $\tilde{u}$ modulo $\bar{I}$ in $\bar{A}$ und wollen uns überlegen, daß die Elemente $1, \tilde{u}, \ldots, \tilde{u}^{s-1} \in \bar{A}$ linear unabhängig sind: Angenommen, es gibt eine Relation der Form $\sum_{\ell=0}^{s-1} \lambda_\ell \tilde{u}^\ell = 0$ mit $\lambda_\ell \in K$. Das Polynom $\sum_{\ell=0}^{s-1} \lambda_\ell u^\ell \in K[X_1, \ldots, X_n]$ liegt dann in $\bar{I}$, verschwindet also für jedes der s Elemente von $V_K(I)$. Da u für jedes dieser Elemente einen anderen Wert annimmt, hat das Polynom $\sum_{\ell=0}^{r-1} \lambda_\ell U^\ell \in k[U]$ einerseits mindestens s verschiedene Nullstellen in K, andererseits ist sein Grad kleiner als s. Das ist nur für das Nullpolynom möglich; somit verschwinden alle Koeffizienten λ_ℓ, was die behauptete lineare Unabhängigkeit beweist. Damit enthält $\bar{A}$ mindestens s linear unabhängige Elemente, d.h. $r = \dim_K \bar{A} \geq s = \#V_K(I)$. Damit ist die Behauptung und auch der gesamte Satz bewiesen.

■

Betrachten wir als Beispiel das von $f = X^2 + Y^2 - 1$ und $g = X - Y$ erzeugte Ideal $I \lhd \mathbb{Q}[X, Y]$. Seine Nullstellenmenge ist, geometrisch gesehen, der Schnitt des Einheitskreises mit der ersten Winkelhalbierenden, besteht also aus den beiden Punkten $\left(\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}\right)$ und $\left(-\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2}\right)$.

Der Polynomring $\mathbb{Q}[X, Y]$ hat als $\mathbb{Q}$-Vektorraum eine Basis bestehend aus allen Monomen $X^a Y^b$ mit $a, b \in \mathbb{N}_0$. Modulo I sind X und Y äquivalent, da $g = X - Y$ in I liegt, und damit ist auch $X^a Y^b$ äquivalent zu X^{a+b} für alle $a, b \in \mathbb{N}_0$. Außerdem ist $2X^2$ äquivalent zu $X^2 + Y^2$, und das wiederum ist wegen f äquivalent zu 1, d.h. $X^2 \sim \frac{1}{2}$. Daher ist jedes Monom äquivalent entweder zu einer Konstanten (falls $a + b$ gerade

ist) oder einem skalaren Vielfachen von X. Da I kein Polynom der Form $\lambda X + \mu$ enthält, sind X und 1 modulo I linear unabhängig; somit bilden ihre Restklassen eine Basis des Vektorraums $\mathbb{Q}[X,Y]/I$. Damit ist dieser zweidimensional; das Gleichungssystem $f(x,y) = g(x,y) = 0$ hat also nach dem gerade bewiesenen Satz höchstens zwei Lösungen, was wir hier natürlich einfacher direkt gesehen haben.

Ersetzen wir in diesem Beispiel g durch $X^2 - Y^2 = (X+Y)(X-Y)$, so schneiden wir den Kreis mit beiden Winkelhalbierenden und haben nun eine vierelementige Lösungsmenge

$$V_{\mathbb{C}}(I) = \left\{ \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) \right\}.$$

Modulo dem neuen Ideal I sind X und Y nicht mehr äquivalent, sondern nur noch X^2 und Y^2. Jedes Monom ist somit äquivalent entweder zu einer X-Potenz oder zu einem Monom der Form $X^a Y$. Da auch hier X^2 äquivalent zu $\frac{1}{2}$ ist, gibt es somit in jeder Äquivalenzklasse eines Monoms ein skalares Vielfaches eines der vier Monome $1, X, Y$ oder XY. Da keine Linearkombination dieser Monome in I liegt, bilden ihre Restklassen eine Basis von $\mathbb{Q}[X,Y]/I$, so daß dieser Faktorring als $\mathbb{Q}$-Vektorraum die Dimension vier hat.

In diesen beiden Beispielen waren sowohl die Lösungsmengen als auch Basen der Faktorringe einfach zu finden; im Allgemeinen ist das eher nicht der Fall. Wenn wir aber eine GRÖBNER-Basis des Ideals I kennen, können wir leicht eine Vektorraumbasis des Faktorrings konstruieren:

Definition: $I \lhd k[X_1, \ldots, X_n]$ sei ein Ideal und G sei eine GRÖBNER-Basis bezüglich irgendeiner Monomordnung auf $k[X_1, \ldots, X_n]$. Ein Monom in $X_1, \ldots, X_n$ heißt *Standardmonom* (bezüglich G), wenn es für kein $g \in G$ durch das führende Monom von g teilbar ist.

Satz: Für jede GRÖBNER-Basis G eines Ideals $I \lhd k[X_1, \ldots, X_n]$ bilden die Restklassen der Standardmonome eine Vektorraumbasis von $k[X_1, \ldots, X_n]/I$.

Beweis: Zunächst sind diese Restklassen linear unabhängig, denn jede nichttriviale Linearkombination der Null entspräche einem Polynom

h aus I, dessen sämtliche Monome Standardmonome sind. Da die führenden Monome der Elemente von G das Ideal FM(I) erzeugen, müßte daher FM(h) Vielfaches eines FM(g) mit $g \in G$ sein, was der Definition eines Standardmonoms widerspricht.

Für ein beliebiges $f \in k[X_1, \ldots, X_n]$ liefert uns der Divisionsalgorithmus eine Darstellung

$$f = \sum_{g \in G} a_g g + r \quad \text{mit} \quad a_g, r \in k[X_1, \ldots, X_n]\,,$$

wobei r eine k-Linearkombination von Standardmonomen ist. Da die Summe der $a_g g$ in I liegt, ist f also äquivalent zu einer k-Linearkombination von Standardmonomen, so daß seine Restklasse die entsprechende Linearkombination von deren Restklassen ist. ∎

Dieser Satz gilt unabhängig davon, ob $k[X_1, \ldots, X_n]/I$ als Vektorraum endlichdimensional ist oder nicht. Er liefert uns ein einfaches Kriterium dafür, wann dieser Vektorraum endliche Dimension hat und wann somit die Lösungsmenge $V_K(I)$ endlich ist:

Lemma: G sei eine GRÖBNER-Basis eines Ideals $I \lhd k[X_1, \ldots, X_n]$ bezüglich irgendeiner Monomordnung. $V_K(I)$ ist genau dann endlich, wenn G für jedes i ein Polynom enthält, dessen führendes Monom eine X_i-Potenz ist.

Beweis: Falls die GRÖBNER-Basis für jedes i ein Polynom mit führendem Monom $X_i^{d_i}$ enthält, ist jedes Monom, in dem ein X_i mit einem Exponenten größer oder gleich d_i vorkommt, durch das führende Monom eines Elements der GRÖBNER-Basis teilbar. In den Monomem $X_1^{e_1} \cdots X_n^{e_n}$, für die das nicht der Fall ist, muß daher jedes e_i echt kleiner als das entsprechende d_i sein, Somit gibt es nur endlich viele Standardmonome, d.h. A ist endlichdimensional, und $V_K(I)$ ist endlich.

Ist umgekehrt $V_K(I)$ endlich, so enthält $\bar{I}$ für jedes i ein Polynom aus $K[X_i]$ – siehe Schritt 2 im Beweis des obigen Satzes. Da die GRÖBNER-Basis von I gleichzeitig eine GRÖBNER-Basis von $\bar{I}$ ist, muß das führende Monom eines ihrer Elemente die höchste X_i-Potenz in diesem Polynom teilen, muß also selbst eine Potenz von X_i sein. ∎

Bei diesem Satz ist wichtig, daß er für jede beliebige Monomordnung gilt. Wenn wir mit der lexikographischen Ordnung (oder einer anderen Eliminationsordnung) arbeiten, gibt uns die GRÖBNER-Basis natürlich deutlich mehr Information über $V_K(I)$ als nur die, ob die Menge endlich ist, aber erfahrungsgemäß dauert die Berechnung einer solchen GRÖBNER-Basis auch erheblich länger als beispielsweise bezüglich der graduiert lexikographischen Ordnung. Wenn es nur darum geht, ob $V_K(I)$ endlich ist oder nicht, können wir daher mit einer „billigen" Monomordnung arbeiten.

Eine weitere Anwendung des obigen Satzes soll noch kurz erwähnt werden: Jedes Polynom $f \in k[X_1, \ldots, X_n]$ definiert für jedes Ideal I eine Abbildung

$$\begin{cases} V_K(I) \to K \\ (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n) \end{cases}.$$

Für $f \in I$ wird dabei jeder Punkt auf Null abgebildet, und allgemeiner können wir sagen, daß zwei Polynome f, g die gleiche Abbildung definieren, wenn ihre Differenz in I liegt. (Falls I ein Radikalideal ist, können wir nach der starken Form des HILBERTschen Nullstellensatzes sogar sagen, daß sie *genau* dann die gleiche Funktion definieren. Jede durch Polynome gegebene Funktion $V_K(I) \to K$ läßt sich daher auch beschreiben mit einem Polynom, das eine Linearkombination der Standardmonome bezüglich einer GRÖBNER-Basis ist.

Ist $V_K(I)$ endlich, so gibt es für jede Abbildung $V_K(I) \to K$ ein Interpolationspolynom, das auf $V_K(I)$ mit dieser Funktion übereinstimmt; somit kann dann *jede* Abbildung $V_K(I) \to K$ durch eine Linearkombination von Standardmonomen beschrieben werden. Dies nutzt beispielsweise die algebraische Statistik aus, die zu einer endlichen Stichprobe $M \subset \mathbb{R}^n$ zunächst ein Ideal $I \lhd \mathbb{R}[X_1, \ldots, X_n]$ bestimmt, für das $V_K(I) = M$ ist. (Wie das geht, haben wir uns im Zusammenhang mit Eliminationsidealen überlegt.) Über die Standardmonome erhält man dann eine Übersicht über mögliche lineare Modelle, deren Parameter sich anhand der Stichprobe schätzen lassen. Hier interessiert dann nicht nur die GRÖBNER-Basis bezüglich einer festen Monomordnung, sondern der gesamte sogenannte GRÖBNER-Fächer, d.h. die Menge aller

möglicher Mengen von Standardmonomen bezüglich irgendwelcher Monomordnungen. Obwohl es unendlich viele mögliche Monomordnungen gibt, kann man zeigen, daß dieser Fächer stets endlich ist.