Das Buchberger Kriterium

Martin Albrecht

A*l*ZAGK Seminar
WiSe 2004/2005

Inhaltsverzeichnis

1 Konventionen & Bezeichnungen

Im folgenden werden folgende Bezeichnungen wiederholt verwendet:

- $G = \{g_1, \ldots, g_t\}$ ist eine Gröbner Basis zu einem Ideal I.
- I ist ein Ideal in $k[x_1, \ldots, x_n]$.
- k ist der Körper, über dem der Polynomring definiert ist.
- $f, g = \sum_\alpha a_\alpha x^\alpha$ sind Polynome $\neq 0$ in $k[x_1, \ldots, x_n]$.
- r ist der Rest $\in k[x_1, \ldots, x_n]$ einer Polynomdivision; siehe Lemma 1.
- $multideg(f) = max(\alpha \; in \; \mathbb{N}_0^n : a_\alpha \neq 0)$
- $\mathrm{LC}(f) = a_{multideg(f)} \in k$ ist der Leitkoeffizient eines Polynoms f.
- $\mathrm{LM}(f) = x^{multideg(f)}$ ist das Leitmonom eines Polynoms f.
- $\mathrm{LT}(f) = \mathrm{LC}(f) \cdot \mathrm{LM}(f)$ ist der Leitterm eines Polynoms f.
- In den meisten Beispielen wird als Monomialordnung GRLEX mit $x > y > z$ verwendet.

2 Das Problem, ob $f \in I$

Zur Motivation des Folgenden sei noch einmal auf die Frage eingegangen, wann ein gegebenes f in I liegt. Da es sich hierbei um eine Wiederholung handelt, wird zum Teil auf Beweise verzichtet.

Lemma 1. *Sei $G = \{g_1, \ldots, g_t\}$ eine Gröbner Basis für ein Ideal $I \subset k[x_1, \ldots, x_n]$. Zu gegebenen f existiert ein eindeutiges $r \in k[x_1, \ldots, x_n]$ mit folgenden Eigenschaften:*

1. *Kein Term von r ist durch ein $\mathrm{LT}(g_1) \ldots \mathrm{LT}(g_t)$ teilbar*
2. *Es existiert ein $g \in I$ so dass man die Darstellung $f = g + r$ hat.*

Ohne Beweis.

Korollar 1. *Sei $G = \{g_1, \ldots, g_t\}$ eine Gröbnerbasis für ein Ideal $I \subset k[x_1, \ldots, x_n]$ und sei $f \in k[x_1, \ldots, x_n]$. Dann liegt f im Ideal I genau dann, wenn bei der Division von f durch G der Rest $r = 0$ ist.*

Beweis. $\Rightarrow$: Wenn $r = 0$ ist, dann hat f eine Darstellung als $a_1g_1 + \cdots + a_tg_t$ und liegt damit im Ideal, das von $\{g_1, \ldots, g_t\}$ aufgespannt wird.

$\Leftarrow$: Das Polynom f hat eine Darstellung als $a_1g_1 + \cdots + a_tg_t + r$, wenn es durch G dividiert wird. Da nun sowohl f als auch die $g_i \in I$, gilt $r = f - a_1g_1 - \cdots - a_tg_t \in I$. Ist $r \neq 0$, dann gilt $\mathrm{LT}(r) \in \langle \mathrm{LT}(I) \rangle = \langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_t) \rangle$. Also muss $\mathrm{LT}(r)$ durch ein $\mathrm{LT}(g_i)$ teilbar sein, was dem Begriff des Restes bei Division durch G widerspricht[1].

□

Damit lässt sich die Frage beantworten, ob ein gegebenes Polynom f in einem Ideal I mit Gröbner Basis G liegt. Man teilt einfach f durch G, genau dann wenn der Rest $r = 0$ ist, dann gilt $f \in I$.

3 Das Problem, ob G eine Gröbner Basis ist

Um nun herauszufinden, ob eine Polynommenge $G = \{f_1, \ldots, f_t\}$ eine Gröbner Basis zu einem gegebenen Ideal I ist, erinnere man sich, dass für die Gröbnerbasis G folgendes gelten muss: $\langle \mathrm{LT}(f_1), \ldots, \mathrm{LT}(f_t) \rangle = \langle \mathrm{LT}(I) \rangle$. Kann man nun ein Element aus $\langle \mathrm{LT}(I) \rangle$ konstruieren, das $\notin \langle \mathrm{LT}(f_1), \ldots, \mathrm{LT}(f_t) \rangle$, dann ist G keine Gröbnerbasis. Man betrachtet dafür – für zwei geeignete Elemente aus G –, ob sich im Term $ax^\alpha f_i - bx^\beta f_j$ die Leitterme $\mathrm{LT}(f_i)$ und $\mathrm{LT}(f_j)$ so gegenseitig wegheben, dass $\mathrm{LT}(ax^\alpha f_i - bx^\beta f_j) \notin \langle \mathrm{LT}(f_1), \ldots, \mathrm{LT}(f_t) \rangle$. Da aber auf der anderen Seite $ax^\alpha f_i - bx^\beta f_j \in I$ gilt, liegt $\mathrm{LT}(ax^\alpha f_i - bx^\beta f_j)$ in $\langle \mathrm{LT}(I) \rangle$. Damit wäre G keine Gröbnerbasis.

[1] Siehe dazu [Cox, S.63]

Beispiel. *Sei* $f_1 = x^3 - 2xy$ *und* $f_2 = x^2y - 2y^2 + x$ *sowie* $I = \langle f_1, f_2 \rangle$. *Als Monomialordnung wird* GRLEX *mit* $x > y$ *verwendet. Die Leitterme ergeben sich dann als*

$$\begin{aligned} \mathrm{LT}(f_1) &= x^3 \\ \mathrm{LT}(f_2) &= x^2y. \end{aligned}$$

Es gilt:

$$\begin{aligned} x^2 &= x \cdot (x^2y - 2y^2 + x) - y \cdot (x^3 - 2xy) \\ x^2 &= x \cdot f_2 - y \cdot f_1 \in I \end{aligned}$$

Also ist $x^2 \in \mathrm{LT}(I)$, *aber* x^3 *und* x^2y *produzieren niemals* x^2. *Damit liegt* x^2 *nicht im Ideal* $\langle \mathrm{LT}(f_1), \mathrm{LT}(f_2) \rangle$. *Demnach ist* $\{f_1, f_2\}$ *keine Gröbner Basis.*

Definition 1. *Seien* $f, g \in k[x_1, \ldots, x_n]$ *Polynome* $\neq 0$.

1. *Wenn* α *der Multigrad von* f *ist* β *der von* g, *dann sei* $\gamma = (\gamma_1, \ldots, \gamma_n)$, *wobei* $\gamma_i = max(\alpha_i, \beta_i)$ *für jedes* $i \leq n$. x^γ *ist das kleinste gemeinsame Vielfache*[2] *von* $\mathrm{LM}(f)$ *und* $\mathrm{LM}(g)$, *geschrieben* $x^\gamma = LCM(\mathrm{LM}(f), \mathrm{LM}(g))$.
2. *Das S-Polynom von* f *und* g *wird definiert als*

$$S(f, g) = \frac{x^\gamma}{\mathrm{LT}(f)} \cdot f - \frac{x^\gamma}{\mathrm{LT}(g)} \cdot g.$$

Beispiel. *Seien* $f_1 = x^3 - 2xy$ *und* $f_2 = x^2y - 2y^2 + x$ *wie oben. Die Leitterme sind* $\mathrm{LT}(f_1) = x^3$ *und* $\mathrm{LT}(f_2) = x^2y$, *damit ist* $x^\gamma = x^3y$. *Das S-Polynom ergibt sich damit als:*

$$\begin{aligned} S(f_1, f_2) &= \frac{x^\gamma}{\mathrm{LT}(f_1)} \cdot f_1 - \frac{x^\gamma}{\mathrm{LT}(f_2)} \cdot f_2 \\ S(f_1, f_2) &= \frac{x^3y}{x^3} \cdot (x^3 - 2xy) - \frac{x^3y}{x^2y} \cdot (x^2y - 2y^2 + x) \\ S(f_1, f_2) &= y \cdot (x^3 - 2xy) - x \cdot (x^2y - 2y^2 + x) \\ S(f_1, f_2) &= x^3y - 2xy^2 - x^3y + 2y^2x - x^2 \\ S(f_1, f_2) &= -x^2 \end{aligned}$$

Wie man im Beispiel sehen kann, ist $S(f_i, f_j)$ so gebaut, dass sich die Leitterme wegheben. Es ist aber darüber hinaus so, dass, wann immer sich Leitterme von Polynomen gleichen Multigrades wegheben, dies auf S-Polynome zurückgeführt werden kann.

Lemma 2. *Habe jedes Element aus* $\Sigma_{i=1}^{t} c_i x^{\alpha(i)} g_i$, *mit Konstanten* $c_1, \ldots, c_n$, *Multigrad* δ *wenn* $c_i \neq 0$, *also* $\alpha(i) + multideg(g_i) = \delta \in \mathbb{N}_0^n$. *Hat nun die Summe kleineren Multigrad, dann*

[2] sozusagen der Super-Leitterm von f und g

existieren Konstanten c_{jk} so dass

$$\sum_{i=1}^{t} c_i x^{\alpha(i)} g_i \quad = \quad \sum_{j=1}^{t-1} c_{jk} x^{\delta-\gamma_{jk}} S(g_j, g_k) \tag{1}$$

Wobei $k = j+1$ und $x^{\gamma_{jk}} = LCM(\mathrm{LM}(g_j), \mathrm{LM}(g_k))$. Weiterhin hat jedes $x^{\delta-\gamma_{jk}} S(g_j, g_k)$ einen Multigrad $< \delta$.

Auf der linken Seite in Gleichung (1) heben sich die Grade der Leitterme erst nach dem Addieren auf, auf der rechten Seite jedoch haben die Terme bereits geringeren Multigrad, die Terme sind also schon weggehoben. Damit sind die S-Polynome für das Wegheben verantwortlich.

Beweis. Sei $d_i = LC(g_i)$, so dass $c_i d_i$ der Leitkoeffizient von $c_i x^{\alpha(i)} g_i$. Da die $c_i x^{\alpha}(i) g_i$ alle den Multigrad δ haben, ihre Summe jedoch einen kleineren Multigrad besitzt, muss $\Sigma_{i=1}^{t} c_i d_i = 0$ sein.

Sei weiterhin $p_i = x^{\alpha(i)} g_i \ / \ d_i$ und damit der Leitkoeffizient von $p_i = 1$. Man betrachtet nun die „Teleskopsumme“:

$$\begin{aligned}
\sum_{i=1}^{t} \quad & c_i x^{\alpha(i)} g_i \quad && = \\
\sum_{i=1}^{t} \quad & c_i \frac{d_i}{d_i} x^{\alpha(i)} g_i \quad && = \\
\sum_{i=1}^{t} \quad & c_i d_i p_i \quad && = c_1 d_1 (p_1 - p_2) \ + \ (c_1 d_1 \ + \ c_2 d_2)(p_2 \ - \ p_3) \ + \ \cdots \\
& && \quad + (c_1 d_1 \ + \ \cdots \ + \ c_{t-1} d_{t-1})(p_{t-1} - p_t) \\
& && \quad + (c_1 d_1 \ + \ \cdots \ + \ c_t d_t) p_t \\
\sum_{i=1}^{t} \quad & c_i d_i p_i \quad && = c_1 d_1 (p_1 - p_2) \ + \ (c_1 d_1 \ + \ c_2 d_2)(p_2 \ - \ p_3) \ + \ \cdots \\
& && \quad + (c_1 d_1 \ + \ \cdots \ + \ c_{t-1} d_{t-1})(p_{t-1} - p_t) \\
& && \quad + 0
\end{aligned}$$

Nun sei $\mathrm{LT}(g_i) = d_i x^{\beta(i)}$. Nach Voraussetzung gilt $\alpha(i) + \beta(i) = \delta$ für alle i, da alle $c_i x^{\alpha(i)} g_i = c_i x^{\alpha(i)} d_i x^{\beta(i)}$ Multigrad δ haben. Jedes Leit-Monom $\mathrm{LM}(g_i) = x^{\beta(i)}$ teilt damit x^{δ}. Ebenso teilt $x^{\gamma_{jk}} = LCM(\mathrm{LM}(g_j), \mathrm{LM}(g_k))$ das Monom x^{δ}, damit ist $x^{\delta-\gamma_{jk}}$ ein Monom und es ergibt sich:

$$\begin{aligned}
x^{\delta-\gamma_{jk}} S(g_j, g_k) \quad &= \quad x^{\delta-\gamma_{jk}} \left(\frac{x^{\gamma_{jk}}}{\mathrm{LT}(g_j)} g_j \ - \ \frac{x^{\gamma_{jk}}}{\mathrm{LT}(g_k)} g_k \right) \\
&= \quad \frac{x^{\delta}}{d_j x^{\beta(j)}} g_j \ - \ \frac{x^{\delta}}{d_k x^{\beta(k)}} g_k \\
&= \quad \frac{x^{\alpha(j)}}{d_j} g_j \ - \ \frac{x^{\alpha(k)}}{d_k} g_k \\
&= \quad p_j \ - \ p_k
\end{aligned} \tag{2}$$

Dies wird nun in die Teleskopsummengleichung (2) von oben eingesetzt, und es ergibt sich:

$$\begin{aligned}\sum_{i=1}^{t} c_i x^{\alpha(i)} g_i &= c_1 d_1 x^{\delta-\gamma_{12}} S(g_1, g_2) + (c_1 d_1 + c_2 d_2) x^{\delta-\gamma_{23}} S(g_2, g_3) \\ &\quad + \cdots + (c_1 d_1 + \cdots + c_{t-1} d_{t-1}) x^{\delta-\gamma_{t-1,t}} S(g_{t-1}, g_t) \\ &= \sum_{j=1}^{t-1} c_{jk} x^{\delta-\gamma_{jk}} S(g_j, g_k)\end{aligned}$$

Da die p_j und p_k Multigrad[3] δ und Leit-Koeffizient 1 haben, hat die Differenz $p_j - p_k$ kleineren Multigrad. Da nach Gleichung (2) $p_j - p_k = x^{\delta-\gamma_{jk}} S(g_j, g_k)$ ist, gilt diese Behauptung auch für $x^{\delta} - \gamma_{jk} S(g_j, g_k)$, und das Lemma ist damit bewiesen. □

4 Das Buchberger Kriterium

Mit S-Polynomen und Lemma 2 kann ein Kriterium dafür formuliert und bewiesen werden, wann eine Basis eine Gröbner Basis ist.

Theorem 1. *Sei I ein Ideal. $G = \{g_1, \ldots, g_t\}$ ist genau dann eine Gröbner Basis für I, wenn für alle Paare $i \neq j$, der Rest bei der Division von $S(g_i, g_j)$ durch G Null ist.*

Beweis. $\Rightarrow$: Wenn G eine Gröbner Basis ist, dann ist der Rest $r = 0$, da $S(g_i, g_j) \in I$.

$\Leftarrow$:

Umformulierung des Problems Sei $f \in I$ ein Polynom. Es ist zu zeigen, dass ,wenn alle S-Polynome Rest $r = 0$ haben, $\mathrm{LT}(f) \in \langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_t) \rangle$. Da f in I liegt, gibt es eine Darstellung von f als

$$f = \sum_{i=1}^{t} h_i g_i \tag{3}$$

Damit folgt, dass

$$multideg(f) \leq max(multideg(h_i g_i)) = \delta \tag{4}$$

Wenn nun alle Möglichkeiten für die Gleichung (3) betrachtet werden, dann gibt es ein kleinstes δ für die Gleichung (4), da die Monominalordnung eine Wohlordnung ist. Wenn nun die Gleichung (3) so gewählt wird, dass δ in Gleichung (4) minimal ist, dann – so die Behauptung – gilt $multideg(f) = \delta$. Damit wäre $\mathrm{LT}(f) \in \langle \mathrm{LT}(g_1), \ldots \mathrm{LT}(g_t) \rangle$ und das Theorem bewiesen. Wir beweisen nun, dass $multideg(f) = \delta$, wenn δ minimal ist.

[3] $x^{\alpha(i)} g_i$ hat Multigrad δ und d_i ist konstant

Problem Reduzierung Es kann nur gelten, dass entweder $multideg(f) = \delta$ oder $multideg(f) < \delta$ ist. Wird also eine der beiden Eigenschaften zum Widerspruch geführt, dann muss die andere gelten. Um die Aussage $multideg(f) < \delta$ zum Widerspruch zu führen, schreiben wir die Gleichung (3) so um, dass die Terme mit Multigrad δ von den Termen mit kleinerem Multigrad getrennt werden. Es ergibt sich

$$\begin{aligned} f &= \sum_{m(i)=\delta} h_i g_i + \sum_{m(i)<\delta} h_i g_i \\ &= \sum_{m(i)=\delta} \mathrm{LT}(h_i) g_i + \sum_{m(i)=\delta} (h_i - \mathrm{LT}(h_i)) g_i + \sum_{m(i)<\delta} h_i g_i. \end{aligned} \tag{5}$$

Da wir in der zweiten Zeile in der zweiten Summe des Terms den Leitterm subtrahieren, wird das Ergebnis einen Multigrad $< \delta$ haben, in der dritten Summe haben schon vor dem Summieren die Terme kleineren Multigrad. Nach unserer Annahme $multideg(f) < \delta$ muss die erste Summe auch Multigrad $< \delta$ haben.

Umformulierung als Term in S-Polynomen Sei nun $\mathrm{LT}(h_i) = c_i x^{\alpha(i)}$. Damit hat die erste Summe

$$\sum_{m(i)=\delta} \mathrm{LT}(h_i) g_i = \sum_{m(i) = \delta} c_i x^{\alpha(i)} g_i$$

genau die Form aus dem Lemma 2. Die $c_i x^{\alpha(i)} g_i$ haben Multigrad δ und die Summe kleineren Multigrad, also gilt:

$$\sum_{m(i)=\delta} LT(h_i) g_i = \sum_{m(i)=\delta} c_i x^{\alpha(i)} g_i = \sum_{j=1}^{t-1} c_{jk} x^{\delta - \gamma_{jk}} S(g_j, g_k), \tag{6}$$

mit $c_{jk} \in k$ und $x^{\gamma_{jk}} = LCM(\mathrm{LM}(g_j), \mathrm{LM}(g_k))$. Weiterhin gilt, dass die $c_{jk} x^{\delta - \gamma_{jk}} S(g_j, g_k)$ Multigrad $< \delta$ haben.

Ersetzungsmöglichkeit der S-Polynome Da alle $S(g_j, g_k)$ bei Division durch G Rest $r = 0$ ergeben, haben sie die Darstellung als

$$S(g_j, g_k) = \sum_{i=1}^{t} a_{ijk} g_i, \tag{7}$$

wobei die a_{ijk} aus $k[x_1, \ldots, x_n]$.

Durch den Divisionsalgorithmus wissen wir[4], dass

$$multideg(a_{ijk} g_i) \leq multideg(S(g_j, g_k)) \tag{8}$$

für alle i, j, k gilt. Dass heißt, dass unter der Bedingung, dass der Rest $r = 0$ ist, Ausdrücke für $S(g_j, g_k)$ gefunden werden können, in denen weniger Leitterme weggehoben werden.

Ersetzung S-Polynome Um diese Möglichkeit auszunutzen/wirklich werden zu lassen, multi-

[4]Siehe [Cox, S.63ff]

plizieren wir den Ausdruck $S(g_j, g_k)$ mit $x^{\delta-\gamma_{jk}}$ und erhalten

$$x^{\delta-\gamma_{jk}} S(g_j, g_k) \;=\; \sum_{i=1}^{t} b_{ijk} g_i, \tag{9}$$

mit $b_{ijk} = x^{\delta-\gamma_{jk}} a_{ijk}$. Aus Gleichung (8) und Lemma 2 folgt dann, dass

$$multideg(b_{ijk} g_i) \;\leq\; multideg(x^{\delta-\gamma_{jk}} S(g_j, g_k)) \;<\; \delta \tag{10}$$

Jetzt setzen wir in Gleichung (6) die Substitution (9) ein, und erhalten

$$\begin{aligned}
\sum_{m(i)=\delta} \mathrm{LT}(h_i) g_i &= \sum_{j,k} c_{jk} x^{\delta-\gamma_{jk}} S(g_j, g_k) \\
&= \sum_{j,k} c_{jk} \left(\sum_{i} b_{ijk} g_i \right) \\
&= \sum_{i} \left(\sum_{j,k} c_{jk} b_{ijk} \right) g_i \\
&= \sum_{i} \tilde{h}_i g_i.
\end{aligned}$$

. Es folgt nun aus Gleichung (10), dass

$$multideg(\tilde{h}_i g_i) \;<\; \delta,$$

da die c_{jk} Konstanten sind und am Multigrad nichts ändern.

Widerspruch zur Annahme δ sei minimal Abschließend setzen wir $\Sigma_i \tilde{h}_i g_i$ in die Gleichung (5) ein und erhalten

$$f \;= \sum_{i} \tilde{h}_i g_i \;+\; \sum_{m(i)=\delta} (h_i \;-\; \mathrm{LT}(h_i)) g_i \;+\; \sum_{m(i)<\delta} h_i g_i$$

Damit ist f in den g_i ausgedrückt, wobei alle Terme kleineren Multigrad als δ haben. Das widerspricht aber der Annahme, dass

$$\delta \;=\; max(\; multideg(h_1 g_1),\; \ldots\; , multideg(h_t g_t))$$

minimal sei. Die Aussage $multigrad(f) \;<\; \delta$ ist damit zum Widerspruch geführt, und es muss gelten $multigrad(f) \;=\; \delta$. Damit gilt auch das Theorem 1.

□

Beispiel. *Sei $I \;=\; \langle y - x^2, z - x^3 \rangle$. Die Behauptung ist nun, dass $G \;=\; \{y - x^2, z - x^3\rangle$ eine Gröbner Basis bei Ordnung* LEX *mit $y \;>\; z \;> x$.*

Das S-Polynom ist

$$S(y - x^2, z - x^3) \;=\; \frac{yz}{y}(y - x^2) \;-\; \frac{yz}{z}(z - x^3) \;=\; -zx^2 + yx^3$$

und mit dem Divisionsalgorithmus lässt sich dieses darstellen als

$$-zx^2 + yx^3 \;=\; x^3 \;\cdot\; (y - x^2) \;+\; (-x^2) \;\cdot\; (z - x^3) + 0.$$

Also ist der Rest $r \;=\; 0$ *und damit gezeigt, dass* G *eine Gröbner Basis ist.*

Beispiel. *Sei* $I \;=\; \langle x^3 - 2xy, x^2y - 2y^2 \rangle$*. Die Behauptung ist nun, dass* $G \;=\; \{x^3 - 2xy, x^2y - 2y^2\rangle$ *keine Gröbner Basis bei Ordnung* GRLEX *ist.*

Das S-Polynom ist

$$S(x^3 - 2xy, x^2y - 2y^2) \;=\; -x^2$$

und mit dem Divisionsalgorithmus lässt sich dieses darstellen als

$$-x^2 \;=\; 0 \;\cdot\; (x^3 - 2xy) \;+\; 0 \;\cdot\; (x^2y - 2y^2) \;- x^2$$

Also ist der Rest $r \;=\; -x^2$ *und damit gezeigt, dass* G *keine Gröbner Basis ist.*

Literatur

[Cox] Cox, Little, O'Shea, *Ideals, Varieties, and Algorithms* ; Springer, New York [u.a.] 1992