

Ein symbolischer Ansatz für finite Differenzenverfahren partieller Differentialgleichungen

Christian Dinger

Vom Fachbereich Mathematik
der Technischen Universität Kaiserslautern
zur
Verleihung des akademischen Grades
Doktor der Naturwissenschaften
(Doctor rerum naturalium, Dr. rer. nat)
genehmigte

Dissertation

1. Gutachter: Prof. Dr. Gerhard Pfister
2. Gutachter: Prof. Dr. Eva Zerz

Datum der Disputation: 20. Oktober 2010

D 386

Inhaltsverzeichnis

Vorwort	v
1 Computeralgebra	1
1.1 Ringe und Ideale	1
1.2 Polynomringe und Ordnungen	6
1.3 Gröbnerbasen	13
1.3.1 Gröbnerbasen für Ringe	13
1.3.2 Gröbnerbasen für Moduln	23
2 Finite Differenzenschemata	29
2.1 Einführung	29
2.2 Konvergenztheorie	31
2.2.1 Konsistenz und Konvergenz	33
2.2.2 Stabilität	36
2.2.3 Die Methode nach Von Neumann	37
3 Quantorenelimination im \mathbb{R}^r	45
3.1 Motivation	45
3.2 Zylindrische Algebraische Zerlegung im \mathbb{R}^r	46
3.3 Lösungsformel	53
4 Algebraische Methoden für Finite Differenzenschemata	55
4.1 Motivation	55
4.2 Theorie der algebraischen Methode	56
4.2.1 Kontinuierlicher Fall	57
4.2.2 Diskreter Fall	58
4.3 Lineare Partielle Differentialgleichungen	78
5 SINGULAR und QEPCAD	87
5.0.1 SINGULAR	87
5.0.2 QEPCAD	88

Inhaltsverzeichnis

5.0.3 findiff.lib 90

Literaturverzeichnis **99**

Vorwort

Überblick

Diese Arbeit beschäftigt sich mit der algebraischen Lösung von numerischen Problemen, die im Zusammenhang mit finiten Differenzenverfahren partieller Differentialgleichungen auftreten. Ein finites Differenzenverfahren ist der Versuch, nach Diskretisierung des Raumes durch Differenzenquotienten eine Approximation der Differentialquotienten einer unbekanntem Funktion herbeizuführen, somit eine Relation der auftretenden diskreten Funktionswerte untereinander zu bestimmen und diese iterativ aus gegebenen Anfangswerten konkret zu berechnen. Neben dem hohen Rechenaufwand, der im Lösen von sehr großen linearen Gleichungssystemen besteht, müssen hinsichtlich der Beschaffenheit dieser Relationen der diskreten Lösungswerte einige Bedingungen beachtet werden. Zum einen sind Approximationen immer mit Fehlern behaftet, sodaß ein Ziel sicherlich die Generierung von *konsistenten* Differenzenverfahren ist, das sind Verfahren, die für Schrittweiten $\Delta t, \Delta x \rightarrow 0$ einen Fehler besitzen, der ebenfalls gegen Null konvergiert. Zum anderen ist wichtig, daß dieser intrinsische *lokale* Fehler des Verfahrens sich in den Iterationsschritten nicht vergrößert, damit für errechnete Werte aus späteren Iterationsschritten eine Aussage über die Größe des Fehlers zu den wirklichen Werten möglich ist. Diese Eigenschaft wird Stabilität genannt und ist in einer besonderen Situation eine notwendige und hinreichende Bedingung für die Konvergenz. Die Bestimmung der Stabilität eines Verfahrens und insbesondere das Finden von Bedingungen für Stabilität ist die eigentliche Schwierigkeit in der Untersuchung eines Verfahrens, da in diesem Schritt aus polynomialen Ungleichungen äquivalente Bedingungen der Parameter untereinander hergeleitet werden müssen für den Fall, daß dies überhaupt möglich ist. An dieser Stelle schafft Quantorenelimination mittels zylindrischer algebraischer Zerlegung Abhilfe: es werden in Implementierungen dieses Konzepts äquivalente Bedingungen hergeleitet, die Aufschluß über die Stabilität des Verfahrens geben, sofern diese prinzipiell möglich ist.

Die Erzeugung eines Differenzenverfahrens besteht im wesentlichen darin, Approximationen in die ursprüngliche Differentialgleichung einzusetzen und eine Relation der diskreten Werte zum Zweck ihrer iterativen Berechnung herbeizuführen. Diese Aufgabe können Gröbnerbasen in einem geeigneten Ring erledigen.

Aufbau

Diese Arbeit wie folgt aufgebaut:

Kapitel 1: Es werden Eigenschaften von Idealen, Ringen sowie die Grundzüge der Gröbnerbasentheorie im Fall von globalen Ordnungen erklärt und entsprechende Beispiele in SINGULAR vorgeführt. Referenzen hierfür sind [GP08, AM69, Eis95, CLO97].

Kapitel 2: Die Einführung in die numerischen Methoden mithilfe finiter Differenzenverfahren (insbesondere Zweischrittverfahren) wird zusammen mit der Beschreibung der Konvergenztheorie und der besonderen Schwerpunktlegung auf den Von-Neumann-Ansatz zur Stabilitätsanalyse für lineare Gleichungen vorgenommen. Referenzen für dieses Kapitel sind [LeV92, Tho95, Kai08].

Kapitel 3: Ausgehend von einer Formel in polynomialen Ausdrücken wird die Konstruktion einer äquivalenten Lösungsformel unter Zuhilfenahme der zylindrischen algebraischen Zerlegung im System QEPCAD beschrieben. Die Darstellung ist vornehmlich durch [Bro99, ACM98a] motiviert.

Kapitel 4: Dieses Kapitel präsentiert neue Ergebnisse aus [DL10] zur Generierung von finiten Differenzenschemata für eine Gleichung und den Fall von Systemen. Hier werden neben den Bedingungen für die Existenz auch unterschiedliche (aber äquivalente) Wege der Erzeugung eines Schemas diskutiert und Beispiele in SINGULAR gezeigt. Der Ansatz ist möglichst allgemein gewählt, sodaß keine Bedingung außer gewissen Eigenschaften für die Existenz einer Ordnung vorausgesetzt werden. Anschließend werden für den speziellen Fall linearer Gleichungen Resultate aus [LM10] ausgeführt.

Kapitel 5: Die SINGULAR-Bibliothek `findiff.lib` ([Din10]) ermöglicht im Falle linearer Gleichungen die Anwendung algebraischer Methoden, die in Kapitel 4 genannt werden, sodaß eine vollständige Automatisierung der Verfahrensuntersuchung allein mit algebraischen Mitteln bereitgestellt wird. Es werden zusammen mit `findiff.lib` die Systeme SINGULAR ([DGPS09]) und QEPCAD ([Bro10]) vorgestellt, die frei erhältlich sind, und mit deren Hilfe

diese algebraischen Methoden zum Einsatz kommen.

Danksagung

Der Autor bedankt sich sehr herzlich bei seinem Betreuer Prof. Dr. Gerhard Pfister, der ihm die Möglichkeit zur Forschung in diesem Grenzgebiet zwischen Algebra und Numerik gab. Ein ganz besonderer Dank gebührt dem Co-Betreuer Dr. Viktor Levandovskyy (Aachen), mit dem der Autor einen großen Teil der Diskussionen über diese Arbeit und die zugrundeliegenden Fragestellungen führte, der ihm zu Motivation in diesem ungewohnten Gebiet verhalf und wertvolle Denkanstöße zu den neuesten Entwicklungen beitrug. Prof. Dr. Axel Klar (Kaiserslautern) sei zusammen mit aufrichtigen Genesungswünschen gedankt für die stete Bereitschaft zu Diskussionen beim Thema Stabilität von Differenzenverfahren. Weitere Danksagungen werden an dieser Stelle gegenüber Dr. Hans Schönemann (Kaiserslautern) für das schnelle Helfen bei Problemen mit SINGULAR, Prof. Dr. Christopher W. Brown (Annapolis) für Hilfen zum System QEPCAD und Prof. Dr. Eva Zerz (Aachen) für ihre Bereitschaft zur Aushilfe in einer schwierigen Situation ausgesprochen.

Kapitel 1

Computeralgebra

In diesem Kapitel wird der Teil der Theorie der kommutativen Algebra dargestellt, der im weiteren Verlauf dieser Arbeit wesentlich ist. Besonderes Gewicht liegt auf dem Konzept der Gröbnerbasis in einem Polynomring in endlich vielen Variablen, sowie den in diesem Zusammenhang auftretenden Begriffen.

1.1 Ringe und Ideale

Vor allem steht die Definition eines Rings, der mit Ausnahme der Invertierbarkeit von Elementen, die ungleich Null sind, alle Eigenschaften besitzt, die ein Körper besitzt:

Definition 1.1. Ringe und Körper

Ein *kommutativer Ring mit Eins* ist eine Menge $(R, +, \cdot)$, wobei R mit den zwei binären Operationen

$$+, \cdot : R \times R \longrightarrow R$$

die Eigenschaft besitzt, daß für alle $r_1, r_2, r_3 \in R$ gilt:

$$(R, +, 0) \text{ ist eine abelsche (kommutative) Gruppe}$$

$$r_1 \cdot (r_2 \cdot r_3) = (r_1 \cdot r_2) \cdot r_3 \text{ (Assoziativität)}$$

$$r_1 \cdot (r_2 + r_3) = r_1 \cdot r_2 + r_1 \cdot r_3$$

$$(r_2 + r_3) \cdot r_1 = r_2 \cdot r_1 + r_3 \cdot r_1 \text{ (Distributivität)}$$

$$1 \cdot r_1 = r_1 \text{ (neutrales Element 1)}$$

Kapitel 1 Computeralgebra

$$r_1 \cdot r_2 = r_2 \cdot r_1 \text{ (Kommutativität)}$$

Wenn überdies in einem Ring R noch gilt

$$\forall r \in R \setminus \{0\} \exists r' \in R : r \cdot r' = 1$$

dann heißt R ein *Körper*.

Es können Ringe auch ohne die Forderung nach einem Eins-Element oder Kommutativität in der Multiplikation definiert werden, jedoch ist diese Allgemeinheit hier nicht erforderlich, weshalb die Beschränkung auf Ringe mit Eins im weiteren vorausgesetzt wird. *Ring* sei ab jetzt Bezeichnung für einen kommutativen Ring mit Eins.

Einige elementare Beispiele zur obigen Definition sind kanonisch:

1. Die natürlichen Zahlen \mathbb{N} erfüllen bezüglich $+$ nicht die Eigenschaft einer Gruppe, weil keine additiven Inversen zu Elementen ungleich Null existieren. Somit ist \mathbb{N} auch kein Ring.
2. Ein Beispiel für einen Ring sind die ganzen Zahlen $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$ mit der bekannten Addition und Multiplikation. Sie stellen aber keinen Körper dar, da zum Beispiel für $3 \in \mathbb{Z}$ keine ganze Zahl $z \in \mathbb{Z}$ mit existiert mit $3 \cdot z = 1$.
3. Die rationalen Zahlen \mathbb{Q} sind ein Körper, da hier für alle Elemente $\neq 0$ auch multiplikative Invertierbarkeit vorliegt.
4. Die reellen Zahlen \mathbb{R} sind ebenfalls ein Körper.

In Ringen R sind auch bestimmte Teilmengen $I \subseteq R$ mit besonderer Struktur interessant:

Definition 1.2. Ideale

1. Eine Menge $I \subseteq R$ heißt *Ideal* in einem Ring $(R, +, \cdot)$ wenn gilt
 - $(I, +)$ ist eine abelsche Untergruppe von $(R, +, 0)$
 - $R \cdot I \subseteq I$, und somit ist I multiplikativ abgeschlossen:

$$\forall r \in R \forall i \in I : r \cdot i \in I$$

Man schreibt

$$I \trianglelefteq R$$

2. Für eine beliebige Menge $M \subseteq R$ sei mit

$$\langle M \rangle_R := \bigcap_{M \subseteq N \trianglelefteq R} N$$

das von M erzeugte Ideal bezeichnet. Dies ist der Durchschnitt aller Ideale N , die die Menge M enthält.

3. Ein von nur einem Element erzeugtes Ideal

$$I = \langle m \rangle_R = \{rm \mid r \in R\} \trianglelefteq R$$

heißt *Hauptideal*.

Aus der Definition in 1.2 folgen einige Beobachtungen:

Bemerkung 1.3.

- Wenn für ein Ideal I gilt, daß $1 \in I$, dann folgt aus der Abgeschlossenheit der Multiplikation, daß $I = R$.
- $\langle M \rangle_R$ ist das (mengentheoretisch) kleinste Ideal, welches M enthält. Im angenommenen Fall eines kommutativen Ringes mit Eins gilt, daß

$$\langle M \rangle_R = \left\{ \sum_{j=1}^n r_j \cdot m_j \mid m_j \in M, r_j \in R \right\}$$

und somit das erzeugte Ideal $\langle M \rangle_R$ der Menge aller endlichen Linearkombinationen entspricht.

- Für einen Ring R gilt

$$K \text{ ist ein Körper} \iff \langle 0 \rangle_R \text{ und } R \text{ sind die einzigen Ideale}$$

Für zwei Ringe A, B sind strukturerhaltende Abbildungen, auch Homomorphismen genannt, nützlich:

Kapitel 1 Computeralgebra

Definition 1.4. Seien $(A, +_A, \cdot_A)$ und $(B, +_B, \cdot_B)$ zwei Ringe und $\phi : A \rightarrow B$ eine Abbildung mit

$$\phi(a \cdot_A \acute{a}) = \phi(a) \cdot_B \phi(\acute{a})$$

$$\phi(a +_A \acute{a}) = \phi(a) +_B \phi(\acute{a})$$

$$\phi(1_A) = 1_B$$

für alle $a, \acute{a} \in A$, dann heißt ϕ (*Ring-*)*Homomorphismus* und B ist eine *A-Algebra*, die mittels der Abbildung ϕ die Struktur von A erhält.

Das Urbild von 0_B

$$\text{Ker } \phi := \{a \in A \mid \phi(a) = 0_B\}$$

heißt *Kern* der Abbildung ϕ .

Trivialerweise ist der Kern eines Homomorphismus ϕ ein Ideal im Definitionsbereich:

$$\text{Ker } \phi \trianglelefteq A$$

Ein Beispiel für einen Homomorphismus bietet die sogenannte *Restklassenabbildung*:

Definition 1.5. Es sei $I \trianglelefteq A$ ein Ideal in einem Ring und eine Äquivalenzrelation \equiv auf A definiert mit

$$a \equiv b \pmod{I} \iff a - b \in I$$

dann bezeichne $[a] := a + I$ die Äquivalenzklasse eines jeden Elements a . Die Menge aller Äquivalenzklassen $A/I = \{[a] \mid a \in A\}$ mit der natürlichen Vererbung der binären Verknüpfungen $+$ und \cdot via

$$[a + \acute{a}] := [a] + [\acute{a}]$$

$$[a \cdot \acute{a}] := [a] \cdot [\acute{a}]$$

stellt offensichtlich einen Ring dar. Die Abbildung ϕ mit

$$\phi : A \longrightarrow A/I$$

$$a \longmapsto [a]$$

ist ein surjektiver Homomorphismus mit den in 1.4 definierten Eigenschaften. Für den Kern dieser Abbildung gilt

$$\text{Ker } \phi = I$$

Der Begriff der *Endlichkeit* ist in der Mathematik für die Praxis aber auch in der Theorie von enormer Relevanz. Man denke an endlich erzeugte Vektorräume, an abzählbare Umgebungsbasen für Punkte in topologischen Räumen (hier ist Abzählbarkeit die Endlichkeitsbedingung) oder auch an Mengen mit endlichen Maßen in der Maßtheorie.

In der kommutativen Algebra ist mit Endlichkeit meist *endliche Erzeugung* gemeint und es hat sich somit der Begriff der *noetherschen Eigenschaft* herausgebildet, welcher nachfolgend definiert wird:

Definition 1.6. Ein Ring A heißt *noethersch*, wenn jede aufsteigende Kette $(I_s)_{s \in \mathbb{N}}$ von Idealen in A

$$I_0 \subseteq I_1 \subseteq \dots \subseteq I_k \subseteq I_{k+1} \subseteq \dots \subseteq A$$

ab einem Index n_0 *stationär* ist: $\exists n_0 \quad \forall n \geq n_0 \quad I_n = I_{n_0}$

Es gilt ein wichtiger Zusammenhang zwischen der noetherschen Eigenschaft in 1.6 und der endlichen Erzeugbarkeit von Idealen für einen Ring:

Theorem 1.7. *Für einen Ring A sind die folgenden beiden Eigenschaften äquivalent:*

1. A ist ein noetherscher Ring
2. Jedes Ideal I in A ist endlich erzeugt, das heißt

$$\forall I \trianglelefteq A \quad \exists a_1, \dots, a_m \in A : \quad I = \langle a_1, \dots, a_m \rangle_A$$

wobei m von Ideal zu Ideal variieren kann.

Beweis. Der Beweis folgt unmittelbar Betrachtung einer aufsteigenden Idealkette $\langle a_1 \rangle_A \subset \langle a_1, a_2 \rangle_A \dots$ mit beliebigen $a_i \in I$. \square

1.2 Polynomringe und Ordnungen

Es wurden in obiger Definition von Idealen und Ringen keinerlei Endlichkeitseigenschaften vorausgesetzt. Ab jetzt wird eine spezielle Art von Ringen mit spezifischen Eigenschaften, nämlich Polynomringe in endlich vielen Veränderlichen, für den weiteren Verlauf wichtig sein. Die noethersche Eigenschaft wird durch das Hilbertsche Basistheorem gewährleistet (vgl. 1.20). Doch zunächst werden an dieser Stelle die maßgeblichen Objekte vorgestellt.

Definition 1.8. Ein *Monom* in den Veränderlichen x_1, \dots, x_n ist ein Produkt der Form

$$x^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n}$$

für nicht-negative Zahlen $\alpha_1, \dots, \alpha_n$ als auftretende Exponenten. Für $\alpha = (0, \dots, 0)$ soll gelten $x^\alpha = 1$.

Die Menge aller Monome, die aus den Veränderlichen x_i gebildet werden können, wird mit

$$Mon_n := Mon(x_1, \dots, x_n) := \{x^\alpha \mid \alpha \in \mathbb{N}^n\}$$

bezeichnet.

Die multiplikative Halbgruppe Mon_n läßt sich offensichtlich bijektiv mit der additiven Halbgruppe \mathbb{N}^n identifizieren und somit $Mon_n = \mathbb{N}^n$. Ein Polynom ist nun eine entsprechende Linearkombination von Monomen:

Definition 1.9. Ein *endliche* Linearkombination f von Monomen in Mon_n und Koeffizienten in einem Ring R

$$f = \sum_{\alpha} r_{\alpha} \cdot x^{\alpha} \quad r_{\alpha} \in R$$

heißt Polynom in den Veränderlichen x_1, \dots, x_n .

Die Menge aller Polynome in x_1, \dots, x_n mit Koeffizienten in R heißt *Polynomring* und wird mit $R[x_1, \dots, x_n]$ bezeichnet.

Bemerkung 1.10. Die in 1.9 als *Polynomring* definierte Menge aller Polynome ist offensichtlich ein Ring.

Die bisher genannten Begriffe werden durch ein sehr einfaches Beispiel erläutert:

Beispiel 1.11. Sei $A = K[x, y]$ ein Polynomring in x, y sowie $M = \{x^2, y^3\}$ eine Menge von Polynomen. Dann gilt für das von M erzeugte Ideal $\langle M \rangle_A$

$$\langle M \rangle_A = \{f \cdot x^2 + g \cdot y^3 \mid f, g \in A\}$$

und somit für das Polynom $f = x^6y^6 - 5x^5y + x^2y + 7xy^6 - xy^4 + xy^3$:

$$f = x^2(x^4y^6 - 5x^3y + y) + y^3(7xy^3 - xy + x)$$

und daher folgt $f \in \langle M \rangle_A$.

Mithilfe von expliziten Erzeugern läßt sich also die Zugehörigkeit eines beliebigen Polynoms zu einem Ideal feststellen. Für eine Automatisierung dieses Vorgehens mit Eindeigkeitseigenschaften werden allerdings die noch weitergehenden Konzepte der Gröbnerbasis und der Normalform benötigt, welche wiederum den Begriff der monomialen Ordnung in 1.12 zugrunde legen:

Definition 1.12. Ordnungen

1. Eine *monomiale Ordnung* ist $<$ eine totale Ordnung auf $Mon_n = \mathbb{N}^n$ mit der Eigenschaft

$$x^\alpha < x^\beta \quad \Rightarrow \quad x^\gamma \cdot x^\alpha < x^\gamma \cdot x^\beta \quad \forall \alpha, \beta, \gamma \in \mathbb{N}^n$$

2. Eine *Wohlordnung* ist eine totale Ordnung $<$ auf einer Menge, sodaß jede Menge, die nicht die leere Menge ist, ein bezüglich der Ordnung $<$ minimales Element besitzt.

Dies erlaubt es, ein Polynom $f \in$ in eindeutiger Reihenfolge zu repräsentieren:

$$f = a_\alpha \cdot x^\alpha + a_\beta \cdot x^\beta + \dots + a_\gamma \cdot x^\gamma$$

mit einer hinsichtlich der Ordnung abfallenden Reihenfolge der Monome

$$x^\alpha > x^\beta > \dots > x^\gamma$$

und daher lassen sich die folgenden Begriffe definieren:

Kapitel 1 Computeralgebra

Definition 1.13. Sei $f = \sum_{\alpha} a_{\alpha} \cdot x^{\alpha}$ mit obiger Eigenschaft. Dann heißt bezüglich einer monomialen Ordnung $<$

$lt(f) := a_{\alpha} \cdot x^{\alpha}$ der *Leitterm* von f

$lm(f) := x^{\alpha}$ das *Leitmonom* von f

$le(f) := \alpha$ der *Leitexponent* von f

$lc(f) := a_{\alpha}$ der *Leitkoeffizient* von f

$tail(f) := f - lt(f) = a_{\beta} \cdot x^{\beta} + \dots + a_{\gamma} \cdot x^{\gamma}$ der *Schwanz* von f

Für die Theorie ist die Unterscheidung von lokalen und globalen Ordnungen besonders relevant:

Definition 1.14. Es sei $<$ eine Ordnung auf \mathbb{N}^n , dann gilt:

1. $<$ heißt *global*, falls

$$x^{\alpha} > 1 \quad \forall \alpha \neq (0, \dots, 0)$$

2. $<$ heißt *lokal*, falls

$$x^{\alpha} < 1 \quad \forall \alpha \neq (0, \dots, 0)$$

3. $<$ heißt *gemischt*, wenn sie weder lokal noch global ist.

Eine teilweise Charakterisierung von Ordnungen nach 1.14 kann mittels der folgenden Äquivalenzen vorgenommen werden:

Lemma 1.15.

Es sind für eine Ordnung $<$ äquivalent:

1. $<$ ist eine Wohlordnung

2. $x_i > 1 \quad \forall i \in \{1, \dots, n\}$

3. $x^{\alpha} > 1 \quad \forall \alpha \in \mathbb{N}^n \setminus \{0\}$

4. $(\alpha \geq_{\text{nat}} \beta \wedge \alpha \neq \beta) \Rightarrow x^\alpha > x^\beta$
 und somit ist $<$ eine Verfeinerung der natürlichen Ordnung auf \mathbb{N}^n

Beweis. Für den Beweis siehe [GP08]. □

Einige sehr bekannte Ordnungen sind in der folgenden Definition erklärt:

Definition 1.16. Für Monome x^α und x^β in Mon_n heißt eine Ordnung $<$

1. *lexikographisch* oder $<_{l_p}$, wenn

$$x^\alpha >_{l_p} x^\beta : \Leftrightarrow \exists i \in \{1, \dots, n\} : \\ \alpha_1 = \beta_1, \dots, \alpha_{i-1} = \beta_{i-1}, \alpha_i > \beta_i$$

2. *graduiert lexikographisch* oder $<_{D_p}$, wenn

$$x^\alpha >_{D_p} x^\beta : \Leftrightarrow \sum_i \alpha_i > \sum_i \beta_i \\ \vee \left(\sum_i \alpha_i = \sum_i \beta_i \wedge \exists i \in \{1, \dots, n\} : \right. \\ \left. \alpha_1 = \beta_1, \dots, \alpha_{i-1} = \beta_{i-1}, \alpha_i > \beta_i \right)$$

3. *graduiert umgekehrt lexikographisch* oder $<_{d_p}$, wenn

$$x^\alpha >_{d_p} x^\beta : \Leftrightarrow \sum_i \alpha_i > \sum_i \beta_i \\ \vee \left(\sum_i \alpha_i = \sum_i \beta_i \wedge \exists i \in \{1, \dots, n\} : \right. \\ \left. \alpha_n = \beta_n, \dots, \alpha_{i+1} = \beta_{i+1}, \alpha_i < \beta_i \right)$$

Mithilfe von Lemma 1.15 sieht man, daß es sich bei allen drei Ordnungen um globale Ordnungen handelt. Es wird im weiteren Verlauf die Annahme getroffen, daß auftretende Ordnungen immer globale Ordnungen sind.

An dieser Stelle werden Beispiele zur Illustration obiger Ordnungen und Erklärung von 1.12 und 1.13 präsentiert:

Kapitel 1 Computeralgebra

Beispiel 1.17. *Es sei ein Polynom*

$$f = x^4z - 4x^2y^2z^2 + 2xy^4z - 13 \in \mathbb{R}[x_1, x_2, x_3] = \mathbb{R}[x, y, z]$$

dann gilt für die jeweilige Schreibweise in der Reihenfolge der auftretenden Monome mit Blick auf die einzelnen Ordnungen:

$$\langle_{lp} : f = x^4z - 4x^2y^2z^2 + 2xy^4z - 13 \quad \Longrightarrow \quad lm(f) = x^4z$$

$$\langle_{Dp} : f = -4x^2y^2z^2 + 2xy^4z + x^4z - 13 \quad \Longrightarrow \quad lm(f) = x^2y^2z^2$$

$$\langle_{dp} : f = 2xy^4z - 4x^2y^2z^2 + x^4z - 13 \quad \Longrightarrow \quad lm(f) = xy^4z$$

In einem Polynomring mit einer Ordnung \langle erzeugen die Leitmonome von Polynomen einer beliebigen Menge das sogenannte Leitideal:

Definition 1.18. Sei $M \subseteq K[x_1, \dots, x_n]$ eine beliebige Menge. Dann heißt

$$L(M) := \langle lm(f) \mid f \in M \setminus \{0\} \rangle_{K[x_1, \dots, x_n]}$$

das von M erzeugte *Leitideal*.

Es muß in dieser Hinsicht beachtet werden, daß $L(M)$ das Ideal ist, welches von den Leitmonomen in M und nicht in $\langle M \rangle$ erzeugt wird. Für den Fall, daß M ein Ideal ist, sind beide Ideale identisch.

Ein einfaches Beispiel für ein Leitideal ist

Beispiel 1.19. Sei $G = \{x^2 - y, xy - y^2\}$ für die lexikographische Ordnung \langle_{lp} , wobei $y < x$ gelten soll. Dann ist

$$L(G) = \langle x^2, xy \rangle_{K[x,y]} = \{f \cdot x^2 + g \cdot xy \mid f, g \in K[x, y]\}$$

Der Polynomring $A[x]$ in einer Variable x erbt eine sehr angenehme Endlichkeitsbedingung von seinem Grundring A . Dies ist die Aussage des folgenden Theorems:

Theorem 1.20. Hilbertsches Basistheorem

Wenn der Ring A noethersch ist, dann ist auch $A[x]$ noethersch.

Beweis. Sei I ein Ideal in $A[x]$. Für ein beliebiges $f \in A[x]$ mit $f = c_n x^n + \dots + c_0 x^0$ wird der Leitterm (bezüglich der Ordnung $<_{lp}$) als $c_n x^n$ und der Leitkoeffizient mit c_n bezeichnet. Betrachte ein beliebiges Element $f_1 \in I \setminus \{0\}$ mit minimalem Grad $d_1 := \deg(f_1)$. Dann suche für den Fall, daß $\langle f_1, \dots, f_i \rangle_A \subsetneq I$, iterativ Elemente $f_{i+1} \in I \setminus \langle f_1, \dots, f_i \rangle_A$ mit minimalem Grad $d_{i+1} = \deg(f_{i+1})$ und erhalte für $\langle f_1, \dots, f_i \rangle_A \subsetneq I$ eine Folge von Idealen $(\langle f_1, \dots, f_i \rangle_A)_i$ mit $\langle f_1, \dots, f_i \rangle_A \subsetneq \langle f_1, \dots, f_{i+1} \rangle_A$. Das von den Leitkoeffizienten c_j der f_j erzeugte Ideal $C := \langle c_j \mid j \in \mathbb{N}^+ \rangle_A \trianglelefteq A$ ist nach Voraussetzung endlich erzeugt. Wähle ein minimales k mit $C = \langle c_1, \dots, c_k \rangle_A$.

Behauptung: $I = \langle f_1, \dots, f_k \rangle_A$ für diesen Index k .

Wenn dies nicht der Fall ist, Wähle wie oben f_{k+1} und schreibe für den Koeffizienten $c_{k+1} = \sum_{j=1}^k a_j c_j$. Dann muß gelten, daß $\forall j \in \{1, \dots, k\} : d_{k+1} \geq d_j$. Konstruiere ein Polynom p vom Grad d_{k+1} : $p = \sum_{j=1}^k a_j f_j x^{d_{k+1}-d_j} \in \langle f_1, \dots, f_k \rangle_A$. Dann folgt, daß $f_{k+1} - p \in I \setminus \langle f_1, \dots, f_k \rangle_A$ und $\deg(f_{k+1} - p) < d_{k+1}$. Dieser Widerspruch beweist die Behauptung. \square

Das Theorem 1.20 ist grundlegend in der kommutativen Algebra und stellt zusammen mit den nachfolgenden Korollaren sicher, das man es bei Polynomringen über noetherschen Ringen in endlich vielen Variablen mit endlich erzeugten Objekten zu tun hat. Dies ist in der praktischen Handhabung ebendieser essentiell.

Mittels Induktion folgt aus 1.20 das Korollar:

Korollar 1.21. Für einen noetherschen Ring A ist der Polynomring in n Variablen $A[x_1, \dots, x_n]$ noethersch. \square

Die Theorie läßt sich in weiten Teilen auch für Polynomringe über beliebigen Ringen R formulieren, jedoch wird sich nachfolgend auf den Fall, daß $R = K$ ein Körper ist, beschränkt. Durch diese Körpereigenschaft ergeben sich zusätzliche angenehme Eigenschaften für den resultierenden Polynomring:

Korollar 1.22.

1. Für einen Körper K ist der Polynomring $A = K[x_1, \dots, x_n]$ in endlich vielen Variablen ein noetherscher Ring und damit

$$\forall I \trianglelefteq A \quad \exists f_1, \dots, f_k \in A : \quad I = \langle f_1, \dots, f_k \rangle_A$$

2. Für $n = 1$, also den Fall $A = K[x]$ in einer Variable x , gilt sogar, daß ein Hauptidealbereich vorliegt. Letzteres bedeutet

$$\forall I \trianglelefteq A \quad \exists f \in A : \quad I = \langle f \rangle_A$$

und somit ist jedes Ideal von einem Element erzeugt.

Beweis.

1. Dies folgt einfach aus 1.20 und 1.21, da ein Körper K als Ring stets nur die beiden zwei Ideale $\langle 0 \rangle$ und K enthält und somit noethersch ist.
2. Wähle für ein Ideal I ein Element minimalen Grades (dieses existiert, da die Ordnung nach auftretendem Grad in $K[x]$ eine Wohlordnung ist). Dann folgt aus den Eigenschaften des Euklidischen Divisionsalgorithmus, daß

$$\forall g \in I : \quad f|g$$

und somit, daß $I = \langle f \rangle_A$. Für dieses letzte Argument ist wichtig, daß K ein Körper ist.

□

Bemerkung 1.23. Es sei an dieser Stelle noch einmal erwähnt, daß im folgenden nur globale Ordnungen betrachtet werden. Im Falle von zusätzlich erlaubten lokalen Ordnungen ist die Theorie mit einigen Modifikationen versehen, sowohl hinsichtlich der Begriffswahl als auch der Beweisführung. Dieser allgemeinere Fall ist eher aus algebraischer und theoretischer Hinsicht, jedoch weniger für die Zwecke der Anwendung in symbolischen Ansätzen zur Lösung von Differenzenschemata interessant. Daher besteht in der Beschränkung auf globale Ordnungen keine wirkliche Einschränkung im praktischen Sinne. Für den allgemeinen Fall siehe [GP08].

1.3 Gröbnerbasen

1.3.1 Gröbnerbasen für Ringe

Eine Gröbnerbasis G für ein Ideal I zu einer Ordnung $<$ ist eine endliche Menge von Polynomen, deren Leitmonome das Leitideal dieses Ideals erzeugen. Das Hilbertsche Basistheorem garantiert die endliche Erzeugung des Leitideals (und damit die Existenz einer Gröbnerbasis) sowie die Terminierung des Buchbergerschen Algorithmus 1.37 zur Berechnung einer Gröbnerbasis.

Da ein Leitideal nur von Monomen erzeugt wird, ergeben sich hieraus Informationen über das Ideal selbst.

Definition 1.24. Sei $I \trianglelefteq K[x_1, \dots, x_n]$ ein Ideal, $G \subseteq K[x_1, \dots, x_n]$ eine endliche Menge. Es heißt G eine *Gröbnerbasis* von I wenn gilt

$$G \subset I \quad \wedge \quad L(I) = L(G)$$

wobei letztere Bedingung nichts anderes bedeutet als

$$\forall f \in I \quad \exists g \in G \quad lm(g) \mid lm(f)$$

Man kann für eine monomiale Ordnung in A einzelne Polynome und Mengen von Polynomen nach gewissen Eigenschaften charakterisieren:

Definition 1.25. Es seien $f \in A$ ein Polynom und $M \subset A$ eine beliebige Menge von Polynomen.

1. M heißt *interreduziert*, wenn gilt für $f, g \in M$ gilt:

$$0 \notin M \quad \wedge \quad (f \neq g \Rightarrow lm(f) \nmid lm(g))$$

2. f heißt *reduziert* bezüglich der Menge M , wenn kein Monom in der Potenzreihenentwicklung von f in $L(G)$ enthalten ist.
3. M heißt *reduziert*, wenn M interreduziert ist und für alle $h \in M$ gilt, daß $lc(h) = 1$ und $tail(h)$ reduziert bezüglich M ist.

Eine Verallgemeinerung von Polynomdivision mit Rest stellt der Begriff der Normalform dar. Es ist mit einer errechneten Normalform (dem Divisionsrest) möglich, über die Zugehörigkeit eines Polynoms zu einem Ideal zu entscheiden, wenn gewisse Kriterien erfüllt sind und eine Gröbnerbasis vorliegt (siehe Lemma 1.35).

Definition 1.26. Für einen Polynomring $A = K[x_1, \dots, x_n]$ sei

$$\mathcal{G} = \{ G \mid G \subset A \text{ eine endliche Menge} \}$$

Dann ist eine *Normalform* eine Abbildung

$$NF : A \times \mathcal{G} \longrightarrow A$$

$$(f, G) \longmapsto NF(f \mid G)$$

mit der Eigenschaft, daß $\forall f \in A$ und $\forall G \in \mathcal{G}$ gilt:

1. $NF(0 \mid G) = 0$
2. $NF(f \mid G) \neq 0 \Rightarrow \text{lm}(NF(f \mid G)) \notin L(G)$
3. Für $G = (g_1, \dots, g_s)$ hat $f - NF(f \mid G)$ eine *Standardrepräsentation* bezüglich G , das heißt $f - NF(f \mid G) = 0$ oder

$$f - NF(f \mid G) = \sum_{k=1}^s a_k g_k$$

mit $a_k \in A$ und $s \geq 0$, sodaß

$$\text{lm}\left(\sum_{k=1}^s a_k g_k\right) \geq \text{lm}(a_k g_k) \quad \forall k \text{ mit } a_k g_k \neq 0$$

Die Eigenschaft der Standardrepräsentation bedeutet, daß sich die Leitertme der $a_k g_k$ mit $\text{lm}(a_k g_k) > \text{lm}(f - NF(f \mid G))$ nicht gegenseitig aufheben und somit

$$\exists k : \text{lm}(a_k g_k) = \text{lm}(f - NF(f \mid G))$$

gilt.

Mit Blick auf Definition 1.25 definiert man für eine Normalform NF mit obigen Bezeichnungen:

Definition 1.27. NF ist eine *reduzierte* Normalform, wenn gilt

$$\forall f : \quad NF(f | G) \text{ ist reduziert bezüglich } G$$

Vor der Angabe eines Beispiels für eine Normalform wird das *spoly*, eine Konstruktion zur Eliminierung von Leittermen, eingeführt:

Definition 1.28. Seien $f, g \in A \setminus \{0\}$ mit $lm(f) = x^\alpha$ sowie $lm(g) = x^\beta$. Desweiteren sei $\gamma := (\max(\alpha_1, \beta_1), \dots, \max(\alpha_n, \beta_n))$ und $lcm(x^\alpha, x^\beta) := x^\gamma$. Man definiert

$$spoly(f, g) := x^{\gamma-\alpha} f - \frac{lc(f)}{lc(g)} x^{\gamma-\beta} g$$

als das *spoly* von f und g .

Das nachstehende Beispiel zeigt die Berechnung des *spoly* für zwei Polynome:

Beispiel 1.29. Sei A mit der Ordnung lp und $x > y > z$, sowie zwei Polynome $f = x^3 - zy$ und $g = x^2y - z^2$, dann ist

$$lcm(lm(f), lm(g)) = x^3y$$

und deshalb folgt für das *spoly*:

$$spoly(f, g) = y(x^3 - zy) - x(x^2y - z^2) = xz^2 - y^2z$$

Es ist offensichtlich, daß das *spoly* in einer besonderen Situation eine sehr nützliche Eigenschaft besitzt:

Kapitel 1 Computeralgebra

Bemerkung 1.30. Für den Fall $lm(g) \mid lm(f)$ kann man das spoly schreiben als

$$spoly(f, g) = f - \frac{lc(f)}{lc(g)} x^{\alpha-\beta} g$$

und daher gilt

$$lm(spoly(f, g)) < lm(f)$$

Diese Beobachtung führt zu einer sehr bekannten Normalform, dem Algorithmus NFBuchberger zur Berechnung einer Normalform, der in Pseudocode präsentiert wird:

Algorithmus 1.31. NFBuchberger

Input: $f \in K[x_1, \dots, x_n]$, $G \in \mathcal{G}$

Output: $h \in K[x_1, \dots, x_n]$ eine Normalform bezüglich G

$h := f$;

while ($h \neq 0$ und $G_h := \{g \in G \mid lm(g) \mid lm(h)\} \neq \emptyset$)

 wähle ein beliebiges $g \in G_h$;

$h := spoly(h, g)$;

return (h);

Die Terminierung des Algorithmus folgt, wenn man für einen beliebigen Iterationsschritt mit Bemerkung 1.30 die Richtigkeit von

$$lm(spoly(h_k, g)) < lm(h_k)$$

aufgrund der spoly-Reduktion durch ein $g \in G$ mit

$$spoly(h_k, g) = h_{k-1} - \frac{lc(h_{k-1})}{lc(g)} m_k g$$

für ein geeignetes Monom m_k einsieht. Da aufgrund des Lemmas 1.15 mit einer globalen Ordnung $<$ eine Wohlordnung vorliegt, wird die Folge der $(h_k)_k$ mit $lm(h_k) < lm(h_{k-1})$ ab einem geeigneten Index stationär und die Eigenschaften einer Normalform sind durch die Konstruktion der Reduktionsschritte trivialerweise erfüllt.

Bemerkung 1.32. Man kann den oben präsentierten Normalform-Algorithmus zu einer reduzierten Normalform modifizieren, indem man in jedem Iterationsschritt zusätzlich noch den Schwanz des Polynoms h reduziert.

Ein einfaches Beispiel in SINGULAR zeigt die Reduzierung eines Polynoms bezüglich einer Menge mithilfe einer Normalform und verdeutlicht auch die Grenzen dieses Konzepts zur Feststellung von Idealmitgliedschaften:

Beispiel 1.33. Normalform (Teil 1)

Es sei $A = \mathbb{Q}[x, y, z]$ mit der Ordnung $<_{lp}$. Betrachte die Polynome

$$f = x^3 - yz \quad h = xy^2 + z \quad g = x^2y - z^2$$

und berechne mithilfe des Befehls `reduce` die Normalformen der jeweiligen `spolys`, die sich aus f, g und h bilden lassen:

```
> LIB "teachstd.lib";
// ** loaded /usr/share/Singular/LIB/teachstd.lib (1.13,2009/04/15)
> ring A=0,(x,y,z),lp;
> poly f=x^3-z*y; poly g=x^2*y-z^2; poly h=x*y^2+z;
> ideal I=f,g,h;
> reduce(spoly(f,g),I);
// ** I is no standard basis
xz2-y2z
> reduce(spoly(h,g),I);
// ** I is no standard basis
xz+yz2
> reduce(spoly(f,h),I);
// ** I is no standard basis
-x2z-y3z
```

Das Fazit aus obigen Berechnungen ist, daß man Normalformen (das heißt: Divisionsreste) erhält, die sämtlich von Null verschieden sind, obwohl die `spolys` als Linearkombinationen von Idealelementen kanonisch ebenfalls Elemente des Ideals $\langle f, g, h \rangle_A$ sind.

Kapitel 1 Computeralgebra

Das nächste Beispiel zeigt darüberhinaus, daß eine Normalform keinesfalls eindeutig sein muß:

Beispiel 1.34. *Es sei wie in Beispiel 1.33 $A = \mathbb{Q}[x, y, z]$ mit der lexikographischen Ordnung $<_{lp}$. Diesmal sei das Polynom $f = 2x^2y + 3xz$ bezüglich des von $g_1 = xy + z$ und $g_2 = xy + y^2$ aufgespannten Ideals $\langle xy + z, xy + y^2 \rangle_A$ zu reduzieren. In SINGULAR berechnet sich dies wie folgt:*

```
> ring r=0,(x,y,z),lp;
> poly f=2*x^2*y + 3*x*z;
> poly g1=x*y +z;
> poly g2=x*y+y^2;
> ideal I=g1,g2;
> reduce(f,I);
// ** I is no standard basis
3xz+2y3
```

Allerdings ergibt sich nach Vertauschung der Erzeuger g_1 und g_2 :

```
> ideal J=g2,g1;
> reduce(f,J);
// ** J is no standard basis
xz
```

Es ist offensichtlich die Normalform von der Reihenfolge der Idealerzeuger abhängig.

Beispiele 1.33 und 1.34 zeigen, daß für ein beliebiges Polynom f in einem Polynomring A die Normalform $NF(f | M)$ nur begrenzte Aussagefähigkeit über die Mitgliedschaft von f in dem von M erzeugten Ideal $\langle M \rangle_A$ hat, wenn die Normalform bezüglich einer beliebigen Menge M ohne weitergehende Eigenschaften berechnet wird. Es wird im folgenden deutlich, daß das Vorliegen der Menge M als Gröbnerbasis der Normalform zur gewünschten Aussagekraft über die Idealmitgliedschaft verhilft und auch Eindeutigkeit gewährleistet ist:

Lemma 1.35. *Sei $I \trianglelefteq A = K[x_1, \dots, x_n]$ ein Ideal, $G \subset I$ eine Gröbnerbasis von I und $NF(\dots | G)$ eine Normalform bezüglich G . Dann gelten folgende Aussagen:*

1. Für ein Polynom $f \in A$ erhält man

$$f \in I \iff NF(f | G) = 0$$

2. Das Ideal I wird als A -Ideal von G erzeugt

$$I = \langle G \rangle_A$$

3. Wenn $NF(\dots | G)$ eine reduzierte Normalform bezüglich G ist, dann ist sie eindeutig.

Beweis.

1. Wenn $NF(f | G) = 0$, dann gilt aufgrund der Standardrepräsentation und $G \subset I$, daß $f \in I$. Wenn $NF(f | G) \neq 0$, dann muß $lm(NF(f | G)) \notin L(G) = L(I)$ (da G eine Gröbnerbasis ist) und somit gilt $NF(f | G) \notin I$. Daraus folgt $f \notin I$.

2. Dies ist eine Folgerung aus 1.

3. Es seien zwei reduzierte Normalformen h_1 und h_2 von f bezüglich G . Dann gilt $h_1 - h_2 = (f - h_2) - (f - h_1) \in \langle G \rangle_A$. Wenn nun $h_1 - h_2 \neq 0$, dann muß $lm(h_1 - h_2) \in L(G)$. Dies steht im Widerspruch zur Annahme, daß h_1 und h_2 reduziert sind (vgl. 1.27). \square

Eine der wichtigsten Charakterisierungen von Gröbnerbasen ist das Buchberger Kriterium:

Theorem 1.36. Buchberger Kriterium

Für eine Menge $G = \{g_1, \dots, g_k\} \subset I$ in einem Ideal $I \trianglelefteq A$ und eine eine Normalform $NF(\dots | G)$ bezüglich G sind die beiden folgenden Aussagen äquivalent:

1. G ist eine Gröbnerbasis von I
2. I wird von G erzeugt und

$$NF(\text{spoly}(g_i, g_j) | G) = 0 \quad \forall i, j = 1, \dots, k$$

Beweis. Die Implikation (1. \Rightarrow 2.) folgt aus Lemma 1.35.

Für die Umkehrung wird auf die entsprechenden Abschnitte in [GP08] verwiesen. \square

Nachdem durch die bisherigen Aussagen deutlich wurde, daß es insbesondere in den praktischen Berechnungen sehr angenehm ist, bezüglich einer monomialen Ordnung eine Gröbnerbasis für ein Ideal zu kennen, stellt sich nun die Frage der konkreten Berechnung einer solchen.

Die Äquivalenz in der Aussage des Buchberger Kriteriums wird bei der Berechnung einer Gröbnerbasis im Algorithmus 1.37 verwendet:

Algorithmus 1.37. Gröbnerbasen-Algorithmus (Buchberger)

Input: $G \in \mathcal{G}$, NF eine Normalform, $<$ eine Ordnung

Output: $S \subset A$ eine Gröbnerbasis von $\langle G \rangle_A$ bezüglich $<$

$S := G$;

$P := \{(f, g) \mid f, g \in S, f \neq g\}$; //die Paarmenge

while ($P \neq \emptyset$)

wähle ein beliebiges $(f, g) \in P$;

$P := P \setminus \{(f, g)\}$;

$h := NF(\text{spoly}(f, g) \mid S)$;

if ($h \neq 0$)

$P := P \cup \{(h, f) \mid f \in S\}$;

$S := S \cup \{h\}$;

return (S);

Lemma 1.38. *Algorithmus 1.37 ergibt eine Gröbnerbasis und terminiert.*

Beweis. Die Terminierung folgt, da in der Iterationsschleife für $h \neq 0$ und folglich $lm(h) \notin L(S)$ durch die Hinzunahme von h in die Menge S eine aufsteigende Folge von Idealen $(L(S_i))_i$ mit $L(S_i) \subset L(S_{i+1})$ entsteht. Durch die noethersche Eigenschaft von $A = K[x_1, \dots, x_n]$ wird diese zu gegebenem Zeitpunkt stationär und für alle Paare von Polynomen f, g gilt ab diesem Zeitpunkt $NF(\text{spoly}(f, g) \mid S) = 0$. Damit terminiert der Algorithmus nach Entleerung der Paarmenge P .

Die Korrektheit ist eine Folge von Theorem 1.36. □

Bemerkung 1.39. Die Berechnung einer Gröbnerbasis nach dem ursprünglichen Algorithmus ist von einer doppelt exponentiellen Komplexität hinsichtlich der Anzahl der auftretenden Variablen (vergleiche [Buc83, JS06]).

Mit der Berechnung einer Gröbnerbasis läßt sich das Beispiel 1.33 zu einem befriedigenden Ergebnis führen:

Beispiel 1.40. Normalform (Teil 2)

(Fortsetzung von Beispiel 1.33)

Nach Berechnung einer Gröbnerbasis des Ideals $I = \langle f, g, h \rangle$

```
> I;//das ursprüngliche Ideal I
I[1]=x3-yz
I[2]=x2y-z2
I[3]=xy2+z
```

ergibt sich mithilfe des Befehls *std* in SINGULAR für die Reduktion der *spolys*:

```
> def J=std(I);J;//die Gröbnerbasis von I
J[1]=z8+z2
J[2]=yz2+z4
J[3]=y2z+yz3
J[4]=xz+yz2
J[5]=xy2+z
J[6]=x2y-z2
J[7]=x3-yz
> reduce(spoly(f,g),J);
0
> reduce(spoly(h,g),J);
0
> reduce(spoly(f,h),J);
0
```

Durch die Hinzunahme der zusätzlichen Elemente in die Menge, bezüglich welcher die Reduktionen vorgenommen werden, läßt sich nunmehr die Idealmitgliedschaft der *spolys* verifizieren.

Eine wichtige Aufgabe im Zusammenhang mit Gröbnerbasen ist die Eliminierung von Variablen. Zu einem Ideal $I \subseteq K[x_1, \dots, x_n]$ mit gegebenen Erzeugern f_1, \dots, f_l möchte man Erzeuger zum k -ten *Eliminationsideal* des Ideals I

$$I_k := I \cap K[x_{k+1}, \dots, x_n]$$

Kapitel 1 Computeralgebra

für ein $k < n$ errechnen. Der Name Elimination bedeutet das Eliminieren von Erzeugern in I , welche von x_1, \dots, x_k abhängig sind.

Es wird nun eine Eigenschaft von Ordnungen erklärt, die für die Eliminierung wichtig ist:

Definition 1.41. Eine monomiale Ordnung $<$ in einem Polynomring $A = K[x_1, \dots, x_k, x_{k+1}, \dots, x_n]$ heißt *Eliminationsordnung* für x_1, \dots, x_k , wenn sie die *Eliminierungseigenschaft* für die ersten k Variablen besitzt, das heißt

$$lm(g) \in K[x_{k+1}, \dots, x_n] \implies g \in K[x_{k+1}, \dots, x_n]$$

für ein beliebiges Polynom $g \in A$.

Die Eliminierungseigenschaft bedeutet für ein Polynom g bei nicht Auftreten einer Menge von Variablen im Leitmonom die Mitgliedschaft von g in einem Unterring, der diese Variablen nicht enthält.

Beispiele für Eliminationsordnungen sind $<_{lp}$ oder auch *Block-* oder *Produktordnungen*, das heißt Ordnungen $<= (<_1, <_2)$ auf $Mon(x_1, \dots, x_k, x_{k+1}, \dots, x_n) = \{(\alpha, \beta) \mid \alpha \in \mathbb{N}^k, \beta \in \mathbb{N}^{n-k}\}$, sodaß $<_1$ und $<_2$ global sind, mit

$$\begin{aligned} x^{\alpha_1} x^{\beta_1} > x^{\alpha_2} x^{\beta_2} & \iff x^{\alpha_1} >_1 x^{\alpha_2} \\ & \vee (x^{\alpha_1} = x^{\alpha_2} \wedge x^{\beta_1} >_2 x^{\beta_2}) \end{aligned}$$

Damit ist die Blockordnung $<= (<_{dp}, <_{dp})$ eine Eliminationsordnung, die ebenfalls global ist.

Als nächstes Lemma ist das Zusammenspiel von Gröbnerbasen und dem Eliminationsideal formuliert:

Lemma 1.42. *Eliminationslemma für Ringe*

Es sei $<$ eine Eliminationsordnung für die ersten k Variablen x_1, \dots, x_k im Ring $K[x_1, \dots, x_k, x_{k+1}, \dots, x_n]$ und G eine Gröbnerbasis des Ideals $I \trianglelefteq K[x_1, \dots, x_n]$. Dann ist die Menge

$$\tilde{G} := \{g \in G \mid lm(g) \in K[x_{k+1}, \dots, x_n]\} \subset K[x_{k+1}, \dots, x_n]$$

eine Gröbnerbasis von

$$\tilde{I} := I \cap K[x_{k+1}, \dots, x_n].$$

Beweis. Sei $h \in \tilde{I}$ ein beliebiges Polynom, dann $\exists g \in G$ mit $lm(g) | lm(h)$, weil G eine Gröbnerbasis von I ist, und ebenso gilt wegen der Eliminationseigenschaft von $<$, daß $g \in \tilde{G}$. \square

1.3.2 Gröbnerbasen für Moduln

Das Konzept der Gröbnerbasen läßt sich sehr leicht auf den Fall eines Moduls übertragen. Hier werden nun die entsprechenden Begriffe definiert:

Definition 1.43. Ein *Modul* über einem Ring A ist eine Menge M mit zwei Verknüpfungen, der Addition in M

$$+ : M \times M \longrightarrow M$$

und der Wirkung des Ringes A auf M

$$\cdot : A \times M \longrightarrow M$$

die folgenden Gesetzen für alle $a_1, a_2 \in A$ und $m_1, m_2 \in M$ genügen:

$(M, +, 0)$ ist eine abelsche (kommutative) Gruppe

$a_1 \cdot (a_2 \cdot m_1) = (a_1 \cdot a_2) \cdot m_1$ (Assoziativität)

$a_1 \cdot (m_1 + m_2) = a_1 \cdot m_1 + a_1 \cdot m_2$

$(a_1 + a_2) \cdot m_1 = a_1 \cdot m_1 + a_2 \cdot m_1$ (Distributivität)

$1_A \cdot m_1 = m_1$ (Identität)

Trivialerweise ist ein Ring immer ein Modul über sich selbst und die Ideale sind die entsprechenden Untermoduln. Für den angenehmen Fall, daß A ein Körper ist, ist ein Modul ein Vektorraum.

Definition 1.44. Ein *A-Modulhomomorphismus* $\phi : N \rightarrow M$ ist eine Abbildung zwischen zwei A -Moduln N und M , sodaß

1. $\phi(an_1) = a\phi(n_1)$

2. $\phi(n_1 + n_2) = \phi(n_1) + \phi(n_2)$

für alle $n_1, n_2 \in N$ und alle $a \in A$ gelten.

Wenn ϕ eine Bijektion darstellt, heißt die Abbildung *Isomorphismus* und man schreibt $N \cong M$.

Interessant sind wie bei Ringen vor allem endlich erzeugte Moduln:

Definition 1.45. Ein Modul M heißt endlich erzeugt, wenn jeder Untermodul von M endlich erzeugt ist.

Die noethersche Eigenschaft wird wie bei Ringen in 1.6 definiert, nur wird der Begriff des Ideals gegen den des Untermoduls ausgetauscht:

Definition 1.46. Ein Modul M heißt *noethersch*, wenn jede aufsteigende Kette $(M_l)_{l \in \mathbb{N}}$ von Untermoduln in M

$$M_0 \subseteq M_1 \subseteq \dots \subseteq M_k \subseteq \dots \subseteq M$$

endlich ist.

Wie bei Ringen ist auch im Modulfall die Eigenschaft in 1.46 äquivalent zur endlichen Erzeugung:

Lemma 1.47. *Ein Modul M ist noethersch, genau dann, wenn M endlich erzeugt ist.*

Beweis. Die Argumentation ist mit der im Falle von Ringen identisch. \square

Von besonderem Interesse sind freie Moduln und insbesondere direkte Summen endlich vieler Moduln über demselben Ring:

Definition 1.48. Für zwei Moduln M und N über dem Ring A definiere die *direkte Summe* mittels

$$N \oplus M := \{(n, m) \in N \times M\}$$

mit der vererbten Modulstruktur via $a(n, m) = (an, am)$ und komponentenweiser Addition. Für die direkte Summe von endlich vielen A -Moduln M_i

schreibe $\bigoplus_{i=1}^n M_i$. Mit einem Ring A definiert man den endlich erzeugten Modul A^n entsprechend als

$$A^n := \bigoplus_{i=1}^n A \cdot e_i$$

Dabei sind für alle $i = 1, \dots, n$ die i -ten Einheitsvektoren e_i kanonisch definiert

$$e_i := (0, \dots, \underbrace{1}_i, \dots, 0)$$

Die Restklassenbildung eines Moduls M nach einem Untermodul $N \subset M$ ist analog zu der bei Ringen und Idealen und daher ergibt sich

$$M/N := \{m + N \mid m \in M\}$$

mit den natürlichen Verknüpfungen.

Eine wichtige Charakterisierung von Moduln liefert das nächste Theorem:

Theorem 1.49. *Ein endlich erzeugter Modul M über einem noetherschen Ring A ist ein noetherscher Modul.*

Der Beweis folgt induktiv aus den beiden Aussagen des folgenden Lemmas:

Lemma 1.50. *Sei M ein A -Modul und $N \subset M$ ein Untermodul. Dann gelten die Aussagen:*

1. *M ist endlich erzeugt, genau dann, wenn für ein $l \in \mathbb{N}$ und einen Untermodul $D \subset A^l$ die Isomorphie*

$$M \cong A^l/D$$

gilt.

2. *M ist noethersch, genau dann, wenn N und der Quotientenmodul M/N noethersch sind.*

Kapitel 1 Computeralgebra

Beweis. Die Beweise sind elementare Modultheorie und folgen der Beobachtung, daß für einen surjektiven Homomorphismus $\phi : W \rightarrow V$ zwischen zwei A -Moduln stets die Isomorphie $V \cong W/\ker \phi$ gilt. \square

Die Verallgemeinerung des Polynomrings auf den Modulfall geschieht durch die komponentenweise Definition der Monome:

Definition 1.51. Es sei für einen Polynomring $A = K[x_1, \dots, x_n]$ der endlich erzeugte Modul $A^r = \bigoplus_{i=1}^r A \cdot e_i$ wie oben. Ein *Monom* ist ein Ausdruck der Form $x^\alpha \cdot e_i = (0, \dots, x^\alpha, \dots, 0)$.

Analog zu Definition 1.12 wird auf $Mon_{A^r} = \{x^\alpha e_i \mid \alpha \in \mathbb{N}^n, i = 1, \dots, r\}$ zu einer gegebenen Monomordnung in A eine Modulordnung definiert:

Definition 1.52. Sei $<$ eine monomiale Ordnung auf $A = K[x_1, \dots, x_n]$ und A^r für $r \in \mathbb{N} \setminus \{0\}$, dann heißt eine totale Ordnung $<_{A^r}$ auf $\{x^\alpha \cdot e_i \mid \alpha \in \mathbb{N}^n, i = 1, \dots, r\}$ *Modulordnung*, wenn sie kompatibel mit der Ringordnung $<$ ist und $\forall \alpha, \beta, \gamma \in \mathbb{N}^n, \forall i, j \in \{1, \dots, r\}$ gilt:

1.

$$x^\alpha \cdot e_i >_{A^r} x^\beta \cdot e_j \quad \Rightarrow \quad x^{\alpha+\gamma} \cdot e_i >_{A^r} x^{\beta+\gamma} \cdot e_j$$

2.

$$x^\alpha > x^\beta \quad \Rightarrow \quad x^\alpha \cdot e_i >_{A^r} x^\beta \cdot e_i$$

Die zweite Kompatibilitätsbedingung bedeutet, daß die Eigenschaften *global*, *lokal* und *gemischt* von der Ringordnung auf die Modulordnung übertragen werden.

Beispiel 1.53. Man kann mithilfe der Ordnung $\leq = (<, c)$ zuerst nach den Monomen in A

$$x^\alpha \cdot e_i > x^\beta \cdot e_j \quad :\Leftrightarrow \quad x^\alpha > x^\beta \vee (x^\alpha = x^\beta \wedge i < j)$$

oder mithilfe von $\leq = (c, <)$ zuerst nach den Komponenten in A^r ordnen:

$$x^\alpha \cdot e_i > x^\beta \cdot e_j \quad :\Leftrightarrow \quad i < j \vee (i = j \wedge x^\alpha > x^\beta)$$

Der Begriff des Leitideals wird in diesem Kontext leicht modifiziert und durch den Begriff Leitmodul ersetzt:

Definition 1.54. Für eine Modulordnung $<$ in A^r und eine Menge $M \subset A^r$ heißt der von den Leitmonomen in M erzeugte Modul

$$L(M) := \langle lm(m) \mid m \in M \rangle_A \subset A^r$$

der *Leituntermodul* von $\langle M \rangle_A$. Wenn M ein Untermodul von A^r ist, dann heißt $L(M)$ auch *Leitmodul* von M .

Teilbarkeit zweier Monome liegt nur vor, wenn sie in der gleichen Komponente auftreten:

$$x^\alpha \cdot e_i \mid x^\beta \cdot e_j \quad :\Leftrightarrow \quad i = j \wedge x^\alpha \mid x^\beta$$

Damit kann auch der Begriff der Gröbnerbasis im A^r präzisiert werden:

Definition 1.55. Für einen Untermodul $M \subset A^r$ heißt eine endliche Menge $G \subset M$ *Gröbnerbasis*, wenn gilt, daß $L(G) = L(M)$ und somit

$$\forall m \in M \setminus \{0\} \quad \exists g \in G : \quad lm(g) \mid lm(m)$$

Bemerkung 1.56. Der Begriff der *Normalform*, der Normalform-Algorithmus, das Buchberger Kriterium und der Gröbnerbasen-Algorithmus übertragen sich nach Austauschen von A durch A^r wörtlich, sodaß hier auf eine detaillierte Angabe verzichtet wird.

Die Analogien der Eliminierungseigenschaften von Ordnungen im Ring- und Modulfall werden kurz aufgezeigt: Eliminierung in Moduln heißt Eliminierung von *Komponenten*, wohingegen im Ring mithilfe einer Eliminationsordnung *Variablen* eliminiert werden. Somit stellt sich an dieser Stelle die Aufgabe, zu einem Untermodul

$$M = \langle m_1, \dots, m_k \rangle_A \subset A^r = \bigoplus_{i=1}^r A \cdot e_i$$

Kapitel 1 Computeralgebra

entsprechende Erzeuger für den Modul

$$M \cap \bigoplus_{i=s+1}^r A \cdot e_i$$

zu finden. Die Lösung bietet das nächste Lemma:

Lemma 1.57. Eliminationslemma für Moduln

Sei $M \subset A^r = \bigoplus_{i=1}^r A \cdot e_i$ ein Untermodul, sowie $G \subset M$ eine Gröbnerbasis von M bezüglich der Modulordnung $\leq = (c, <)$, welche in 1.53 definiert ist und den Komponenten den Monomen in A gegenüber, sowie den höheren Komponenten den niedrigeren Komponenten gegenüber den Vorzug gibt.

Es gilt, daß $\forall s = 0, \dots, r-1$

$$G' := G \cap \bigoplus_{i=s+1}^r A \cdot e_i$$

ist eine Gröbnerbasis von

$$M' := M \cap \bigoplus_{i=s+1}^r A \cdot e_i$$

bezüglich der Ordnung $<$.

Beweis. Der Beweis verwendet die Eliminationseigenschaft der Modulordnung $<$ für die ersten s Komponenten und argumentiert wie im Eliminationslemma für Ringe 1.42. \square

Kapitel 2

Finite Differenzenschemata

In diesem Kapitel soll eine Gattung der numerischen Methoden zur Lösung von linearen partiellen Differentialgleichungen, die der finiten Differenzenverfahren, vorgestellt werden. Die mathematische Theorie der zugrundeliegenden Gleichungen wird, soweit für das Verständnis nötig, in diese Darstellung eingebettet.

2.1 Einführung

Betrachte die wohlbekannte Advektionsgleichung als Beispiel für ein Cauchy-Problem in einer unbekanntenen, reellwertigen Funktion $u = u(x, t)$ in zwei reellen Argumenten x und t :

$$u_t + au_x = 0$$

mit der gegebenen Anfangsbedingung

$$u(x, 0) = u_0(x)$$

Das Theorem von Cauchy-Kowalewsky gewährleistet, daß für eine Umgebung des Randes, an welchem die Anfangsbedingung gilt, eine eindeutige analytische Lösung für das obige Problem existiert.

Somit stellt sich die Frage, ob durch Diskretisierungen der obigen Gleichung eine zufriedenstellende Approximation erreicht werden kann: es sei für feste reelle Konstanten $(h, k) = (\Delta x, \Delta t) \in \mathbb{R}^+ \times \mathbb{R}^+$ der Näherungswert der unbekanntenen Funktion u in den diskreten Punkten $(ih, jk) \in \mathbb{R} \times \mathbb{R}^+$ und $(i, j) \in \mathbb{Z} \times \mathbb{N}$ mittels

$$u_i^j = u(ih, jk)$$

bezeichnet. Die (x, t) -Ebene des Definitionsbereichs der Funktion u wird durch ein orthogonales und äquidistantes Gitter wie in Abbildung 2.1 ersetzt, auf welchem mit festen Schrittweiten h und k die Approximationen u_i^j betrachtet werden.

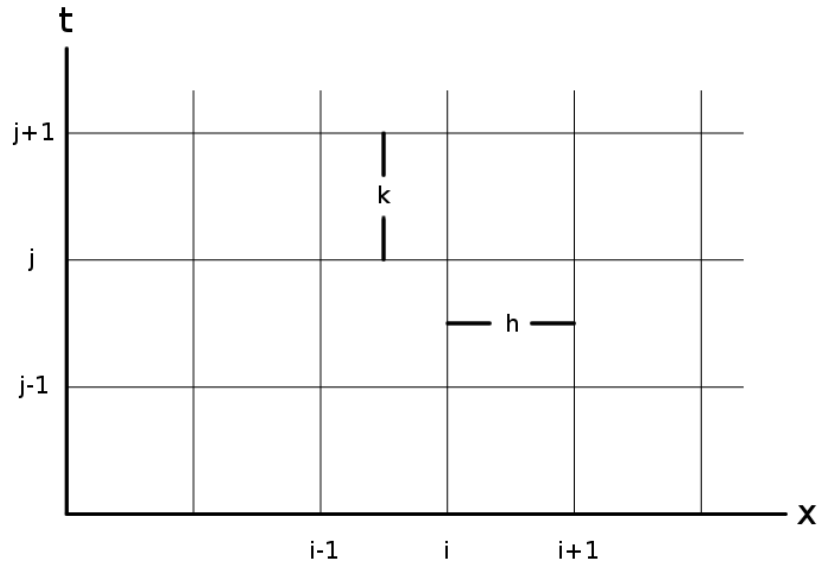


Abbildung 2.1: Orthogonales und äquidistantes Gitter in (x, t) -Ebene

Nun werden die auftretenden *Differentialquotienten* u_t und u_x in der zu lösenden Gleichung $u_t + au_x = 0$ durch approximative *Differenzenquotienten* um den Punkt (ih, jk) ersetzt. Hierbei treten endlich viele u_m^l für gewisse $(m, l) \in N_1 \times N_2 \subset \mathbb{Z} \times \mathbb{N}$ in eine Relation untereinander. Die Illustration dieses Konzepts wird an zwei verschiedenen Beispielen deutlich gemacht.

1. Das naive Verfahren: Es seien für u_t

$$(u_t)_i^j = \frac{u_i^{j+1} - u_i^j}{k}$$

und für u_x

$$(u_x)_i^j = \frac{u_{i+1}^j - u_{i-1}^j}{h}$$

Approximationen gewählt, dann erhält man die Abhängigkeit der Werte u_i^{j+1} des höchsten Zeitschritts $j + 1$ durch Werte der u_m^j mit $m \in \{i - 1, i, i + 1\}$:

$$u_i^{j+1} = u_i^j - \frac{ak}{2h} \left(u_{i+1}^j - u_{i-1}^j \right)$$

2. Das Lax-Friedrich-Verfahren: Mithilfe der Approximationen

$$(u_t)_i^j = \frac{u_i^{j+1} - \frac{1}{2} \left(u_{i+1}^j + u_{i-1}^j \right)}{k}$$

für u_t sowie

$$(u_x)_i^j = \frac{u_{i+1}^j - u_{i-1}^j}{2h}$$

für u_x ergibt sich nach Ersetzung

$$\frac{u_i^{j+1} - \frac{1}{2} \left(u_{i+1}^j + u_{i-1}^j \right)}{k} + a \frac{u_{i+1}^j - u_{i-1}^j}{2h} = 0$$

die für obige Approximationen spezifische Relation, die, umgeformt und nach den u_i^{j+1} mit dem höchsten Index $j + 1$ aufgelöst, eine rekursive Formel zur Errechnung von u_i^{j+1} aus den u_m^j

$$u_i^{j+1} = \frac{1}{2} \left(u_{i+1}^j + u_{i-1}^j \right) - \frac{ak}{2h} \left(u_{i+1}^j - u_{i-1}^j \right)$$

liefert, die auch als *Lax-Friedrich-Verfahren* bekannt ist.

Der sogenannte *Differenzenstern* zeigt geometrisch die Relation der u_m^l und ist für obiges Schema in Abbildung 2.2 zu sehen.

2.2 Konvergenztheorie

Die bisher in Abschnitt 2.1 skizzierten Begriffe werden nun detailliert erläutert und die Konvergenztheorie von finiten Differenzenverfahren mit Schwerpunktsetzung auf den Von-Neumann-Ansatz dargelegt.

Es wird sich im weiteren Verlauf auf Funktionen u in zwei Argumenten x und t

$$\begin{aligned} u : \quad \mathbb{R} \times \mathbb{R}^+ &\longrightarrow \mathbb{R} \\ (x, t) &\longmapsto u(x, t) \end{aligned}$$

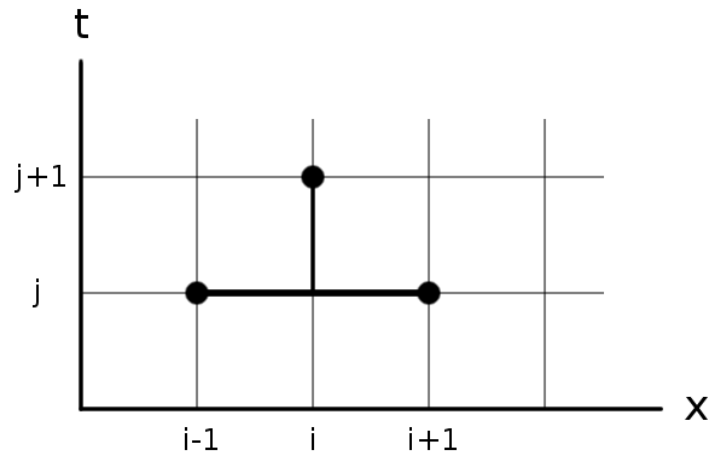


Abbildung 2.2: Differenzenstern für Lax-Friedrich-Verfahren

eingeschränkt, obwohl die Begriffe und Sätze auch für den Fall einer zeitlichen (t) und mehrerer räumlicher (x_1, \dots, x_n) Argumente Gültigkeit besitzen.

Auf einem rechtwinkligen und äquidistanten Gitter wie in Abbildung 2.1 werden die Gitterabstände in x - und t -Richtung für positive $h, k \in \mathbb{R}^+$, die mittels einer Gitterkonstante

$$\lambda := \frac{k}{h}$$

einem konstanten Verhältnis unterliegen, eingeführt und somit ist es möglich durch Variation von k die Verfeinerung des Gitters herbeizuführen.

Die Punkte im Definitionsbereich $\mathbb{R} \times \mathbb{R}^+$ der unbekannt Funktion u , in welchen die Funktionswerte von u berechnet werden sollen, seien mit $(x_i, t_j) := (ih, jk)$ für $(i, j) \in \mathbb{Z} \times \mathbb{N}$ bezeichnet. Ausgehend von einer Anfangsbedingung

$$u(x, 0) = u_0(x)$$

definiere nach der Initialisierung

$$u_i^0 := u(x_i, 0) = u_0(x_i) \quad \forall i \in \mathbb{Z}$$

iterativ alle Werte u_i^k mittels der Definition

$$u_i^{j+1} := H(u_{i-s}^j, \dots, u_{i+s}^j) \quad s \in \mathbb{N}$$

für ein Verfahren H mit

$$H : \mathbb{R}^{2s+1} \longrightarrow \mathbb{R}$$

Verfahren dieser Art werden als *finite Differenzenmethoden* oder *-schemata* bezeichnet

Definition 2.1. Ein Verfahren H , bei dem zur Berechnung eines Wertes u_i^{j+1} im Zeitschritt $j + 1$ nur Terme u_i^j des nächstniedrigeren Zeitschritts j auftauchen, heißt *explizites Zweischrittverfahren*.

Die Stabilitätstheorie und insbesondere der Von-Neumann-Ansatz zur Stabilitätsbestimmung sind laut [Tho95] nur für explizite Zweischrittverfahren definiert, sodaß diese spezielle Art von Verfahren auch hier im Folgenden angenommen wird.

Differenzenverfahren H wie in Definition 2.1 sollen viele Eigenschaften der Funktion u erhalten und mittels der Relation, die aus der Differentialgleichung resultiert, auf die approximative Lösung der u_i^j -Werte übertragen.

2.2.1 Konsistenz und Konvergenz

Ein Maß für die Güte eines Verfahrens ist der Fehler, der sich aus der Wahl der Approximationen in H ergibt. Ein spezielles Konzept der Fehlerbestimmung ist das des *lokalen Fehlers*:

Definition 2.2. Ein Verfahren H heißt *konsistent* mit einer *Konsistenzordnung* p , wenn eine maximale Zahl $p \in \mathbb{N}_{\geq 1}$ existiert mit

$$u(x, t + k) = H(u(x - sh, t), \dots, u(x + sh, t)) + \mathcal{O}(k^{p+1})$$

wobei u die Lösung der zugrundeliegenden Differentialgleichung ist.

Die Abschätzung dieses lokalen Fehlers erfolgt bei einer hinreichend oft differenzierbaren Funktion u durch den Vergleich des Differenzenschemas, welches aus gewissen Approximationen der auftretenden Ableitungen resultiert, mit dem Taylorpolynom an einer geeigneten Entwicklungsstelle. Dazu sei ein Ergebnis aus der reellen Analysis zitiert:

Theorem 2.3. Taylorpolynom

Es sei auf einer offenen Menge $G \subset \mathbb{R}^n$ die Funktion $f : G \rightarrow \mathbb{R}$ $(m + 1)$ -mal stetig differenzierbar für $m \in \mathbb{N}$. Dann existiert für alle $x, \Delta x$ mit $S := \{x + s\Delta x \mid s \in [0, 1]\} \subset G$ ein $\xi \in S$, sodaß

$$f(x + \Delta x) = \sum_{|i| \leq m} \frac{\partial^i f(\Delta x)}{i!} (\Delta x)^i + \sum_{|i|=m+1} \frac{\partial^i f(\xi)}{i!} (\Delta x)^i$$

mit der Multiindexschreibweise als Kurzform für die bessere Schreibbarkeit:

$$i = (i_1, \dots, i_n)$$

$$|i| = i_1 + \dots + i_n$$

$$i! = i_1! \cdot \dots \cdot i_n!,$$

$$\partial^i f = \frac{\partial^{|i|}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}}$$

$$(\Delta x)^i = \Delta x_1^{i_1} \cdot \dots \cdot \Delta x_n^{i_n}$$

Beweis. Der Beweis ist in [Wal02] zu sehen. □

Ein Beispiel für den lokalen Fehler ist anhand der Vorwärtsapproximation der Ableitung u_t zu beobachten:

Beispiel 2.4. Für die Approximation

$$u_t(x, t) = \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t}$$

ergibt sich mittels Taylorentwicklung der hinreichend oft differenzierbaren Funktion u im Punkt $z = (x, t)$ für $\Delta z = (\Delta x, \Delta t)$ mit $\Delta x = 0$ und $\Delta t > 0$, daß

$$u(x, t + \Delta t) = \sum_{n=0}^m \frac{\partial^n u(x, t)}{n!} (\Delta t)^n + \frac{\partial^{m+1} u(x, t + \xi)}{(m + 1)!} (\Delta t)^{m+1}$$

für ein $\xi \in]0, \Delta t[$. Man beachte, daß die Summe über alle Indizes des Betrages wie in Theorem 2.3 eine Summe ausschließlich über die Terme ist, die in der Δt -Komponente ungleich Null sind. Insbesondere verschwinden alle Terme mit $\Delta x = 0$, sodaß eine Taylorentwicklung wie im eindimensionalen Fall vorliegt. Die Restgliedabschätzung ergibt für $m = 1$:

$$u(x, t + \Delta t) = u(x, t) + \Delta t u_t(x, t) + \frac{1}{2} u_{tt}(x, t + \xi) (\Delta t)^2$$

und somit nach Umstellung in diskreter Schreibweise für $\Delta t = k$

$$\frac{u_i^{j+1} - u_i^j}{k} = (u_t)_i^j + k(\dots) = (u_t)_i^j + \mathcal{O}(k^1)$$

Daher läßt sich der lokale Abschneidefehler als $\mathcal{O}(k)$ charakterisieren.

Für die Approximation der Ableitung u_{tt} mit

$$\frac{u_i^{j+1} - 2u_i^j + u_i^{j-1}}{k^2} = (u_{tt})_i^j + \mathcal{O}(k^2)$$

erhält man mit ähnlicher Argumentation die Konsistenzordnung $\mathcal{O}(k^2)$.

Das Abschätzen der *lokalen* Fehler ist allerdings nicht ausreichend für einen Konvergenzbegriff *globaler* Natur, der eine Aussage über die Güte des Verfahrens H für den kompletten Definitionsbereich der Lösungsfunktion u bereit hält.

Als Konzept zur Untersuchung ebendieser wird eine Fehlerfunktion definiert:

Definition 2.5. Für ein äquidistantes Gitter mit einer festen Gitterkonstante $\lambda = \frac{k}{h}$ definiere mittels $(x_i, t_j) = (ih, jk)$ eine stückweise konstante Funktion

$$u_k(x, t) := u_i^j$$

und ebenso

$$\tilde{u}_k(x, t) := u(x_i, t_j)$$

für $(x, t) \in [x_i, x_{i+1}[\times [t_j, t_{j+1}[$.

Mit den Funktionen u_k und \tilde{u}_k wird eine ebenfalls stückweise Fehlerfunktion

$$e_k(x, t) := \tilde{u}_k(x, t) - u_k(x, t)$$

definiert.

Die Eigenschaften und das Verhalten dieser Fehlerfunktion e_k bei immer kleiner werdenden Gitterabständen h, k geben Aufschluss über die Konvergenz des Verfahrens H :

Definition 2.6. Konvergenz

Ein Verfahren H heißt für eine Norm $\|\cdot\|$ *konvergent* gegen die Lösungsfunktion u , wenn im Grenzübergang $k \rightarrow 0$ die Bedingung

$$\forall T \quad \sup_{t \in [0, T]} \|e_k(\cdot, t)\| \xrightarrow{k \rightarrow 0} 0$$

erfüllt ist.

Die Konvergenz ist die erwünschte Eigenschaft eines Verfahrens H , da sie sicherstellt, daß der von H induzierte Fehler über dem gesamten Definitionsbereich für hinreichend kleine Gitterabstände ebenfalls klein ist. Konsistenz ist für Konvergenz ein notwendiges, jedoch kein hinreichendes Kriterium, wie sich im Folgenden zeigen wird.

2.2.2 Stabilität

Eine sehr wichtige Eigenschaft des Verfahrens H ist seine Stabilität. Stabilität bei einem iterativen Verfahren bedeutet anschaulich, daß kleine Fehler in der Initialisierung sich nicht stark auf die erhaltene approximative Lösung auswirken und der Fehler zwischen ebendieser und der wirklichen Lösung klein bleibt. Formal bedeutet Stabilität:

Definition 2.7. Stabilität

Ein Verfahren H heißt bezüglich einer Norm $\|\cdot\|$ *stabil*, wenn

$$\exists K \in \mathbb{R}^+ \quad \forall j : \quad \|H^j\| = \sup_{u \neq 0} \frac{\|H^j u\|}{\|u\|} \leq K$$

Der in der Einführung dieses Kapitels erwähnte Satz von Cauchy-Kowalewsky garantiert, daß es sich bei Cauchy-Problemen um korrekt gestellte Probleme handelt (vergleiche [Had53] und [KS82]):

Definition 2.8. Ein Problem heißt nach Hadamard *korrekt gestellt*, wenn eine Lösung existiert, diese Lösung eindeutig ist und die Lösung stetig (bezüglich einer geeigneten Topologie) von den Anfangsdaten abhängt.

Mit letzterer Definition lassen sich nun die Lax-Theoreme notieren:

Theorem 2.9. Lax Äquivalenz Theorem

Ein konsistentes, explizites Zweischrittverfahren H für ein korrekt gestelltes, lineares Anfangswertproblem ist konvergent, genau dann, wenn H stabil ist.

□

Es kann mit der Konsistenzordnung eines Verfahrens H auf die Konvergenzordnung geschlossen werden:

Theorem 2.10. Lax Theorem

Es sei H ein Verfahren wie in Theorem 2.9. Dann ist die Konvergenzordnung gleich der Konsistenzordnung.

□

Die Beweise der Theoreme 2.9 und 2.10 können in [Tho95] gefunden werden.

2.2.3 Die Methode nach Von Neumann

Somit bleibt eine Stabilitätsbetrachtung, die für ein Verfahren die Konvergenz sichert. Obige Begriffe hatten abstrakt eine Norm $\|\cdot\|$ zugrundeliegend, nach welcher Konvergenz und Stabilität bemessen werden. In der Praxis nimmt man sehr oft die sogenannte L^2 - oder $\|\cdot\|_{L^2}$ -Norm, da hier der Von-Neumann-Ansatz mit Fourieranalyse verwendet werden kann.

Einfache Gleichung

Für eine kontinuierliche, quadratisch integrierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ definiere die L^2 -Norm:

$$\|f\|_{L^2} := \left(\int_{\mathbb{R}} |f(x)|^2 dx \right)^{\frac{1}{2}}$$

Zu einer solchen Funktion f wird die *Fouriertransformierte*

$$\mathcal{F}_\xi \circ f := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ix\xi} dx$$

Kapitel 2 Finite Differenzenschemata

Die Fouriertransformation ist eine *isometrische* Abbildung $\mathcal{F}_\xi : L^2 \rightarrow L^2$, das heißt, sie ist normerhaltend im L^2 gemäß einem Ergebnis aus der Analysis, der Parseval-Identität (siehe [Tho95]):

$$\|f\|_{L^2} = \|\mathcal{F}_\xi \circ f\|_{L^2}$$

Die Definition dieser Begriffe kann auch im diskreten Fall erfolgen. Es sei eine Folge $(a_n)_{n \in \mathbb{Z}}$ im l^2 -Folgenraum mit der Eigenschaft

$$\|(a_n)_n\|_{l^2} := \left(\sum_{n=-\infty}^{\infty} |a_n|^2 \right)^{\frac{1}{2}} < \infty$$

Dann definiere die Fouriertransformation im diskreten Fall:

Definition 2.11. Diskrete Fouriertransformierte

Für eine Folge $a = (a_n)_n \in l^2$ ist die *diskrete Fouriertransformierte* eine Funktion $\hat{a} : [-\pi, \pi] \rightarrow \mathbb{R}$, die für $\xi \in [-\pi, \pi]$ definiert ist als

$$\hat{a}(\xi) := \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} e^{-in\xi} a_n$$

Eine Aussage über die Normerhaltung bei Anwendung auch der diskreten Fourieranalyse gibt das folgende Theorem:

Theorem 2.12. Parseval Identität im diskreten Fall

Für eine Folge $a = (a_n)_n \in l^2$ und die diskrete Fouriertransformierte \hat{a} gilt die Gleichheit der beiden Normen

$$\|a\|_{l^2} = \|\hat{a}\|_{L^2}$$

Beweis. Der Beweis ist in [Tho95] zu sehen. □

Eine besondere Eigenschaft, die sich bei Indexverschiebung ergibt, wird später ausgenutzt werden:

Bemerkung 2.13. Für eine Folge $a = (a_n)_n$ bedeutet eine Indexverschiebung mittels $m = n \pm 1$ bei der Fouriertransformierten \hat{a} eine Multiplikation mit $e^{\pm i\xi}$. Dies sieht man anhand

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} e^{-in\xi} a_{n\pm 1} &= \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} e^{-i(m\pm 1)\xi} a_m \\ &= e^{\pm i\xi} \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} e^{-im\xi} a_m \\ &= e^{\pm i\xi} \hat{a}(\xi) \end{aligned}$$

Es sei wie oben H ein explizites Zweischnittverfahren mit

$$u_k^{j+1} := H(u_{k-m}^j, \dots, u_{k+m}^j) = \left(\sum_{s=-m}^m c_s u_{k+s}^j \right)$$

für ein $s \in \mathbb{N}$. Als Folge in l^2 werden nun

$$u^{j+1} := (u_i^{j+1})_i$$

respektive

$$u^j := (u_i^j)_i$$

betrachtet und es seien nur endlich viele Folgenglieder ungleich Null. Dies garantiert, daß die Folgen im Raum l^2 liegen, und bedeutet keine Einschränkung, da die Diskretisierung eines Teils des Definitionsbereichs durch ein Gitter für praktische Zwecke stets nur endlich viele Gitterpunkte besitzt.

Nach Anwendung der Fouriertransformation ergibt sich für obige Gleichung

$$\begin{aligned} \hat{u}^{j+1}(\xi) &= \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} e^{-ik\xi} u_k^{j+1} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} e^{-ik\xi} \sum_{s=-m}^m c_s u_{k+s}^j \\ &= \sum_{s=-m}^m c_s \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} e^{-ik\xi} u_{k+s}^j \end{aligned}$$

Kapitel 2 Finite Differenzenschemata

und mit Bemerkung 2.13 über die Indexverschiebung:

$$\begin{aligned} &= \sum_{s=-m}^m c_s e^{si\xi} \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} e^{-ik\xi} u_k^j \\ &= \hat{u}^j(\xi) \sum_{s=-m}^m c_s e^{si\xi} \end{aligned}$$

Definition 2.14. Symbol

Der Term

$$\varrho(\xi) := \sum_{s=-m}^m c_s e^{si\xi}$$

heißt das *Symbol* des Verfahrens H .

Man kann mithilfe des Symbols die Stabilität des Verfahrens bezüglich der $\|\cdot\|_{l^2}$ -Norm bestimmen, denn es gilt nach iterativer Anwendung des Verfahrens H :

$$\hat{u}^{j+1}(\xi) = \hat{u}^0(\xi) \varrho(\xi)^{j+1}$$

Und nach Bilden der $\|\cdot\|_{L^2}$ -Norm auf beiden Seiten gilt:

$$\|\hat{u}^{j+1}(\xi)\|_{L^2} = \|\hat{u}^0(\xi)\|_{L^2} |\varrho(\xi)^{j+1}|$$

Aufgrund des Theorems 2.12 kann man schreiben

$$\|u^{j+1}\|_{l^2} = \|u^0\|_{l^2} |\varrho(\xi)^{j+1}|$$

Das heißt, das Verfahren H besitzt Stabilität, wenn

$$\sup_{\xi} |\varrho(\xi)| \leq 1$$

gilt.

Beispiel 2.15. Betrachte das Lax-Friedrich-Verfahren mit $\lambda = \frac{k}{h}$ als Gitterkonstante. Das Verfahren liefert

$$u_i^{j+1} = \frac{1}{2} \left(u_{i+1}^j + u_{i-1}^j \right) - \frac{ak}{2h} \left(u_{i+1}^j - u_{i-1}^j \right)$$

Es gilt somit für das Symbol $\varrho(\xi) = \sum_{s=-m}^m c_s e^{is\xi}$:

$$\varrho(\xi) = \frac{1 + a\lambda}{2} e^{-i\xi} + \frac{1 - a\lambda}{2} e^{i\xi}$$

und nach der Eulerschen Formel $e^{i\phi} = \cos(\phi) + i \sin(\phi)$ ist dies nichts anderes als

$$\cos(\xi) - ia\lambda \sin(\xi)$$

Der quadrierte komplexe Betrag ist

$$|\varrho(\xi)|^2 = \cos^2(\xi) + a^2 \lambda^2 \sin^2(\xi)$$

Man sieht leicht, daß aufgrund von

$$\sup_{\xi} |\varrho(\xi)| \leq 1 \iff |a\lambda| \leq 1$$

die hinreichende Bedingung für Stabilität durch $|a\lambda| \leq 1$ gegeben ist.

Eine Folgerung dieser Beobachtungen ist das Theorem von Von Neumann, welches ein Stabilitätskriterium liefert:

Theorem 2.16. Von Neumann

Ein lineares Verfahren H wie in 2.9 mit der Eigenschaft

$$\sup_{\xi} |\varrho(\xi)| \leq 1$$

ist stabil bezüglich der $\|\cdot\|_{l^2}$ -Norm. □

Wie nachfolgend erkennbar ist die Erfüllung der sogenannten CFL-Bedingung $|a\lambda| \leq 1$ aus Beispiel 2.15 ein hinreichendes Kriterium für Konvergenz nach

Kapitel 2 Finite Differenzenschemata

Theorem 2.16, aber aus offensichtlichen Gründen auch *notwendig*:
die Gleichung der Lösungsfunktion

$$u_t + au_x = 0$$

und die Anfangsbedingung

$$u(x, 0) = u_0(x)$$

liefern mit Anwendung der Kettenregel der Differentialrechnung eine Lösung für die obige Gleichung:

$$u_0(x - at)$$

Dann hängt für einen Punkt (x, t) , der sich mit $x(t) = x_0 + at$ berechnet, alles von dem Startwert x_0 ab:

Definition 2.17. Die Menge aller Punkte, von denen ein Punkt abhängt, heißt *analytisches Abhängigkeitsgebiet*.

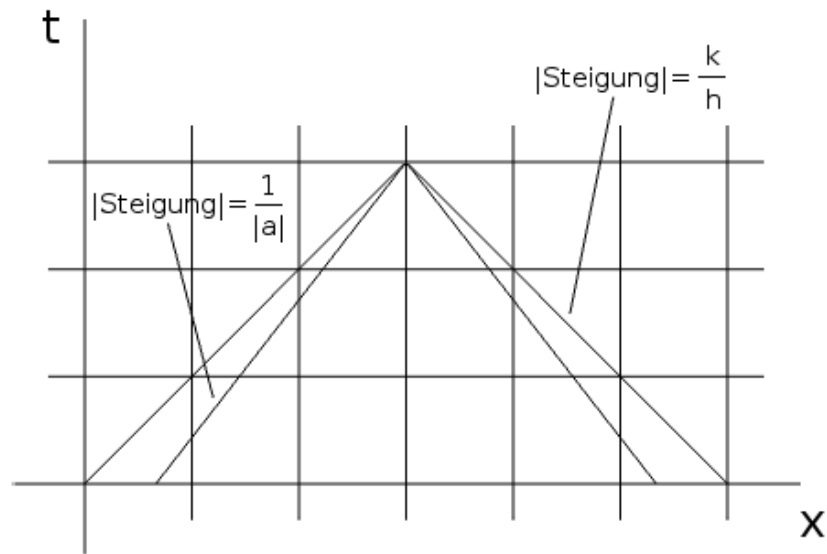
Es ist selbstverständlich, daß die numerischen Abhängigkeiten, die sich aus dem Verfahren H ergeben, die analytischen Abhängigkeiten respektieren müssen. Dies bedeutet anschaulich, daß das analytische Abhängigkeitsgebiet im numerischen Abhängigkeitsgebiet liegen muss, was in Abbildung 2.3 zu sehen ist.

Definition 2.18. CFL-Bedingung

Wenn das analytische Abhängigkeitsgebiet im numerischen Abhängigkeitsgebiet enthalten ist, erfüllt das Verfahren H die *Courant-Friedrichs-Lewy-Bedingung*.

Die Bestätigung der obigen Argumentation erfolgt durch das Theorem (in [Tho95]):

Theorem 2.19. *Die CFL-Bedingung ist notwendig für die Konvergenz eines Verfahrens.* □

Abbildung 2.3: Es muss $|a\lambda| \leq 1$ gelten.

Aber die CFL-Bedingung ist nur eine notwendige und im allgemeinen nicht hinreichende Bedingung für Konvergenz, wie das nächste Beispiel zeigt:

Beispiel 2.20. *Das naive Verfahren*

$$u_i^{j+1} = u_i^j - \frac{a}{2}\lambda \left(u_{i+1}^j - u_{i-1}^j \right)$$

hat das Symbol

$$\varrho(\xi) = 1 - \frac{a}{2}\lambda \left(e^{i\xi} - e^{-i\xi} \right) = 1 - ia\lambda \sin(\xi)$$

und damit gilt für den komplexen Betrag von $\varrho(\xi)$ zum Quadrat

$$|\varrho(\xi)|^2 = 1 + a^2\lambda^2 \sin^2(\xi)$$

dieser Ausdruck ist im Supremum $\sqrt{1 + a^2\lambda^2}$ sicher größer als 1 und damit ist das naive Verfahren instabil.

Systeme

Der Von-Neumann-Ansatz ist auch auf den linearen Fall von Systemen verallgemeinerbar. Sei

$$\partial_t \omega + A \partial_x \omega = 0$$

für $A \in \mathbb{R}^{n \times n}$ eine partielle Differentialgleichung in einer unbekanntem Funktion

$$\omega : (x, t) \mapsto \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_n \end{pmatrix} \in \mathbb{R}^n$$

Für ein Verfahren auf einem Gitter H mit

$$\omega_i^{j+1} = H \left(\omega_{i-l}^j, \dots, \omega_{i+l}^j \right) = \sum_{s=-l}^l c_s \omega_{i+s}^j$$

und $c_s \in \mathbb{R}^{n \times n}$ betrachte in völliger Analogie zum Fall *einer* Gleichung den Einsatz der $\|\cdot\|_2$ -Norm mit dem Symbol (Amplifikationsmatrix)

$$\varrho(\xi) := \sum_{s=-l}^l c_s e^{is\xi}$$

welches diesmal eine $n \times n$ -Matrix ist. Die Stabilität folgt, wenn für den Spektralradius von $\varrho(\xi)$ gilt

$$\sup_{\xi} |\varrho(\xi)| \leq 1$$

Vergleiche [Tho95].

Kapitel 3

Quantorenelimination im \mathbb{R}^r

Das Konzept der zylindrischen algebraischen Zerlegung (englisch: cylindrical algebraic decomposition, kurz: CAD) erlaubt für eine gegebene Menge von Polynomen in r Variablen die Zerlegung des Raumes \mathbb{R}^r in Zellen mit bestimmten Eigenschaften dieser Zellen für die Polynome. Eine wichtige Eigenschaft einer solchen Zerlegung im Zusammenhang mit Quantorenelimination ist die Wahrheitsinvarianz dieser Zellen für polynomiale Formeln, sodaß die Formel auf jeder Zelle bezüglich ihres Wahrheitswerts **TRUE** respektive **FALSE** invariant ist. Damit können wertvolle Aussagen getroffen werden und eine Quantorenelimination von polynomialen Formeln vorgenommen werden, sofern eine solche äquivalente Aussage existiert.

3.1 Motivation

Ein wohlbekanntes Problem der Quantorenelimination ist die Aussage

$$\exists x : \quad x^2 + px + q = 0$$

mit einer dazu äquivalenten Aussage ohne Quantoren

$$\Leftrightarrow 4q - p^2 \leq 0$$

welche auch als p, q -Formel bezeichnet wird. Diese Art äquivalente Aussagen zu finden ist das mathematische Teilgebiet, welches nachfolgend vorgestellt wird.

Kapitel 3 Quantorenelimination im \mathbb{R}^r

Definition 3.1. Formel

Mit *Formel* wird eine Aussage der Form

$$F = (Q_1 x_{k+1})(Q_2 x_{k+2}) \dots (Q_{r-k} x_r) [\underbrace{\acute{F}(x_1, \dots, x_k)}_{\text{frei}}, \underbrace{x_{k+1}, \dots, x_r}_{\text{gebunden}}]$$

wobei $Q_i \in \{\forall, \exists\}$ und \acute{F} ein polynomialer Ausdruck ohne Quantoren ist, bezeichnet. Diese spezielle Form heißt in der Prädikatenlogik *standard prenex formula*.

Das Vorgehen zur Quantorenelimination für eine polynomiale Formel wie in Definition 3.1 mithilfe der CAD-Technik besteht aus zwei Schritten:

1. Berechnung einer CAD des Koordinatenraumes \mathbb{R}^r aller Variablen (auch der gebundenen) mit Wahrheitswertinvarianz aller Zellen für F . Damit wird auch eine Zerlegung des Raumes \mathbb{R}^k der freien Variablen berechnet.
2. Betrachte Polynome, die nur von den freien Variablen x_1, \dots, x_k bestehen und dasselbe Vorzeichen auf den Zellen der Zerlegung besitzen. Berechne aus diesen eine Lösungsformel, die äquivalent zur Formel in F ist.

3.2 Zylindrische Algebraische Zerlegung im \mathbb{R}^r

Es können Zerlegungen nach Eigenschaften charakterisiert werden:

Definition 3.2. Eine Zerlegung des Raumes

$$\mathbb{R}^r = \bigcup_{i=1, \dots, 2m+1} c_i$$

heißt

1. *algebraisch*, wenn die c_i *semi-algebraische* Mengen sind, das heißt, von der Form

$$c_i = \{x = (x_1, \dots, x_r) \mid f(x) \geq 0, f \in \mathbb{R}[x_1, \dots, x_r]\}$$

oder eine endliche Vereinigung, ein Schnitt, ein Komplement solcher Mengen.

3.2 Zylindrische Algebraische Zerlegung im \mathbb{R}^r

2. *zylindrisch algebraisch* oder eine *CAD*, wenn die Zerlegung algebraisch ist und eine zusätzliche Struktur besitzt.

Diese zusätzliche Struktur für \mathbb{R}^r mit $r \geq 2$ ist iterativ definiert und wird für \mathbb{R}^1 initialisiert:

Definition 3.3. CAD im \mathbb{R}^1

Eine Zerlegung

$$\mathbb{R}^1 = \bigcup_{i=1, \dots, 2m+1} c_i$$

heißt *zylindrische algebraische Zerlegung* oder *CAD*, wenn

$$c_1 =]-\infty, \alpha_1[\quad c_2 = [\alpha_1, \alpha_1] \quad \dots \quad c_{2m} = [\alpha_m, \alpha_m] \quad c_{2m+1} =]\alpha_m, \infty[$$

für algebraische Zahlen $\alpha_1, \dots, \alpha_m$.

Bemerkung 3.4. Man kann die Punkte $[\alpha_i, \alpha_i]$ als Graphen (*sections*) der konstanten Funktionen

$$f_i : \mathbb{R}^0 \cong \{x_0\} \longrightarrow \mathbb{R} \text{ mit } x_0 \longmapsto \alpha_i$$

für ein beliebiges x_0 betrachten.

Die offenen Intervalle $]\alpha_i, \alpha_{i+1}[$ zusammen mit den uneigentlichen Funktionen $f_0 : x_0 \longmapsto -\infty$ und $f_{m+1} : x_0 \longmapsto \infty$ sind *Sektoren* (*sectors*) dieser Funktionen, also Mengen der Form

$$\{(x_0, \alpha) \mid f_i(x_0) = \alpha_i < \alpha < \alpha_{i+1} = f_{i+1}(x_0)\}$$

Es ergibt sich eine kanonische Ordnung $c_1 < \dots < c_{2m+1}$ für diese Zellen.

Bevor die iterative Definition einer CAD für den Raum \mathbb{R}^r gegeben werden kann, ist eine wichtige Begriffserklärung nötig:

Definition 3.5. Stapel

Für ein zusammenhängendes Gebiet $\mathcal{A} \subset \mathbb{R}^k$ ist ein *Stapel* über \mathcal{A} eine Zerlegung

$$Z(\mathcal{A}) = \bigcup_{i=1, \dots, 2m+1} c_i$$

Kapitel 3 Quantorenelimination im \mathbb{R}^r

des Zylinders $Z(\mathcal{A}) := \mathcal{A} \times \mathbb{R} \subseteq \mathbb{R}^{k+1}$ (Zylinder über \mathcal{A}) mit der Eigenschaft, daß $\forall \alpha \in \mathcal{A}$ und $\acute{c}_i := c_i \cap (\alpha \times \mathbb{R})$ gilt:

$$\bigcup_{i=1, \dots, 2m+1} \acute{c}_i$$

ist eine CAD von \mathbb{R}^1 mit der kanonischen Ordnung $c_i < c_j \Leftrightarrow \acute{c}_i < \acute{c}_j$.

Wie im Fall \mathbb{R}^1 kann man die Zellen anhand ihres Index charakterisieren, da die ungerade indizierten Zellen den Sektoren (Intervallen) und die gerade indizierten Zellen den Graphen gewisser polynomialer Funktionen entsprechen. Seien umgekehrt

$$f_i : \mathbb{R}^k \supset \mathcal{A} \longrightarrow \mathbb{R}$$

Funktionen mit $f_1 < \dots < f_m$, dann ist die durch die (f_i, f_{i+1}) -Sektoren

$$\{(x, \alpha) \mid f_i(x) < \alpha < f_{i+1}(x)\} \subseteq Z(\mathcal{A})$$

sowie die f_i -Graphen

$$\{(x, \alpha) \mid f_i(x) = \alpha\}$$

induzierte Zerlegung von $Z(\mathcal{A})$ ein Stapel über \mathcal{A} .

Mit diesen Bemerkungen und Definitionen ist es möglich, eine CAD über dem Raum \mathbb{R}^r zu definieren:

Definition 3.6. CAD im \mathbb{R}^r

Eine algebraische Zerlegung D von \mathbb{R}^{i+1} heißt *zylindrisch* (i.e. CAD), wenn mit \acute{D} einer CAD von \mathbb{R}^i für alle Zellen $\acute{c} \in \acute{D}$ eine Menge $D_S \subseteq D$ existiert, welche einen Stapel über \acute{c} darstellt.

Die Zerlegung \acute{D} heißt von D *induziert*. Umgekehrt heißt D eine *Erweiterung* von \acute{D} .

Um Wahrheitswerte der Zellen einer Zerlegung definieren zu können, müssen die Polynome in der Formel auf jeder Zelle bezüglich ihres Vorzeichens invariant sein:

3.2 Zylindrische Algebraische Zerlegung im \mathbb{R}^r

Definition 3.7. Wahrheitswert und Invarianz

Eine Zelle c in einer Zerlegung heißt *vorzeicheninvariant* bezüglich einer Menge von Polynomen $P = (f_i)_{i \in I}$ (oder einfach P -invariant), wenn für alle $i \in I$ gilt:

$$\forall \alpha \in c \quad f_i(\alpha) \sim_i 0 \quad \text{für } \sim_i \in \{<, =, >\}$$

Man definiert den *Wahrheitswert* einer P -invarianten Zelle c bezüglich einer Aussage F in den Polynomen von P als

1. TRUE, falls $F|_c = \text{TRUE}$
2. FALSE, falls $F|_c = \text{FALSE}$

Die Aufgabe besteht nun in der Berechnung einer vorzeicheninvarianten CAD von \mathbb{R}^r bezüglich der Polynome in einer gegebenen Formel. Da Polynome stetige Funktionen sind, bietet sich die reelle Varietät der Polynome im Koordinatenraum als Hilfsmittel an, um maximale Gebiete mit Vorzeicheninvarianz zu erhalten. Dies führt zur nächsten Definition:

Definition 3.8. Delinearisierbarkeit

1. Eine Menge von Polynomen

$$\{f_1, \dots, f_m\} \subseteq \mathbb{R}[x_1, \dots, x_i]$$

heißt *delinearisierbar* über einem zusammenhängenden Gebiet $\mathcal{A} \subseteq \mathbb{R}^{i-1}$, wenn mit der reellen Varietät $V(f_1, \dots, f_m)$ der f_i gilt

$$V(f_1, \dots, f_m) \cap Z(\mathcal{A}) = \{s_1, \dots, s_k\}$$

wobei $s_i \cap s_j = \emptyset$ für $i \neq j$ und die s_i Graphen sind.

2. Für Polynome $F = (f_1, \dots, f_m)$ ist die *natürliche algebraische Zerlegung* von $Z(\mathcal{A})$ eine Zerlegung $Z(\mathcal{A}) = \bigcup_{j=1, \dots, l} R_j$ in maximale zusammenhängende Gebiete R_j , sodaß R_j f_i -invariant ist für alle i und j . Die Delinearisierbarkeit der $\{f_1, \dots, f_m\}$ bedeutet dann, daß die natürliche algebraische Zerlegung $(R_j)_j$ ein Stapel über \mathcal{A} ist.

Kapitel 3 Quantorenelimination im \mathbb{R}^r

Ein Beispiel für eine CAD des Polynoms $2x^4 - 3x^2y + y^4 - 2y^3 + y^2$ ist mithilfe von QEPCAD in Abbildung 3.1 gegeben. Man beachte, daß die Delinearisierbarkeit über der x -Achse erreicht wird, indem die kritischen Punkte der reellen Varietät (Singularitäten und Stellen mit vertikaler Tangente) im \mathbb{R}^2 auf \mathbb{R}^1 projiziert werden und diese als Intervallgrenzen für die zugrundeliegende CAD im \mathbb{R}^1 verwendet werden.

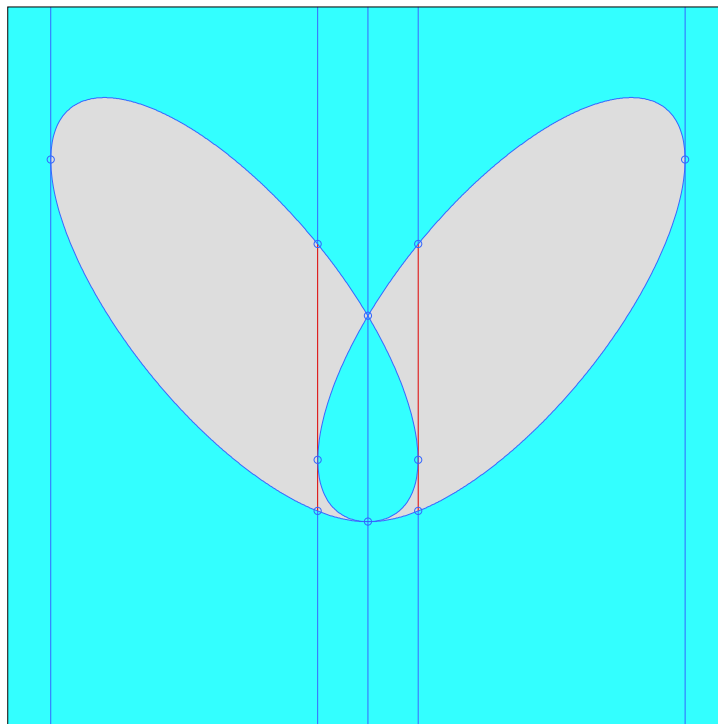


Abbildung 3.1: CAD für $2x^4 - 3x^2y + y^4 - 2y^3 + y^2$

Die Berechnung einer CAD besteht im wesentlichen aus drei Phasen: dem *Projektionsschritt*, dem *Basisschritt* und dem *Stapelschritt*:

3.2 Zylindrische Algebraische Zerlegung im \mathbb{R}^r

Input: F eine Formel von endlich vielen Polynomen $\acute{F} = \{f_i \mid i \in I\} \subset \mathbb{R}[x_1, \dots, x_r]$

Output: eine \acute{F} -invariante CAD (jede Zelle durch einen Punkt α (*sample point*) repräsentiert), eine Projektionsmenge $P \subset \mathbb{R}[x_1, \dots, x_r]$ und das Vorzeichen aller Polynome $p \in P$ auf jeder Zelle

1. (Projektionsschritt)

Berechne $Proj^i(\acute{F}) \subset \mathbb{R}[x_1, \dots, x_{r-i}]$ für alle $i = 1, \dots, r-1$

2. (Basis-Schritt)

Berechne eine $P_1 = Proj^{r-1}(\acute{F})$ -invariante CAD für \mathbb{R}^1

3. (Stapelschritt)

Errechne iterativ aus der CAD von \mathbb{R}^{i-1} eine CAD von \mathbb{R}^i

Bemerkung 3.9. Die Projektionsabbildung soll so definiert sein, daß

$$Proj : \mathbb{R}[x_1, \dots, x_i] \rightarrow \mathbb{R}[x_1, \dots, x_{i-1}]$$

für alle $i = r, \dots, 2$ mit $P_{i-1} = Proj(P_i)$ die Eigenschaft

$$\mathcal{A}_{i-1} \text{ ist } P_{i-1}\text{-invariant} \Rightarrow \mathcal{A}_i = \mathcal{A}_{i-1} \times \mathbb{R} \text{ ist } P_i\text{-invariant}$$

besitzt.

Die Abbildung $Proj$ ist nicht eindeutig und es gibt verschiedene Projektionsoperatoren, vergleiche [Hon98a, McC98, Col75].

Hier wird die Projektion aus [Col75] erklärt:

Definition 3.10. Projektion nach Collins

Betrachte den Polynomring nach Umbenennung von x_i als

$$\mathbb{R}[x_1, \dots, x_{i-1}][x_i] = J[x]$$

mit nur einer Variable $\forall r \geq i \geq 2$.

Dann definiere für eine Menge von Polynomen $A = \{A_1, \dots, A_k\} \subset J[x]$ die Abbildung $Proj$:

$$Proj(A) := Proj_1(A) \cup Proj_2(A)$$

Kapitel 3 Quantorenelimination im \mathbb{R}^r

mit

$$Proj_1(A) := \bigcup_{i=1}^k \bigcup_{p \in R_i} (\{lc(p)\} \cup PSC(p, p'))$$

und

$$Proj_2(A) := \bigcup_{1 \leq i < j \leq k} \bigcup_{p \in R_i, q \in R_j} PSC(p, q)$$

unter Beachtung der Notationen:

1. die Mengen R_i sind definiert als

$$R_i := RED(A_i) := \{red^l(A_i) \mid 0 \leq l \leq deg(A_i) \wedge red^k(A_i) \neq 0\}$$

mit der induktiven Reduktion $red^{l+1}(A_i) = red^l(A_i)$ und $red(A_i) = A_i - lt(A_i)$.

2. die Menge $PSC(p, q)$ ist

$$PSC(p, q) := \{psc_j(p, q) \mid 0 \leq j < \min(deg(p), deg(q)) \wedge psc_j(p, q) \neq 0\}$$

für die j -te prinzipielle Subresultante $psc_j(p, q)$, die definiert ist als

$$psc_j(p, q) := \det(M_{j,j})$$

für die Matrix $M_{j,j}$, die mit $deg(p) = n, deg(q) = m$ aus der Sylvester-Matrix $M = Syl(p, q, x) \in J^{(m+n) \times (m+n)}$ durch Streichung der jeweils letzten (j) Zeilen der Koeffizienten von p und q und der letzten $(2j + 1)$ Spalten mit Ausnahme der $(m + n - 2j)$ -ten Spalte entsteht.

Daß diese Abbildung die Eigenschaft in 3.9 gilt, ist ein Resultat aus [ACM98a]:

Theorem 3.11. *Wenn für einen beliebigen Index $r \geq 2$ ein zusammenhängendes Gebiet $\mathcal{R} \subset \mathbb{R}^{r-1}$ bezüglich der Menge $Proj(A)$ invariant ist und $A \subset R[x_1, \dots, x_r]$, dann ist A über \mathcal{R} entweder delinearierbar oder identisch mit dem Nullpolynom auf $Z(\mathcal{R})$ und für zwei Polynome $p, q \in A$ sind die Graphen entweder disjunkt oder identisch. \square*

Im **Projektionsschritt** erhält man mit $P_i := Proj^{r-i}(\acute{F})$ somit durch iterative Anwendung von $Proj$ die Polynomengen

$$P_{r-1} = Proj^1(\acute{F}) \subset \mathbb{R}[x_1, \dots, x_{r-1}], \dots, P_1 = Proj^{r-1}(\acute{F}) \subset \mathbb{R}[x_1]$$

mit der gewünschten Eigenschaft. Definiere $P(l) := (\bigcup_{i=1}^l P_i) \subset \mathbb{R}[x_1, \dots, x_l]$. Nach hinreichend vielen Iterationen ist man im Ring $\mathbb{R}[x_1]$ und beginnt den **Basisschritt**: da nur noch eine Variable vorliegt

$$P_1 = Proj^{r-1}(\acute{F}) \in \mathbb{R}[x_1]$$

kann man reelle Nullstellen durch geeignete Verfahren bestimmen und errichtet zusammen mit den zugehörigen offenen Intervalle eine P_1 -invariante CAD für \mathbb{R}^1 . Die besondere Konstruktion von $Proj$ gewährleistet, daß man nun im **Stapelschritt** rekursiv aus dieser CAD von \mathbb{R}^1 Zerlegungen der Räume höherer Dimension errechnen kann, die ebenfalls invariant sind:

- 1.) Berechne für $i = 2, \dots, r$ die Wahrheitswerte der Zellen bezüglich F , sowie die Vorzeichen der Polynome $p \in P_{i-1}$ auf diesen Zellen. Wegen der Invarianz reicht ein einziger repräsentativer Versuchspunkt $\acute{\alpha}$ dafür aus.
- 2.) Wenn nun $\acute{\alpha} = (\alpha_1, \dots, \alpha_{i-1})$ ein Versuchspunkt einer Zelle $c \in D_{i-1}$ und D_{i-1} eine P_{i-1} -invariante CAD von \mathbb{R}^{i-1} ist, dann erhalte durch Einsetzung dieser Punkte in die ersten $i-1$ Variablen für alle $p \in P(i)$ mittels $p(\alpha_1, \dots, \alpha_{i-1}, x)$ Polynome, die nur von einer Variablen abhängen, somit können wie im Basisschritt Zellen $c \in D_i$ einer $P(i)$ -invarianten CAD von \mathbb{R}^i und ein Versuchspunkt $\alpha = (\alpha_1, \dots, \alpha_{i-1}, \alpha_i)$ für jede dieser Zellen in D_i erhalten werden.

3.3 Lösungsformel

Hat man eine CAD berechnet, wird nun die Konstruktion einer Lösungsformel die Aufgabe sein. Grob vereinfacht werden Polynome gesucht, die nur von den freien Variablen abhängen und in einer Formel denselben Wahrheitswert auf jeder Zelle des Raumes wie die Ausgangsformel haben. Das nächste Theorem aus [Hon98b] zeigt die Idee:

Kapitel 3 Quantorenelimination im \mathbb{R}^r

Theorem 3.12. Betrachte eine Formel F in den Polynomen $A = \{f_i \mid i \in I\} \subset \mathbb{R}[x_1, \dots, x_r]$ und mit $Q_i \in \{\forall, \exists\}$ die Aussage

$$F = (Q_1 x_{k+1})(Q_2 x_{k+2}) \dots (Q_{r-k} x_r) [\acute{F}(x_1, \dots, x_k, x_{k+1}, \dots, x_r)]$$

Es sei D_k eine bezüglich der Polynome in A invariante CAD für den Raum \mathbb{R}^k . Mit den Wahrheitswerten der Zellen $V_T(c) = F \mid_c \in \{\mathbf{TRUE}, \mathbf{FALSE}\}$ definiere die Vereinigung aller Zellen, auf denen F wahr ist:

$$C_k = \{c \in D_k \mid V_T(c) = \mathbf{TRUE}\} \quad \text{sowie} \quad T_k = \bigcup_{c \in C_k} c$$

Dann gilt:

$$\begin{aligned} (Q_1 x_{k+1})(Q_2 x_{k+2}) \dots (Q_{r-k} x_r) [\acute{F}(\alpha_1, \dots, \alpha_k, x_{k+1}, \dots, x_r)] \\ \iff (\alpha_1, \dots, \alpha_k) \in T_k \end{aligned}$$

□

Es reicht offensichtlich nach obigem Theorem, eine definierende Formel für T_k zu finden.

Definition 3.13. Wenn die definierende Formel aus Polynomen

$$p \in P_1 \cup \dots \cup P_k$$

gewonnen werden kann, heißt die Lösung *projektionsbasiert*.

Über das Verhalten des Algorithmus zu Quantoreneliminierung mittels CAD in der Praxis können folgende Aussagen getroffen werden:

Bemerkung 3.14.

1. Der spezielle Fall der projektionsbasierten Lösung tritt laut [Bro99] oft ein. Im allgemeinen Fall müssen jedoch erweiterte Verfahren hinzugezogen werden.
2. Die Quantoreneliminierung ist von einer doppelt exponentiellen Komplexität bezüglich der Variablenanzahl (vergleiche hierzu [Bro99, FR74, DH88]).

Kapitel 4

Algebraische Methoden für Finite Differenzenschemata

In diesem Kapitel werden Ergebnisse präsentiert, welche die Erzeugung von finiten Differenzenverfahren einer Gleichung mithilfe der Gröbnerbasentheorie erlauben. Es werden an die Art der zu approximierenden Gleichung keine spezifischen Anforderungen gestellt, sodaß dieser Ansatz ein sehr allgemeiner ist. Ein Kriterium für die Existenz eines Differenzenschemas wird mit Blick auf die Besonderheiten der verwendeten Approximationen und gewisse Eigenschaften der zugrundeliegenden Monomordnung gegeben. Für eine Gleichung werden verschiedene, aber zum gleichen Ergebnis führende Methoden angegeben, um ein Differenzenschema zu erzeugen.

Ein Ausblick auf den Fall von Systemen und die dortige Erzeugung von Differenzenschemata wird ebenfalls erfolgen.

4.1 Motivation

Betrachte wie in Kapitel 2 das Lax-Friedrich Verfahren

$$u_i^{j+1} = \frac{1}{2} \left(u_{i+1}^j + u_{i-1}^j \right) - \frac{ak}{2h} \left(u_{i+1}^j - u_{i-1}^j \right)$$

auf einem entsprechenden Gitter. Ansätze wie in [GBM06] regen die Formulierung eines solchen Verfahrens durch Differenzenoperatoren T_x, T_t an, die auf den Argumenten der Funktion $u(x, t)$ operieren:

$$T_x \circ u(x, t) = u(x + h, t)$$

für die x -Komponente und

$$T_t \circ u(x, t) = u(x, t + k)$$

für t .

Das komplette Lax-Friedrich-Verfahren gleicht in Operatorschreibweise dann dem Ausdruck

$$(T_t T_x - (\frac{1}{2}(T_x^2 + 1) - \frac{ak}{2h}(T_x^2 - 1))) \circ u(x, t) = 0$$

Für das naive Verfahren

$$u_i^{j+1} = u_i^j - \frac{ak}{2h}(u_{i+1}^j - u_{i-1}^j)$$

schreibt sich analog als

$$T_t T_x = T_x - \frac{ak}{2h}(T_x^2 - 1) \in \mathbb{K}[T_x, T_t]$$

Im weiteren Verlauf dieses Kapitels wird die oben skizzierte Herangehensweise formalisiert und theoretische Aspekte erläutert.

4.2 Theorie der algebraischen Methode

Um ein finites Differenzenverfahren zu generieren, bedarf es neben der partiellen Differentialgleichung auch geeigneter Approximationen für die auftretenden Ableitungen. *Geeignet* heißt in diesem algebraischen Kontext, daß die Erzeugung des Differenzenschemas möglich ist. Es werden Beispiele gezeigt, die aus algebraische Begründung heraus keine Erzeugung zulassen. In diesem theoretischen Teil ist es nicht die Intention, die Stabilität als ein Kriterium für die Eignung der Approximation herauszustellen. Dieses ist vielmehr ein praktisches Problem, welches für lineare partielle Differentialgleichungen gleichwohl durch reelle algebraische Geometrie und Quantorenelimination mittels CAD-Algorithmus zu lösen möglich ist. In diesem Abschnitt spielen Stabilitätsbetrachtungen keine tragende Rolle, obwohl dies für den Praktiker sicher wichtig ist.

4.2.1 Kontinuierlicher Fall

Es sei eine beliebige endliche Menge S partieller Differentialgleichungen

$$S = \left\{ \sum_{\gamma \in \Gamma_\alpha} (c_{\alpha\gamma} \cdot \prod_{\beta \in B_\gamma} u_{x^\beta}) \mid \alpha \in \Psi \right\}$$

für eine unbekannte Lösungsfunktion u . Es sind somit auch Gleichungen, die nichtlinear in den Ableitungen u_{x^β} sind, möglich. Mit

$$u_{x^\beta} := \frac{\partial^\beta}{\partial x^\beta} u := \frac{\partial^{|\beta|}}{\partial_{x_1}^{\beta_1} \dots \partial_{x_l}^{\beta_l}} u$$

seien die Ableitungen beliebiger Ordnungen bezeichnet, die Koeffizienten $c_{\alpha\gamma}$ seien invertierbar und der Einfachheit wegen sei nur eine Gleichung, also der Fall $|S| = 1$, betrachtet. Eine kurze Diskussion für den Fall mehrerer Gleichungen wird am Ende des Kapitels in Theorem 4.28 stattfinden.

In der Menge Λ seien alle auftretenden Ableitungen u_{x^β}

$$\Lambda := \bigcup_{\alpha \in \Psi} \bigcup_{\gamma \in \Gamma_\alpha} \bigcup_{\beta \in B_\gamma} \{u_{x^\beta}\}$$

Diese Ableitungen $u_{x^{\alpha_i}}$ sollen Variablen in einem Polynomring über einem Körper \mathbb{F} sein. \mathbb{F} sei eine algebraische Körpererweiterung über \mathbb{Q} mit Transzendenzgrad 0. Dies ist für die Berechenbarkeit aller auftretenden Objekte eine wesentliche Voraussetzung.

Vom prinzipiellen Standpunkt aus sind alle Koeffizienten $c_{\alpha\gamma}$ in S als Funktionen zugelassen, zum Beispiel $c_{\alpha\gamma} \in C_1$ mit $C_1 = C^\infty$. Allerdings werden später aus Gründen der Nichtkommutativität im Falle von nicht-konstanten Koeffizienten bestimmte Einschränkungen vorgenommen.

Mit den Variablen $\Lambda = \{u_{x^{\alpha_1}}, \dots, u_{x^{\alpha_l}}\}$ sei der Ring D_0 via

$$D_0 := \mathbb{F} \otimes_{\mathbb{F}} C_1[\Lambda] = \mathbb{F} \otimes_{\mathbb{F}} C_1[u_{x^{\alpha_1}}, \dots, u_{x^{\alpha_l}}]$$

definiert. Diese Definition mit obigen Bemerkungen garantiert aus algebraischer Sicht, daß alle Parameter invertierbar sind und Termumformungen vorgenommen werden können.

Bemerkung 4.1.

1. D_0 ist nicht unbedingt ein *differentieller Ring*, das heißt, ein Ring mit einer existierenden Abbildung $\partial : D_0 \rightarrow D_0$, sodaß für alle $d_1, d_2 \in D_0$ gilt:

1. $\partial(d_1 + d_2) = \partial(d_1) + \partial(d_2)$
2. $\partial(d_1 \cdot d_2) = \partial(d_1) \cdot d_2 + d_1 \cdot \partial(d_2)$

Aber D_0 ist sicher ein Unterring eines differentiellen Ringes.

2. Ein Beispiel für einen differentiellen Ring wäre $(C_1, (\partial_{x_i})_i)$ für $C_1 = C^\infty(\mathbb{F}, \mathbb{F})$ mit den partiellen Differentialoperatoren ∂_i .

3. Man beachte, daß die Menge $\Lambda = \{u_{x^{\alpha_1}}, \dots, u_{x^{\alpha_l}}\}$ aus formalen Ableitungen besteht, diese jedoch einfach als Variablen im Ring D_0 anzusehen sind:

$$\{u_{x^{\alpha_1}}, \dots, u_{x^{\alpha_l}}\} = \{z_1, \dots, z_l\}$$

4.2.2 Diskreter Fall

Es werden Funktionen u betrachtet, die als Definitionsbereich \mathbb{F}^n und als Zielbereich \mathbb{F} besitzen:

$$u : \mathbb{F}^{n+1} \longrightarrow \mathbb{F}$$

$$(x_1, \dots, x_n, t) \longmapsto u(x_1, \dots, x_n, t)$$

Bemerkung 4.2. Dies stellt keine Einschränkung dar, da ohnehin für praktische Zwecke die Verwendung von transzendenten Zahlen nicht möglich ist.

Um ein unendliches Gitter auf dem Definitionsbereich einzuführen, definiert man für positive $\Delta x_1, \dots, \Delta x_n, \Delta t$:

$$\Omega := (\mathbb{Z}\Delta x_1 \times \dots \times \mathbb{Z}\Delta x_n \times \mathbb{Z}\Delta t) \cap \mathbb{R}^{n+1}$$

Nun sind die Argumente der Funktion u im \mathbb{R} -Raum \mathbb{R}^{n+1} und dieser ist kein endlich erzeugter \mathbb{Z} -Modul, sodaß erst die Definition von Ω die kanonische Isomorphie als \mathbb{Z} -Modul sicherstellt:

$$\mathbb{Z}^{n+1} \cong \Omega$$

Die Idee zur Definition von Differenzenoperatoren T_i ist die Wirkung ebendieser auf die Argumente einer Funktion:

$$T_{x_i} \circ u(x_1, \dots, x_i, \dots, x_n, t) = u(x_1, \dots, x_i + \Delta x_i, \dots, x_n, t)$$

für $(x_1, \dots, x_i, \dots, x_n, t) \in \Omega$.

Für ein $\delta \in \mathbb{Z}^{n+1}$ mit $x_{\delta_i} = \delta_i \Delta x_i \forall i$ sowie $t_{\delta_{n+1}} = \delta_{n+1} \Delta t$ und einer Ableitung u_{x^α} ist durch die Identifizierung

$$u_{x^\alpha}(x_{\delta_1}, \dots, x_{\delta_n}, t_{\delta_{n+1}}) \stackrel{\text{approx}}{\approx} (u_{x^\alpha})^\delta$$

die Grundlage geschaffen, daß für eine beliebige gegebene Menge von Ableitungen $\{u_{x^\theta} | \theta \in \Theta\}$ die daraus resultierende Menge aller möglichen Gitterwerte der *Funktionswerte* (nicht der Argumente x_1, \dots, x_n, t) definiert wird:

$$D_1 := \mathbb{F}[\{(u_{x^\theta})^\delta \mid \theta \in \Theta, \delta \in \mathbb{Z}^{n+1}\}]$$

Die Wirkung der Differenzenoperatoren wird nun kanonisch auf die unendliche Menge D_1 übertragen:

Definition 4.3. Differenzenoperatoren

Die Differenzenoperatoren $T_i : D_1 \rightarrow D_1$ genügen nach Definition den folgenden Eigenschaften:

1. Für $\delta \in \mathbb{Z}^{n+1}$ mit $x_{\delta_i} = \delta_i \Delta x_i$ und $t_{\delta_{n+1}} = \delta_{n+1} \Delta t$ ist die Wirkung des i -ten Differenzenoperators T_i für ein beliebiges $u_{x^\alpha} \in D_1$

$$T_i : u_{x^\alpha}(x_{\delta_1}, \dots, x_{\delta_i}, \dots, x_{\delta_n}, t_{\delta_{n+1}}) \mapsto u_{x^\alpha}(x_{\delta_1}, \dots, x_{\delta_{i+1}}, \dots, x_{\delta_n}, t_{\delta_{n+1}})$$

2. Für $\delta, \gamma \in \mathbb{Z}^{n+1}$ und $T^\gamma = \prod_i^{n+1} T_i^{\gamma_i}$ ist die Wirkung des zusammengesetzten Differenzenoperators T^γ

$$T^\gamma \circ (u_{x^\alpha})^\delta = (u_{x^\alpha})^{\delta+\gamma}$$

3. Die Operatoren sind *multiplikativ*, das heißt $\forall (u_{x^\alpha})^{\mu_1}, (u_{x^\beta})^{\mu_2} \in D_1$ gilt

$$T^\gamma \circ ((u_{x^\alpha})^{\mu_1} \cdot (u_{x^\beta})^{\mu_2}) = T^\gamma \circ ((u_{x^\alpha})^{\mu_1}) \cdot T^\gamma \circ ((u_{x^\beta})^{\mu_2})$$

Eingedenk dieser Eigenschaft wird ab jetzt die Operatorwirkung nicht mehr explizit mit \circ gekennzeichnet.

Nach den Definitionen sind einige Bemerkungen nützlich:

Bemerkung 4.4.

1. Die Einführung eines formal unendlichen Gitters erlaubt es auf einer hinreichend großen Menge zu operieren, ohne daß Randbedingungen bei der Generierung von Differenzenschemata berücksichtigt werden müssen.

2. Es sind in der Definition der Differenzenoperatoren auch negative Exponenten erlaubt. Damit ist es für einen repräsentativen generischen Punkt δ möglich, durch Wirkung eines geeigneten Operators jeden beliebigen Punkt zu erreichen. Die Abbildung 4.1 zeigt ein passendes Beispiel.

3. Die Identifizierung $u_{x^\alpha}(x_{\delta_1}, \dots, x_{\delta_n}, t_{\delta_{n+1}}) \approx (u_{x^\alpha})^\delta$ erkennt keine Abschneidefehler und ist eher philosophischer Natur, da numerische Aspekte nicht in hohem Maße berücksichtigt werden. Es wird naiv von einer guten Approximation ausgegangen.

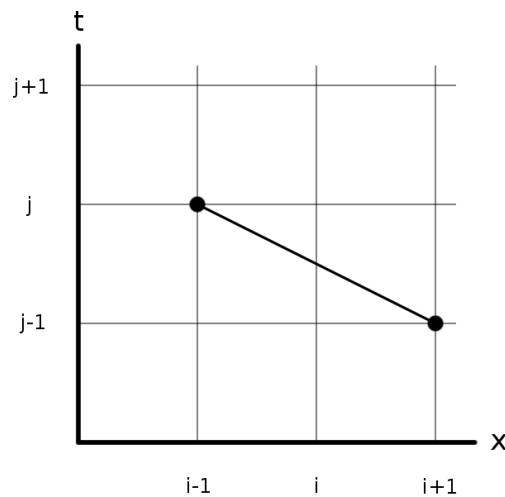


Abbildung 4.1: Wirkung des Differenzenoperators $(T_t T_x^{-2}) \circ (w_{i+1}^{j-1}) = w_{i-1}^j$

Zu einer Menge von partiellen Differentialgleichungen (als *kontinuierliche* Funktionen)

$$S = \left\{ \sum_{\gamma \in \Gamma_\alpha} (c_{\alpha\gamma} \cdot \prod_{\beta \in B_\gamma} u_{x^\beta}) \mid \alpha \in \Psi \right\} \subset D_0$$

mit den auftretenden Variablen

$$\Lambda = \{u_{x^{\alpha_1}}, \dots, u_{x^{\alpha_l}}\}$$

sei eine beliebige endliche Menge von *diskreten* Approximationen gegeben:

$$A^\delta = \{(u_{x^\lambda})^\delta - \sum_{\omega \in \Omega_\lambda} (c_\omega \cdot \prod_{\nu \in N_\omega} T^{\gamma_\nu}(u_{x^\nu})^\delta) \mid \lambda \in \Phi, c_\omega \neq 0 \ \forall \omega \in \Omega_\lambda\}$$

Auch hier sind prinzipiell nichtlineare Terme unter dem Summenzeichen erlaubt. Mit Blick auf diese Approximationen betrachte die dort auftretenden Ableitungen der Form $(u_{x^\varsigma})^\delta$:

$$M^\delta = \{(u_{x^{\beta_1}})^\delta, \dots, (u_{x^{\beta_m}})^\delta\} = \bigcup_{\lambda \in \Phi} \bigcup_{\omega \in \Omega_\lambda} \bigcup_{\nu \in N_\omega} \{(u_{x^\nu})^\delta\} \cup \{(u_{x^\lambda})^\delta\}$$

Dies sind *alle* formalen Ableitungen in A^δ , jedoch ist für die im Folgenden verwendeten Begriffe eine Untermenge $L^\delta \subseteq M^\delta$ interessant, nämlich die Menge aller Terme auf der *linken* Seite in A^δ :

$$L^\delta = \bigcup_{\lambda \in \Phi} \{(u_{x^\lambda})^\delta\}$$

Bemerkung 4.5.

1. Diese Menge L^δ kann verstanden werden als die Menge aller Ableitungen in A^δ , die wiederum durch andere Ableitungen approximiert werden. Diese Vorstellung ist nützlich für zukünftige Betrachtungen.

2. Man beachte, daß die Menge aller Variablen $\Lambda = \{u_{x^{\alpha_1}}, \dots, u_{x^{\alpha_l}}\}$ im kontinuierlichen Ring D_0 sind, jedoch die Approximationen in einer diskreten Struktur auftauchen.

Es wird für Berechnungen mit den Approximationen in A^δ ein geeigneter Ring gesucht. Daher wird an späterer Stelle eine Abbildung φ vorgestellt, die kontinuierliche Objekte auf ihre diskreten Analoga abbildet. Diese Abbildung bezweckt, daß die Bilder der Menge S und Λ unter φ in derselben diskreten Struktur auftreten wie die Approximationen A^δ .

Im kontinuierlichen Ring

$$D_0 = \mathbb{F} \otimes_{\mathbb{F}} C_1[\Lambda]$$

werden die auftretenden Koeffizienten durch C_1 beschrieben, können jedoch nach Abhängigkeit von x_1, \dots, x_n, t unterteilt werden in konstante Koeffizienten $c' \in C_1^c$ und nicht-konstante Koeffizienten $c'' \in C_1^d$. Bezüglich der Differentiation bilden die konstanten Koeffizienten einen Unterring $C_1^c \subseteq C_1$, sodaß gilt

$$\partial^\alpha \circ c = 0 \quad \forall c \in C_1^c \quad \forall \alpha \in \mathbb{N} \setminus \{0\}$$

Entsprechend bilden die nicht-konstanten Koeffizienten C_1^d eine \mathbb{F} -Algebra. Betrachte $C_1 = C_1^c \otimes_{\mathbb{F}} C_1^d$, dann ist C_1 eine C_1^c -Algebra: jedes $c \in C_1$ kann als endliche Summe $c = \sum c_i \otimes t_i$ mit $c_i \in C_1^c$ und $t_i \in C_1^d$ geschrieben werden. Ein Beispiel hierzu kann leicht angegeben werden:

Beispiel 4.6. Betrachte den Term $(2 \cdot a + 5) \cdot \sin(t) + (3 \cdot b) \cdot \cos(t)$ mit der Relation $b^2 + 1 = 0$. Dann definiere C_1^d und C_1^c mit

$$C_1^d = \mathbb{F}(\cos(t), \sin(t))$$

für die nicht-konstanten Funktionen und

$$C_1^c = \mathbb{F}(a)[b]/\langle b^2 + 1 \rangle$$

für die konstanten Funktionen. Es ergibt sich

$$(2 \cdot a + 5) \otimes \sin(t) + (3 \cdot b) \otimes \cos(t) \in C_1^c \otimes_{\mathbb{F}} C_1^d.$$

Bemerkung 4.7. Das Verhalten der Elemente in C_1^c und C_1^d unter der Differentiation ist sicherlich unterschiedlich. Somit ist auch in Bezug auf die Anwendung von Differenzenoperatoren eine Unterscheidung dieser beiden Typen von Koeffizienten erforderlich, da formal eine besondere Beziehung zwischen Differenzen- und Differentialoperatoren besteht:

$$\partial_x = \lim_{\Delta x \rightarrow 0} \frac{T_x - 1}{\Delta x}$$

Somit ergibt sich die Definition der Diskretisierungsabbildung:

Definition 4.8. Diskretisierung

Mit obigen Bezeichnungen definiere

$$\begin{aligned} \varphi : C_1 = C_1^c \otimes_{\mathbb{F}} C_1^d &\longrightarrow \{\phi \mid \phi : \Omega \rightarrow \mathbb{R}\} \\ c' \otimes c'' &\longmapsto c' \cdot \delta'(c'') \end{aligned}$$

wobei $\delta'(c'') = c''(\delta)$ der Wert von c'' im generischen Punkt $\delta \in \mathbb{Z}^{n+1}$ ist.

Bemerkung 4.9.

1. Die Abbildung φ ist wohldefiniert, da die Koeffizienten bekannt sind.
2. Der Definitionsbereich der Bilder unter φ ist allerdings ein anderer.

Es wird nun ein diskreter Ring D_2 definiert, der sowohl die ursprüngliche Differentialgleichungen als auch die Approximationen jeweils zusammen mit allen auftretenden Ableitungen und Koeffizienten berücksichtigt:

Definition 4.10. Es sei mit $\Lambda^\delta = \varphi(\Lambda)$ das Bild unter φ für einen symbolischen Punkt δ und mit C_2 die diskreten Koeffizienten in den Approximationen A^δ , dann definiere den diskreten Ring D_2 mittels

$$D_2 = \mathbb{F} \otimes_{\mathbb{F}} C_1^\delta \otimes_{\mathbb{F}} C_2(T_1, \dots, T_n, T_{n+1}) [\Lambda^\delta \cup M^\delta]$$

oder kürzer

$$D_2 = \mathbb{F} \otimes_{\mathbb{F}} C_1^\delta \otimes_{\mathbb{F}} C_2(\underline{T}) [\Lambda^\delta \cup M^\delta]$$

Diese Definition gibt Anlass zu einer Bemerkung:

Bemerkung 4.11.

1. Es ist die Philosophie zu erkennen, den Ring D_2 durch Einführung der Differenzenoperatoren möglichst *groß* zu machen, dann reicht umgekehrt *ein* generischer Punkt alleinig zur Beschreibung aller anderen Punkte mittels der Differenzenoperatoren aus (vgl. Bemerkung 4.4).
2. Da die Operatoren \underline{T} nach Definition von D_2 auch in Brüchen auftreten können, entsteht der günstige Fall, daß

$$D_2 = \mathbb{F} \otimes_{\mathbb{F}} C_1^\delta \otimes_{\mathbb{F}} C_2(\underline{T}) [\Lambda^\delta \cup M^\delta] = \mathbb{K} [\Lambda^\delta \cup M^\delta]$$

für einen Körper \mathbb{K} . Diese Tatsache wird im weiteren Verlauf eine sehr wichtige sein.

Allerdings werden diese Zugeständnisse an die Eigenschaften der \underline{T} nur formaler Natur sein, da in der Praxis für den Fall, daß Operatoren in einem Nenner sind, durch Multiplikation mit dem kleinsten gemeinsamen Vielfachen aller Nenner etwaige Operatoren in den Nennern beseitigt werden.

Bis zu dieser Stelle sind hinreichend viele Begriffe definiert worden, daß nun ein entscheidendes algebraisches Kriterium für die Existenz (später kommt noch die Eindeutigkeit) von Differenzenschemata formuliert werden kann. Das Konzept einer monomialen Ordnung in Definition 1.12 spielt die wichtigste Rolle, da die Existenz einer geeigneten Ordnung im Ring D_2 gefordert wird:

Definition 4.12. Die Menge $A^\delta \subset D_2$ heißt *hinreichend zur Erzeugung eines Verfahrens*, wenn eine (globale) Ordnung $>$ auf den Monomen Mon_{D_2} des Ringes D_2 existiert mit

$$(u_{x^\lambda})^\delta > \prod_{\nu \in N_\omega} (u_{x^\nu})^\delta \quad \forall \lambda \in \Phi \quad \forall \omega \in \Omega_\lambda$$

Man sagt auch, daß die Ordnung $<$ *hinreichend zur Erzeugung eines Verfahrens* ist, wenn klar ist, um welche Approximationen es sich handelt.

Dies heißt anschaulich, daß eine Ordnung existieren muß, sodaß in den Approximationen A^δ die Terme auf der linken Seite stets größer sind als diejenigen, die in der Summe auf der rechten Seite auftreten. Dies ist im Allgemeinen nicht selbstverständlich, wie das folgende Korollar und das anschließende Beispiel zeigen:

Korollar 4.13. Kriterium für Approximationen

Wenn die Ordnung $<$ auf $\Lambda^\delta \cup M^\delta$ hinreichend im obigen Sinne ist, dann gilt

$$|L^\delta| < |M^\delta|$$

Es ergibt sich somit die notwendige Bedingung $|M^\delta \setminus L^\delta| \geq 1$.

Beweis. Da ohnehin $L^\delta \subseteq M^\delta$ gilt, muß die Striktheit dieser Enthaltung nachgewiesen werden. Wenn aber $L^\delta = M^\delta$ der Fall wäre, führte dies zu einem Widerspruch hinsichtlich der Transitivitätseigenschaft von $<$. \square

Beispiel 4.14 zeigt die Argumentation von Korollar 4.13:

Beispiel 4.14. Betrachte für eine partielle Differentialgleichung

$$S^\delta = \{u_{tt} + a \cdot u_{xx}\}$$

die Approximationen

$$A^\delta = \{u_{tt} - (u_x + u_t), u_t - (u + u^2), u_{xx} - (u_x + u), u_x - (u + u^3), u - (u_t + u_x)\}$$

Dann gilt, daß

$$L^\delta = \{u_{tt}, u_t, u_{xx}, u_x, u\} = M^\delta$$

Nun müssen das zweite und letzte Element in A^δ die Relation

$$u_t > u > u_t$$

erfüllen. Dies ist aber ein Widerspruch zur Transitivität von $<$.

Daß die Umkehrung der Implikation in Korollar 4.13 im Allgemeinen nicht gilt, beweist das Beispiel 4.15:

Beispiel 4.15. Es sei

$$S^\delta = \{u_t + a \cdot u_x\}$$

mit den Approximationen

$$A^\delta = \{u_t - (u_x + u_x \cdot u), u_x - (u_t + u^2)\}$$

Man sieht, daß

$$L^\delta = \{u_t, u_x\} \subsetneq \{u_t, u_x, u\} = M^\delta.$$

Wegen 4.13 ist die Existenz einer Ordnung $<$ mit

$$u_t > u_x \quad \text{and} \quad u_x > u_t$$

notwendig. Auch dies ist nicht möglich.

Bemerkung 4.16. Im Falle von nicht-konstanten Koeffizienten und damit Funktionen, die von x_1, \dots, x_n, t abhängen, müssen bezüglich der Wirkung der Differenzenoperatoren einige Besonderheiten beachtet werden: die Multiplikativität der Operatoren

$$T \circ (u_1 \cdot u_2) = T \circ u_1 \cdot T \circ u_2$$

war in der Definition nur für je zwei Variablen $u_1 = (u_\alpha)^{\delta_1}$ und $u_2 = (u_\beta)^{\delta_2}$ gefordert. Es ist natürlich, diese multiplikative Eigenschaft auch in der Wirkung auf beliebige Ringelemente in D_2 zu fordern.

Betrachte somit zwei Elemente $u_1 = x_k$, die identische Abbildung auf die k -te räumliche Variable, und $u_2 = u_{x_k}$ für eine Funktion u sowie $T = T_k$. Die kanonische Diskretisierung der nicht-konstanten Funktion x_k in einem beliebigen Punkt $a = (i_1, \dots, i_{n+1})$ des Gitters Ω ist mit $x_k(a) = i_k$ für das k -te Argument gegeben. Ohne Unterscheidung zwischen \cdot und \circ im Ring D_2 ergibt sich im kontinuierlichen Fall

$$T \cdot (x \cdot u_x) = (T \cdot x) \cdot (T \cdot u_x) = (x + \Delta x) \cdot T \cdot u_x$$

und ebenso für den diskreten Fall

$$T_k \cdot (i_k \cdot u_{x_k}) = (T_k \cdot i_k) \cdot (T_k \cdot u_{x_k}) = (i_k + 1) \cdot T_k \cdot u_{x_k}$$

Damit ist die Multiplikation der Operatoren T_k (als Elemente im Ring D_2 betrachtet) mit nicht-konstanten Funktionen im Allgemeinen nicht kommutativ und es gilt

$$T_k \cdot i_k = (i_k + 1) \cdot T_k \quad \forall k \in \{1, \dots, n + 1\}$$

Für beliebige nicht-konstante Funktionen ist diese Argumentation übertragbar und führt zum gleichen Ergebnis.

Die Multiplikation von Operatoren mit Konstanten ist aus offensichtlichen Gründen nicht davon betroffen.

Die Einschränkung auf konstante Koeffizienten umgeht diesen unangenehmen Fall aus obiger Bemerkung und es soll ab jetzt gelten: $C_1 = C_1^c$.

Ziel ist nun die Definition eines finiten Differenzenschemas mit algebraischen Begriffen:

Definition 4.17. Es seien S^δ , A^δ und die Mengen M^δ , L^δ wie oben und eine monomiale Ordnung $<$ auf dem Differenzen-Ring

$$D_2 = \mathbb{F} \otimes_{\mathbb{F}} C_1^\delta \otimes_{\mathbb{F}} C_2(\underline{T}) [\Lambda^\delta \cup M^\delta]$$

Es sei darüber hinaus die Menge A^δ hinreichend zur Erzeugung eines Verfahrens zusammen mit der Ordnung $<$. Dann heißt das Ideal D , welches definiert wird durch

$$D := (\langle S^\delta \rangle_{D_2} + \langle A^\delta \rangle_{D_2}) \cap \mathbb{F} \otimes_{\mathbb{F}} C_1^\delta \otimes_{\mathbb{F}} C_2(\underline{T}) [(\Lambda^\delta \cup M^\delta) \setminus L^\delta]$$

ein *Finites-Differenzenschema-Ideal*.

Der Einfachheit halber wird ab jetzt von einem *Differenzenschema* gesprochen, wenn der Kontext unmißverständlich ist.

Bemerkung 4.18. Es wird nicht angenommen, daß die Ideale $\langle S^\delta \rangle_{D_2}$ und $\langle A^\delta \rangle_{D_2}$ *Differenzen-Ideale* sind, das sind Ideale I mit

$$\Delta_i \circ f \in I \quad \forall f \in I \forall \Delta_i$$

wobei Δ_i die auftretenden Differenzenoperatoren sind. Differenzen-Ideale sind nach Definition abgeschlossen unter der Wirkung dieser Operatoren. Ebenso ist auch $D \trianglelefteq D_2$ nicht notwendigerweise ein Differenzen-Ideal (auch wenn die Bezeichnung *Finites-Differenzenschema-Ideal* dies nahe legt).

Da nun Objekte vorliegen, die mithilfe der Theorie in Kapitel 1 beschrieben und berechnet werden können, werden an dieser Stelle einige Beobachtungen gemacht:

Da $\mathbb{F} \otimes_{\mathbb{F}} C_1^\delta \otimes_{\mathbb{F}} C_2(\underline{T}) = \mathbb{K}$ ein Körper ist, wird die Frage beantwortet, warum das Ideal

$$D \trianglelefteq \mathbb{K} [(\Lambda^\delta \cup M^\delta) \setminus L^\delta]$$

ein Differenzenschema im numerischen Sinne sein soll. Aufgrund von Korollar 1.22 gilt im Falle

$$|(\Lambda^\delta \cup M^\delta) \setminus L^\delta| = 1$$

daß der Ring D_2 ein Hauptidealbereich ist und somit das Ideal D von einem Element erzeugt wird

$$D = \langle d \rangle_{D_2}$$

Dieser Erzeuger d ist bis auf Multiplikation mit Konstanten eindeutig und in einer solchen Situation kann von *dem* Differenzenschema gesprochen werden, wohingegen für

$$|(\Lambda^\delta \cup M^\delta) \setminus L^\delta| \geq 2$$

der Ring D_2 im allgemeinen kein Hauptidealbereich mehr ist und D nicht notwendigerweise von nur einem Element erzeugt werden kann. Dies würde mit den entsprechenden Erzeugern zu einem System von mehreren Gleichungen führen.

Bemerkung 4.19. Die Idee der Definition 4.17 ist: alle Ableitungen (Variablen), die von anderen Ableitungen (Variablen) approximiert werden, sollen eliminiert werden, damit nur die Ableitungen, die nicht mehr selbst approximiert werden, übrig bleiben.

Zwei Beispiele in SINGULAR sollen den Begriff des Differenzenschemas praktisch erläutern:

Beispiel 4.20. *Betrachte die wohlbekannte Advektionsgleichung*

$$u_t + a \cdot u_x = 0$$

Wähle $C_1 = \mathbb{F}(a) = C_1^c$ und approximiere die Ableitungen u_t, u_x durch einfache Differenzenquotienten mit den Schrittgrößen $k = \Delta t, h = \Delta x$:

$$u_t = \frac{u(x, t+k) - u(x, t)}{k} = \frac{T_t \circ u - u}{k}$$

respektive

$$u_x = \frac{u(x+2h, t) - u(x, t)}{2h} = \frac{T_x^2 \circ u - u}{2h}$$

In der Notation wie oben schreibt sich dies als

$$S^\delta = \{u_t + au_x\}$$

und

$$A^\delta = \{k \cdot u_t - (T_t - 1)u, 2h \cdot u_x - (T_x^2 - 1)u\}$$

4.2 Theorie der algebraischen Methode

Es ist klar, daß die Unterscheidung zwischen \cdot und \circ nicht stattfindet und invertierbare Koeffizienten h, k für u_x and u_t angenommen werden. Damit ergibt sich

$$L^\delta = \{u_t, u_x\} \subsetneq \{u_t, u_x, u\} = M^\delta$$

Um ein Differenzenschema wie in 4.17 mittels SINGULAR zu berechnen, muss erst der geeignete Ring $A = \mathbb{Q}(a, h, k)[u_t, u_x, u]$ und die Ordnung $<$ gemäß 4.12 erklärt werden. In SINGULAR wird definiert man zusätzliche Variablen T_t und T_x (als Differenzenoperatoren) zusammen mit einer Blockordnung, welche T_t und T_x den Variablen u_t, u_x, u unterordnet:

```
> ring A=(0,a,h,k),(Ut,Ux,U,Tt,Tx),(lp(3),lp);
```

definiere die Ideale

```
> ideal S=Ut+a*Ux;
> ideal L=Ut-(Tt-1)/k*U,Ux-(Tx^2-1)/(2*h)*U;
> ideal I=S+L; //das Ideal aus der Vereinigung von S und L
```

und berechne das Differenzenschema

$$(\langle S^\delta \rangle_{D_2} + \langle A^\delta \rangle_{D_2}) \cap \mathbb{F} \otimes C_1^\delta \otimes C_2(\underline{T}) [(\Lambda^\delta \cup M^\delta) \setminus L^\delta]$$

das heißt, eliminiere die ersten zwei Variablen u_t und u_x im Ideal:

```
> eliminate(I,intvec(1,2));
_[1]=(2*h)*U*Tt+(a*k)*U*Tx^2+(-a*k-2*h)*U
```

Man kann nach Geschmack auch die Reduktion der Erzeuger von S^δ durch die Erzeuger des Ideals A^δ vornehmen, um modulo Multiplikation mit $2hk$ dasselbe Ergebnis zu erzielen:

```
> reduce(S,std(L));
_[1]=1/(k)*U*Tt+(a)/(2*h)*U*Tx^2+(-a*k-2*h)/(2*h*k)*U
```

Der Test der Gleichheit

```
> reduce(S,L)[1]*2hk==eliminate(I,intvec(1,2))[1];
1
```

Kapitel 4 Algebraische Methoden für Finite Differenzenschemata

bestätigt dies. Man erhält das Differenzenschema

$$(T_t - 1 + \frac{ak}{2h}(T_x^2 - 1)) \circ u$$

oder aufgelöst nach dem höchsten Zeitschritt

$$T_t \circ u = u - \frac{ak}{2h} \cdot (T_x^2 \circ u - u)$$

respektive

$$u_i^{j+1} = u_i^j - \frac{ak}{2h} \cdot (u_{i+2}^j - u_i^j)$$

Den Differenzenstern dieses Verfahrens sieht man in Abbildung 4.2.

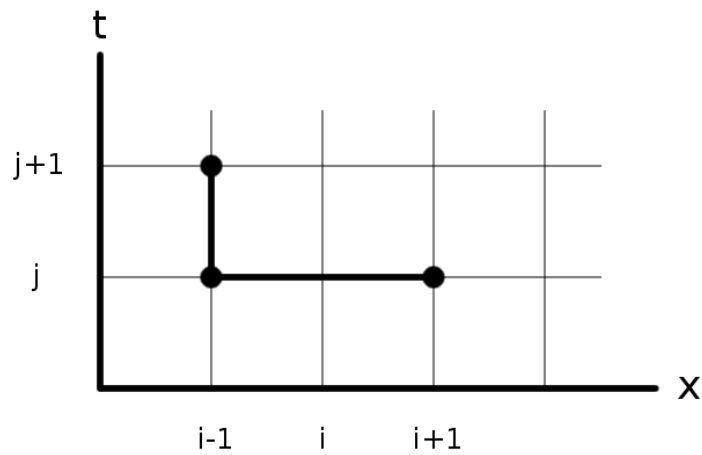


Abbildung 4.2: Differenzenstern von Beispiel 4.20

Im nächsten Beispiel wird das Lax-Wendroff-Verfahren dargestellt:

Beispiel 4.21. Auch hier sei die Advektionsgleichung betrachtet

$$S^\delta = \{u_t + a \cdot u_x\}$$

4.2 Theorie der algebraischen Methode

mit der Vorwärts-Approximation für die Zeitrichtung

$$(u_t)_i^j = \frac{u_i^{j+1} - u_i^j}{k}$$

und einer Gewichtung aus Upwind- und Downwind-Approximation für die räumliche Richtung

$$(u_x)_i^j = \theta \frac{u_i^j - u_{i-1}^j}{h} + (1 - \theta) \frac{u_{i+1}^j - u_i^j}{h}$$

mit konkreten Gewichten

$$\theta = \frac{1 + \frac{ak}{h}}{2} \quad (1 - \theta) = \frac{1 - \frac{ak}{h}}{2}$$

Damit ist

$$A^\delta = \left\{ u_t - \frac{(T_t - 1)}{k} u, u_x - \frac{\left(1 + \frac{ak}{h}\right) (1 - T_x^{-1})}{h} u - \frac{\left(1 - \frac{ak}{h}\right) (T_x - 1)}{h} u \right\}$$

Hier werden Differenzen-Operatoren in positiver und negativer Richtung verwendet. Dementsprechend definiere den Ring in SINGULAR via

```
> ring A=(0,a,h,k),(Ut,Ux,U,Ttp,Ttm,Txp,Txm),(lp(3),lp);
```

und stelle die gegenseitige Wirkung der Differenzenoperatoren durch das Ideal

```
> ideal C=Ttp*Ttm-1,Txp*Txm-1;
```

sicher. Dann definiere S^δ und A^δ :

```
> ideal S=Ut+a*Ux;
```

```
> ideal L=Ut-(Ttp-1)*1/k*U,Ux-1/2*(1+k/h*a)*(1-Txm)/h*U
-1/2*(1-k/h*a)*(Txp-1)/h*U;
```

und eliminiere die ersten zwei Variablen u_t und u_x im Ideal

$$\langle S^\delta \rangle_{D_2} + \langle A^\delta \rangle_{D_2}$$

im Grundring:

```
> ideal I=reduce(S+L,std(C));
```

```
> intvec v=1,2;
```

```
> reduce(eliminate(I,v),std(C));
```

```
_ [1]=(-2*h^2)*U*Ttp+(a^2*k^2-a*h*k)*U*Txp+(a^2*k^2+a*h*k)*U*Txm
+(-2*a^2*k^2+2*h^2)*U
```

Kapitel 4 Algebraische Methoden für Finite Differenzenschemata

Wie im vorigen Beispiel ergibt sich das gleiche Resultat durch Reduktion von S^δ durch A^δ

```
> reduce(reduce(S, std(L)), std(C));
_[1]=1/(k)*U*Ttp+(-a^2*k+a*h)/(2*h^2)*U*Txp
+(-a^2*k-a*h)/(2*h^2)*U*Txm+(a^2*k^2-h^2)/(h^2*k)*U
```

Der Unterschied ist auch hier eine Konstante, nämlich $(-2h^2k)$:

```
> reduce(reduce(S, std(L)), std(C)) [1]*-2*h^2*k
==reduce(eliminate(I, v), std(C)) [1];
1
```

Mit der Schreibweise $c = \frac{ak}{h}$ erhalte das Differenzenschema

$$T_t u_t = T_x u \frac{(c^2 - c)}{2} + u(1 - c^2) + T_x^{-1} u \frac{(c^2 + c)}{2}$$

oder klassisch

$$u_i^{j+1} = u_{i+1}^j \frac{(c^2 - c)}{2} + u_i^j (1 - c^2) + u_{i-1}^j \frac{(c^2 + c)}{2}$$

Der Differenzenstern ist in Abbildung 4.3 zu sehen.

Eine sehr schöne Eigenschaft besitzt die Menge A^δ :

Lemma 4.22. A^δ ist eine Gröbnerbasis bezüglich der Ordnung $<$.

Beweis. Da alle Leitmonome u_{x^β} paarweise verschiedene Variablen im Ring

$$D_2 = \mathbb{F} \otimes_{\mathbb{F}} C_1^\delta \otimes_{\mathbb{F}} C_2(\underline{T}) [\Lambda^\delta \cup M^\delta]$$

sind, sieht man aufgrund des Produkt-Kriteriums in Lemma 4.23, daß A^δ eine Gröbnerbasis ist. \square

Das Produkt-Kriterium ist sehr nützlich und verbessert die Laufzeit des Algorithmus 1.37 zur Berechnung von Gröbnerbasen erheblich, da die Paarmenge ohne aufwendige Normalformberechnungen bezüglich der Paare von Elementen bereinigt werden kann, die in ihren Leitmonomen koprim sind.

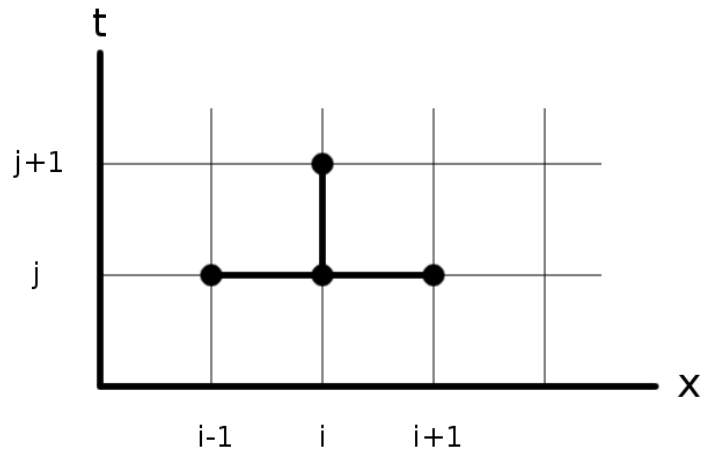


Abbildung 4.3: Differenzenstern des Lax-Wendroff-Schemas im Beispiel 4.21

Lemma 4.23. Produkt-Kriterium

Es seien a und b zwei Polynome mit

$$lm(a) \cdot lm(b) = lcm(lm(a), lm(b))$$

dann

$$NF(\text{spoly}(a, b) \mid \{a, b\}) = 0.$$

Beweis. Dies ist eine simple Konsequenz aus dem Reduktionsprozess und der Konstruktion des *spoly* zweier Polynome (vgl. [GP08]). \square

Eine nützliche Folgerung aus dem Produkt-Kriterium ist:

Korollar 4.24. Für a und b wie in 4.23 gilt, daß $\{a, b\}$ eine Gröbnerbasis des Ideals $\langle a, b \rangle$ ist. \square

Mit den Ergebnissen aus Kapitel 1 kann nun ein Theorem angegeben werden, das zwei äquivalente Möglichkeiten zur Berechnung eines Differenzschemas bereitstellt:

Theorem 4.25. *Es sei $S^\delta = \{h\}$ eine beliebige partielle Differentialgleichung einer unbekannteten Funktion u und es sei A^δ eine Menge von Approximationen, die hinreichend für die Erzeugung eines Verfahrens ist, zusammen mit der entsprechenden Ordnung $<$ auf den Monomen des Ringes*

$$D_2 = \mathbb{F} \otimes_{\mathbb{F}} C_1^\delta \otimes_{\mathbb{F}} C_2(\underline{T}) [\Lambda^\delta \cup M^\delta] = \mathbb{K} [\Lambda^\delta \cup M^\delta]$$

wie in 4.12. Die Ordnung $<$ habe die Eliminationseigenschaft (Definition 1.41) für die Menge L^δ . Es sei wie oben L^δ die Menge der auftretenden Ableitungen auf der linken Seite in A^δ , M^δ die Menge aller Ableitungen in A^δ und schließlich Λ^δ die Menge aller Ableitungen in h . Wenn diese Mengen die Eigenschaft

$$|(\Lambda^\delta \cup M^\delta) \setminus L^\delta| = 1$$

besitzen, dann existieren zwei äquivalente Methoden zur Berechnung eines Differenzenverfahrens:

1. $h \xrightarrow{\text{red.}} NF(h | A^\delta) = a_1 \in D_2$ für eine reduzierte Normalform NF
2. $(\langle S^\delta \rangle_{D_2} + \langle A^\delta \rangle_{D_2}) \cap \mathbb{K} [(\Lambda^\delta \cup M^\delta) \setminus L^\delta] = \langle a_2 \rangle_{D_2}$

Äquivalent bedeutet die Identität von a_1 und a_2 modulo Multiplikation mit einer Konstanten:

$$a_1 = c \cdot a_2$$

Beweis. Es sei ohne Beschränkung der Allgemeinheit $\Lambda^\delta \cup M^\delta = \{u_m, \dots, u_1, u\}$ und $(\Lambda^\delta \cup M^\delta) \setminus L^\delta = \{u\}$ in der Reihenfolge $u_m > \dots > u_1 > u$ bezüglich der Ordnung $<$. Als eine reduzierte Normalform sei der reduzierte Buchberger-Algorithmus (Algorithmus 1.31) mit Schwanzreduktion gewählt. Nun ist in der Iterationsschleife der Normalform die Berechnung des *spoly* wesentlich:

$$\text{spoly}(h, m) = h - \frac{lc(h)}{lc(m)} \cdot \underline{u}^{\alpha-\beta} m$$

für ein Element $m \in A^\delta$ mit $lm(m) = \underline{u}^\alpha | \underline{u}^\beta = lm(h)$. Offensichtlich gilt dann

$$lm(\text{spoly}(h, m)) < lm(h)$$

und im Laufe des Algorithmus erhalte eine Folge von Leitmonomen, die stationär wird:

$$lm(h) > lm(h_1) > \dots > lm(h_s)$$

Behauptung: $h_s = u$. Dies folgt einfach, da die Approximationen in A^δ entsprechende Ersetzungen jeder auftretenden Ableitung bereitstellen und diese Ringvariablen sind. Die Annahme im Theorem stellt sicher, daß die Behauptung gilt und somit

$$h \in \mathbb{K}(\underline{T})[u]$$

Darüber hinaus garantiert die Darstellung in Standardrepräsentation, daß

$$h \in (\langle S^\delta \rangle_{D_2} + \langle A^\delta \rangle_{D_2})$$

An dieser Stelle ist die *Reduziertheit* der Normalform essentiell, da diese Eigenschaft die Eindeutigkeit gewährleistet (siehe 1.35) und die Tatsache, daß A^δ eine Gröbnerbasis ist (wegen 4.22), in diesem Zusammenhang ausgenutzt wird. Nach der Reduktion ist die Menge $A^\delta \cup \{h\}$ eine Gröbnerbasis von $\langle A^\delta \cup \{h\} \rangle_{D_2}$ wegen des Buchberger-Kriteriums aus 1.36 und mit gleicher Begründung wird klar, daß

$$(A^\delta \cup \{h\}) \cap \mathbb{K}(\underline{T})[u] = \{h\}$$

eine Gröbnerbasis für $(\langle S^\delta \rangle_{D_2} + \langle A^\delta \rangle_{D_2}) \cap \mathbb{K}(\underline{T})[u]$ ist.

Da der Ring D_2 aufgrund der Annahmen im Theorem ein Hauptidealbereich ist, besitzt jedes Ideal einen bis auf Multiplikation mit einer Konstanten, die ungleich Null ist, eindeutigen Erzeuger und somit ist der Beweis vollbracht. \square

Die beiden Beispiele 4.20 und 4.21 haben die Aussage dieses Theorems vorweggenommen und können als Bestätigung gelten.

Bemerkung 4.26. Es wurde im Streben nach größtmöglicher Allgemeinheit bisher keine weitere Annahme und Voraussetzung für die Definition eines Differenzschemas getroffen mit Ausnahme der Forderung nach Existenz einer Ordnung zu einer Menge von Approximationen und der Forderung nach konstanten Koeffizienten. Auch ist es ein fundamentaler Unterschied, daß keine Forderungen nach Linearität der auftretenden Gleichungen nötig sind, wie dies im Janet-Basen-Ansatz von V. Gerdt ([GBM06]) oder im Modul-Ansatz von [LM10] der Fall ist.

In [LM10] werden für den speziellen linearen Fall einige sehr angenehme Eigenschaften der Erzeugung von Differenzschemata gezeigt, diese werden an späterer Stelle (Theorem 4.32) vorgestellt.

Die Formulierung von Bedingungen für die Existenz von Differenzenschemata kann auch für den Fall einer Gleichung in mehreren unbekannt Funktionen u_1, \dots, u_s etc. erfolgen. Die Bedingung in 4.12 muß separat für jede einzelne Funktion u_i mit den Approximationsmengen der Ableitungen in u_i gelten. Bezeichnet man in Analogie zur bisherigen Notation die auftretenden Variablenmengen mit M_{u_i} , dann muß für einen entsprechenden Grundkörper K , der alle Koeffizienten und Differenzenoperatoren enthält, in einem Ring

$$A = K [M_{u_1}, \dots, M_{u_s}]$$

mit einer Blockordnung gearbeitet werden.

Bemerkung 4.27. Es sei $|S^\delta| > 1$ eine Menge von *mehreren* partiellen Differentialgleichungen in einer unbekannt Funktion. Dann kann die Möglichkeit der Erzeugung eines Differenzenschemas auf den Fall *einer* Gleichung zurückgeführt werden, indem man die einzelnen Gleichungen separat behandelt:

$$S^\delta = \{s_1^\delta, \dots, s_m^\delta\} = \{s_1^\delta\} \cup \dots \cup \{s_m^\delta\} = S_1^\delta \cup \dots \cup S_m^\delta$$

Dann erhalte die verschiedenen Differenzenverfahren für jede Gleichung S_j^δ . Es ist klar, daß für die S_j^δ bestimmte Approximationen A_j^δ zu verwenden sind, die gegenseitig kompatibel sein müssen (vergleiche [GBM06]).

Im nächsten Theorem werden Bedingungen und Eigenschaften für die Generierung von Differenzenschemata im Falle eines Systems von mehreren unbekannt Funktionen vorgestellt. Es wird vorausgesetzt, daß eine Menge mit Approximationen für alle Gleichungen im System gültig ist:

Theorem 4.28. *Es sei $S^\delta = \{s_1^\delta, \dots, s_N^\delta\}$ eine Menge von N partiellen Differentialgleichungen in M unbekannt Funktionen $\underline{U} = \{u_1, \dots, u_M\}$ zusammen mit den gemeinsamen Approximationen $A^\delta = \{a_1, \dots, a_k\}$. Es sei die Menge A^δ hinreichend zur Erzeugung eines Verfahrens mit einer Eliminationsordnung $<$ für die Menge L^δ aller Ableitungen auf der linken Seite in den Approximationen A^δ . Wie oben sei Λ^δ die Menge aller Ableitungen in S^δ und es bezeichne M^δ alle Ableitungen in A^δ , sodaß*

$$(\Lambda^\delta \cup M^\delta) \setminus L^\delta = \underline{U}$$

Es sei wie bisher D_2 der Ring

$$D_2 = \mathbb{F} \otimes_{\mathbb{F}} C_1^\delta \otimes_{\mathbb{F}} C_2(\underline{T}) [\Lambda^\delta \cup M^\delta] = \mathbb{K} [\Lambda^\delta \cup M^\delta]$$

Die beiden Ideale I_1 und I_2 seien definiert als

1.

$$I_1 = (\langle G \rangle_{D_2}) \cap \mathbb{K} [\underline{U}]$$

und G sei eine Gröbnerbasis von $S^\delta \cup A^\delta$ bezüglich der Eliminationsordnung $<$

2.

$$I_2 = (\langle h_1, \dots, h_N \rangle_{D_2} + \langle A^\delta \rangle_{D_2}) \cap \mathbb{K} [\underline{U}]$$

mit $h_j = NF(s_j^\delta \mid A^\delta)$ für alle j

Behauptung:

$$I_1 = I_2$$

Beweis. Der Beweis ist offensichtlich und folgt elementarer Idealtheorie. \square

Bemerkung 4.29.

1. Beachte, daß in $\mathbb{K}(\underline{T})[\underline{U}]$ keine approximierten Ableitungen mehr auftreten, da sie durch die Approximationen ersetzt werden mittels der hinreichenden Eigenschaft zur Erzeugung eines Verfahrens und wegen $(\Lambda^\delta \cup M^\delta) \setminus L^\delta = \underline{U}$. Damit ist offensichtlich, daß $h_j \in \mathbb{K}(\underline{T})[\underline{U}]$ und somit $h_j \in I_2 \forall j$. Auch ist $\{h_j\} \cup A^\delta$ eine Gröbnerbasis, aber die Menge $\{h_1, \dots, h_N\} \cup A^\delta$ ist im Allgemeinen nicht notwendigerweise eine Gröbnerbasis von $\langle h_1, \dots, h_N \rangle_{D_2} + \langle A^\delta \rangle_{D_2}$.

2. Die beiden Methoden aus dem letzten Theorem unterscheiden sich vom prinzipiellen und praktischen Standpunkt: das erste Ideal I_1 wird durch die Berechnung einer Gröbnerbasis der gesamten Vereinigung $S^\delta \cup A^\delta$ gewonnen, in welcher alle Gleichungen enthalten sind. Dagegen ergibt sich das Ideal I_2 durch die gegenseitig unabhängige Berechnung der einzelnen Gröbnerbasen $\{h_j\} \cup A^\delta$. Die Komplexität des ersten Ansatzes kann signifikant höher sein als die des zweiten, jedoch könnte unter Umständen ein Vorteil im Aufdecken von Zusammenhängen zwischen den Objekten mittels einer gemeinsamen Berechnung einer Gröbnerbasis bestehen, was in einer isolierten Berechnung nicht

möglich scheint.

3. In [GR10] untersuchen Gerdt und Robertz Konsistenz im Falle variabler Koeffizienten. Es muß beachtet werden, daß in diesem Abschnitt alle Aussagen Gültigkeit besitzen für den Fall konstanter Koeffizienten (insbesondere die Theoreme 4.28 und 4.25). Ein weiterer Unterschied besteht darin, daß in [GR10] *differentielle* Gröbnerbasen berechnet werden, was in dieser Arbeit nicht der Fall ist.

4.3 Lineare Partielle Differentialgleichungen

Für den linearen Fall ist nicht nur die Konvergenztheorie der Schemata in der Numerik mit dem Von-Neumann-Ansatz besonders angenehm, es ist auch algebraisch sehr einfach, mithilfe des Modul-Ansatzes aus Theorem 4.32 ein Differenzenschema zu erzeugen.

Modul-Ansatz

Es sei mit

$$S = \left\{ \sum_{\beta \in B} c_{\beta} u_{x^{\beta}} \right\}$$

eine Menge einer linearen partiellen Differentialgleichung bezeichnet, wobei $B \subset \mathbb{N}^{m+1}$ eine endliche Indexmenge ist und jedes $c_{\beta} \in K$ eine einfache Konstante.

Für ein Gitter $\Omega := (\mathbb{Z}\Delta x_1 \times \dots \times \mathbb{Z}\Delta x_n \times \mathbb{Z}\Delta t) \cap \mathbb{R}^{n+1}$ wie vorher schreibe die Gleichung diskretisiert als

$$S^{\delta} = \left\{ \sum_{\beta \in B} c_{\beta} (u_{x^{\beta}})^{\delta} \right\}$$

Eine Approximation für alle Ableitungen $(u_{x^{\beta}})^{\delta}$ sei durch Ableitungen niedrigeren Grades mittels einer endlichen Summe, die linear in den $(u_{x^{\omega}})^{\nu}$ ist,

$$(u_{x^{\beta}})^{\delta} - \sum_{\omega \in \Omega_{\beta}^1, \nu \in \Omega_{\beta}^2} c_{\omega, \nu} (u_{x^{\omega}})^{\nu}$$

für $|\omega| < |\beta|$, also Ableitungen kleineren Grades in der Summe, und konstante Koeffizienten $c_{\omega, \nu} \in \mathbb{F}$ in einem Körper \mathbb{F} wie im letzten Abschnitt gegeben.

4.3 Lineare Partielle Differentialgleichungen

Mit Differenzenoperatoren schreiben sich die Approximationen als

$$(u_{x^\beta})^\delta - \sum_{\omega \in \Omega_\beta^1, \nu \in \Omega_\beta^2} c_{\omega, \nu} T^{\lambda_\nu} (u_{x^\omega})^\delta$$

in einem generischen Punkt $\delta \in \mathbb{Z}^{n+1}$. Die Menge aller Approximationen sei wie vorher mit A^δ , die auftretenden Ableitungen in S^δ mit Λ^δ , die Menge aller Ableitungen in A^δ mit M^δ und alle Ableitungen auf der linken Seite in A^δ mit L^δ bezeichnet. Es wird hier die Annahme

$$\Lambda^\delta = M^\delta$$

getroffen, somit soll jede auftretende Ableitung in S^δ eine Approximation besitzen. Es habe die Menge der Approximationen A^δ im Ring

$$D_2 = \mathbb{F}(\underline{T})[\Lambda^\delta]$$

die hinreichende Eigenschaft zur Erzeugung eines Verfahrens und es sei eine Ordnung $<$ mit selbiger Eigenschaft. Die Menge Λ^δ sei in der Reihenfolge bezüglich $<$ geschrieben als

$$\Lambda^\delta = \{u_r, \dots, u_1\}$$

mit $u_k > u_l$ wenn $k < l$. Desweiteren sei

$$\Lambda^\delta \setminus L^\delta = \{u_1\}$$

die Menge der Ableitungen, die nicht durch andere Ableitungen approximiert werden, bestehend aus exakt einem Element.

Bemerkung 4.30. Es kann ohne Beschränkung der Allgemeinheit angenommen werden, daß in den Gleichungen

$$(u_{x^\beta})^\delta - \sum_{\omega \in \Omega_\beta^1, \nu \in \Omega_\beta^2} c_{\omega, \nu} T^{\lambda_\nu} (u_{x^\omega})^\delta$$

die Exponenten der Differenzenoperatoren *nicht-negative Einträge* besitzen: man multipliziere im Falle negativer Komponenten in der Gleichung auf beiden Seiten mit einem bezüglich der Exponenten hinreichend großen Differenzenoperator und betrachte die Gleichungen an einem neuen generischen Punkt $\check{\delta}$. Dieser neue generische Punkt muss für alle Gleichungen in S^δ und A^δ derselbe sein.

Betrachte den endlich erzeugten $\mathbb{F}[\underline{T}]$ -Modul $\mathbb{F}[\underline{T}]^r := \bigoplus_{i=1}^r \mathbb{F}[\underline{T}] e_i$ und definiere die Abbildung

$$\begin{aligned} \pi : \mathbb{F}[\underline{T}][\Lambda^\delta] &\longrightarrow \mathbb{F}[\underline{T}]^r \\ u_i &\longmapsto e_i \end{aligned}$$

Dann führt nach der Annahme der nicht-negativen Exponenten der Differenzenoperatoren in Bemerkung 4.30 die kanonische Inklusion der Mengen $S^\delta, A^\delta \subset \mathbb{F}[\underline{T}][\Lambda^\delta]$

$$(S^\delta \cup A^\delta) \subset \mathbb{F}[\underline{T}][\Lambda^\delta]$$

zu der Abbildung

$$\tilde{\pi} = \pi|_{S^\delta \cup A^\delta} : S^\delta \cup A^\delta \longrightarrow \mathbb{F}[\underline{T}]^r$$

sodaß die Bilder der Elemente in $S^\delta \cup A^\delta$ Vektoren in $\mathbb{F}[\underline{T}]^r$ sind.

Bemerkung 4.31. Die Abbildung π ist definiert, da es sich bei den Gleichungen A^δ und S^δ um lineare Gleichungen in den auftretenden Ableitungen handelt.

Ein Ergebnis aus [LM10] ist die Äquivalenz der Berechnung einer Gröbnerbasis im Modul $\mathbb{F}[\underline{T}]^r$ mit der Modulordnung $\leq = (c, <_T)$ via

$$x^\alpha \cdot e_i > x^\beta \cdot e_j \quad :\Leftrightarrow \quad i < j \vee (i = j \wedge x^\alpha >_T x^\beta)$$

aus Beispiel 1.53, die sich bezüglich einer Ordnung $<_T$ in $\mathbb{F}[\underline{T}]$ ergibt, und den beiden Methoden in Theorem 4.25:

Theorem 4.32. Modul-Ansatz

Für eine lineare partielle Differentialgleichung $S^\delta = \{h\}$ mit konstanten Koeffizienten, eine Menge von Approximationen A^δ , die auftretenden Ableitungen in den jeweiligen Mengen $M^\delta, \Lambda^\delta, L^\delta$ wie oben und der hinreichenden Eigenschaft von A^δ für die Erzeugung eines Schemas mit einer Eliminationsordnung für L^δ und

$$\Lambda^\delta \setminus L^\delta = \{u_1\}$$

sei \tilde{M} eine Modul-Gröbnerbasis des Untermoduls

$$M := \langle \tilde{\pi}(S^\delta \cup A^\delta) \rangle_{\mathbb{F}[\underline{T}]} \subset \mathbb{F}[\underline{T}]^r$$

zur Ordnung $(c, <_T)$. Mit

4.3 Lineare Partielle Differentialgleichungen

1. $h \xrightarrow{\text{red.}} NF(h \mid A^\delta) = a_1 \in \mathbb{F}(\underline{T})[u_1]$, NF eine reduzierte Normalform
2. $(\langle S^\delta \rangle_{D_2} + \langle A^\delta \rangle_{D_2}) \cap \mathbb{F}(\underline{T})[u_1] = \langle a_2 \rangle_{\mathbb{F}(\underline{T})[u_1]}$
3. $\tilde{M} \cap \mathbb{F}[\underline{T}] \cdot e_r = I \cdot e_r$ mit $I = \langle a_3 \rangle_{\mathbb{F}[\underline{T}]} \trianglelefteq \mathbb{F}[\underline{T}]$

gilt die Behauptung:

es sind die Ergebnisse dieser drei Methoden identisch in dem Sinne, daß wie in Bemerkung 4.30 die Polynome a_1 und a_2 als Elemente in $\mathbb{F}[\underline{T}][u_1]$ aufgefasst werden können und gilt

$$\pi(a_1) = c_1 \pi(a_2) = c_2 a_3$$

für $c_1, c_2 \in \mathbb{F}$. □

Ein Beispiel zeigt die Möglichkeiten des Theorems 4.32:

Beispiel 4.33. Es sei eine lineare partielle Differentialgleichung

$$S^\delta = \{(u_t)^\delta + a(u_x)^\delta\}$$

und Approximationen

$$A^\delta = \{(2hT_x)(u_x)^\delta + (-T_x^2 + 1)(u)^\delta, (kT_x)(u_t)^\delta + (-T_t T_x + \frac{1}{2}(T_x^2 + 1))(u)^\delta\}$$

dann ist

$$\Lambda^\delta = M^\delta = \{u_x, u_t, u\} \text{ und } \Lambda^\delta \setminus L^\delta = \{u\}$$

Die Ordnung $<$ mit $u_x > u_t > u$ ist hinreichend zur Erzeugung eines Schemas, daher definiere die Abbildung

$$\pi : \mathbb{Q}[T_x, T_t][u_x, u_t, u] \rightarrow \mathbb{Q}[T_x, T_t]^3$$

$$\text{mit } \pi(u_x) = e_1, \pi(u_t) = e_2 \text{ und } \pi(u) = e_3$$

Dann kann der Modul M_1 , der von $\pi(S^\delta \cup A^\delta)$ aufgespannt wird, als eine Matrix verstanden werden, deren Zeilen M_1 erzeugen:

$$M_1 = \begin{pmatrix} aT_x T_t & T_x T_t & 0 \\ 2hT_x & 0 & -T_x^2 - 1 \\ 0 & kT_x & \frac{1}{2}T_x^2 - T_x T_t + \frac{1}{2} \end{pmatrix}$$

Eine Gröbnerbasis M_2 dieses Moduls berechnet sich in SINGULAR mittels:

Kapitel 4 Algebraische Methoden für Finite Differenzenschemata

```

> ring A=(0,a,h,k),(Ux,Ut,U,Tx,Tt),(c,lp);
> vector v1=[Tx*Tt*a,Tx*Tt*1,0];
> vector v2=[2*h*Tx,0,-(Tx^2+1)];
> vector v3=[0,k*Tx,-Tt*Tx+1/2*(Tx^2+1)];
> module M1=v1,v2,v3;
> def M2=std(M1);M2;
M2[1]=[0,0,(a*k-h)*Tx^2*Tt+(2*h)*Tx*Tt^2+(a*k-h)*Tt]
M2[2]=[0,(2*k)*Tx,Tx^2-2*Tx*Tt+1]
M2[3]=[(2*h)*Tx,0,-Tx^2-1]

```

und damit in Matrixschreibweise für die Gröbnerbasis:

$$M_2 = \begin{pmatrix} 0 & 0 & (ak-h)T_x^2T_t + 2hT_xT_t^2 + (ak-h)T_t \\ 0 & 2kT_x & T_x^2 - 2T_xT_t + 1 \\ 2hT_x & 0 & -T_x^2 - 1 \end{pmatrix}$$

Schneidet man diesen Modul M_2 mit dem Erzeugnis des dritten Einheitsvektors, ergibt sich für $M_2 \cap Ae_r$

```

> intersect(M2,module(gen(3)));
_[1]=[0,0,(-a*k+h)*Tx^2*Tt+(-2*h)*Tx*Tt^2+(-a*k+h)*Tt]

```

und damit das Polynom

$$(ak-h)T_x^2T_t + 2hT_xT_t^2 + (ak-h)T_t$$

Der Vergleich mit den Methoden 1. und 2. in Theorem 4.32 liefert in SINGULAR

```

> ideal S=Tx*Tt*(Ut+a*Ux);
> ideal L=2*h*Tx*Ux-Tx^2*U-U,k*Tx*Ut-Tt*Tx*U+1/2*(Tx^2+1)*U;
> ideal I=S+L;
> eliminate(I,intvec(1,2));
_[1]=(a*k-h)*U*Tx^2*Tt+(2*h)*U*Tx*Tt^2+(a*k-h)*U*Tt
> reduce(S,std(L));
_[1]=(a*k-h)/(2*h*k)*U*Tx^2*Tt+1/(k)*U*Tx*Tt^2+(a*k-h)/(2*h*k)*U*Tt
> intersect(M2,module(gen(3)))[1][3]*U==reduce(S,std(L))[1]*-2*h*k;
1

```

und somit ist der Unterschied in der Berechnung nur die Multiplikation mit der Konstante $-2hk$.

Bemerkung 4.34.

1. Wie in Beispiel 4.33 zu sehen ist, besteht die Errechnung einer Modul-Gröbnerbasis M_2 für den Modul M_1 aus dem Gauß-Algorithmus, der die Zeilen der Matrix M_1 in der Komponentenreihenfolge der Modulordnung einsetzt und eine untere Dreiecksmatrix M_2 erzeugt.
2. Die Schreibweise des Moduls M_1 in Matrixform deutet die durch die Gleichungen A^δ und S^δ induzierten Relationen anschaulich durch die Multiplikation mit dem Vektor $(u_x, u_t, u)^T$ an:

$$\begin{pmatrix} aT_xT_t & T_xT_t & 0 \\ 2hT_x & 0 & -T_x^2 - 1 \\ 0 & kT_x & \frac{1}{2}T_x^2 - T_xT_t + \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} u_x \\ u_t \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Eine exemplarische Auflistung von Approximationen ist mithilfe des kanonischen Skalarprodukts gegeben durch:

Forwärts-Appr. $(\Delta x, 1 - T_x)(u_x, u)^T = 0$

Rückwärts-Appr. $(\Delta x \cdot T_x, 1 - T_x)(u_x, u)^T = 0$

Zentrale Appr. (1.Ord.) $(2\Delta x \cdot T_x, 1 - T_x^2)(u_x, u)^T = 0$

Zentrale Appr. (2.Ord.) $(-\Delta x^2 \cdot T_x, (1 - T_x)^2)(u_{xx}, u)^T = 0$

Trapez-Appr. $(\frac{1}{2}\Delta x \cdot (T_x + 1), 1 - T_x)(u_x, u)^T = 0$

Mittelpunkt-Appr. $(2\Delta x \cdot T_x, 1 - T_x^2)(u_x, u)^T = 0$

Pyramiden-Appr. $(\frac{1}{3}\Delta x \cdot (T_x^2 + 4T_x + 1), 1 - T_x^2)(u_x, u)^T = 0$

Fouriertransformation

Die Methoden der Computeralgebra können auch für die Errechnung des Symbols $\varrho(\xi)$ eines expliziten Zweischrittverfahrens eingesetzt werden:

Für ein Verfahren der Form

$$u_k^{j+1} = \sum_{s=-m}^m c_s u_{k+s}^j$$

ist das Symbol definiert als

$$\varrho(\xi) := \sum_{s=-m}^m c_s e^{si\xi}$$

Wenn nun ein Verfahren H nach dem Modul-Ansatz in Theorem 4.32 berechnet wird und sich ein Polynom f in $\mathbb{F}[\underline{T}]$ ergibt, sodaß H nach einer Indexverschiebung der Exponenten von T_t in der Form geschrieben werden kann

$$T_t^1 T_x^\gamma = \sum_{\theta \in \Theta} c_\theta T_t^0 T_x^\theta = \sum_{\theta \in \Theta} c_\theta T_x^\theta$$

für eine endliche Indexmenge $\Theta \subset \mathbb{Z}^n$ mit $T_x^\theta := \prod_{l=1, \dots, n} T_{x_l}^{\theta_l}$, dann definiere einen \mathbb{F} -linearen Ringhomomorphismus Φ :

$$\Phi : \mathbb{F}[T_{x_1}, \dots, T_{x_n}, T] \longrightarrow \mathbb{F}[\sin_{x_1}, \cos_{x_1}, \dots, \sin_{x_n}, \cos_{x_n}, i]/W$$

$$T_{x_l} \longmapsto \cos_{x_l} + i \sin_{x_l} \quad \forall l$$

mit einem Ideal $W \triangleleft \mathbb{F}[\sin_{x_1}, \cos_{x_1}, \dots, \sin_{x_n}, \cos_{x_n}, i]$, welches die Relationen des *Trigonometrischen Pythagoras* und der symbolischen imaginären Einheit i festlegt und diese im Quotientenring rausteilt:

$$W := \langle \sin_{x_1}^2 + \cos_{x_1}^2 - 1, \dots, \sin_{x_n}^2 + \cos_{x_n}^2 - 1, i^2 + 1 \rangle_{\mathbb{F}}$$

Somit erhalte polynomial unter dieser Abbildung das Symbol $\varrho(\xi)$, welches nun aber nicht mehr von ξ sondern von den Ringvariablen \sin_{x_j}, \cos_{x_j} und i abhängt.

Stabilitätsanalyse mit CAD

Soweit wurde deutlich, daß die Computeralgebra Methoden an die Hand gibt, welche unter Einhaltung gewisser Kriterien die Erzeugung eines Differenzverfahrens aus partiellen Differentialgleichungen samt geeigneter Approximationen erlauben. Die Ergebnisse aus diesen automatischen Ansätzen sind denen entsprechend, die im klassischen Fall manuell errechnet werden. In der besonderen Situation einer linearen partiellen Differentialgleichung mit der Möglichkeit der Von-Neumann-Analyse bezüglich der Stabilität eines Verfahrens kann Quantoren-Elimination mithilfe zylindrischer algebraischer Zerlegung (CAD) (vergleiche Kapitel 3) zur Abschätzung des Symbols, welches für

4.3 Lineare Partielle Differentialgleichungen

die Stabilität betragsmäßig kleiner oder gleich 1 sein muss, verwendet werden, um dazu äquivalente Bedingungen zu erhalten.

In diesem Abschnitt wird die Abschätzung der Stabilitätsbedingungen mithilfe des CAD-Systems QEPCAD ([Bro10]) vorgestellt. Dies stellt zusammen mit der automatischen Erzeugung eine komplette Automatisierung der Schema-Untersuchung dar.

Nach dem Theorem 2.16 gilt die Bedingung

$$\sup_{\xi} |\varrho(\xi)| \leq 1$$

als hinreichend für die Stabilität eines linearen Verfahrens nach der $\|\cdot\|_2$ -Norm. Durch die Schreibweise des Symbols in Termen der Sinus- und Kosinusfunktionen folgt, daß der komplexe Betrag des Symbols im Quadrat, $|\varrho(\xi)|^2$, ebenfalls eine Funktion dieser trigonometrischen Funktionen ist, die insbesondere ihr Maximum annimmt für Werte der trigonometrischen Funktionen im Bereich $[-1, 1]$. Dies garantiert, daß der CAD-Ansatz angewendet werden kann, wenn die durch das Polynom $|\varrho(\xi)|^2 - 1$ im \mathbb{R}^d (für ein geeignetes d) definierten Bedingungen untersucht werden.

Betrachte zur Erklärung dieses Vorgehens das Lax-Friedrich-Verfahren, welches in SINGULAR mithilfe der Bibliothek *findiff.lib* ([Din10]) erzeugt und durch Benutzung von QEPCAD innerhalb einer SINGULAR-Sitzung auf Stabilität untersucht wird (für Details siehe Kapitel 5).

```
> list D="Ux","Ut","U";
> list P="a";
> list V="t","x";
> setinitials(V,D,P);
> def s1=scheme(u(Ut)+a*u(Ux),laxfrX(Ux,U,x),laxfrT(Ut,U,x));
> s1;
(-a*dt+dx)/(2*dx)*x^2+(a*dt+dx)/(2*dx)
> def s2=errormap(s1);
> s2;
(-a*dt+dx)/(dx)*i*cx*sx+(a*dt-dx)/(dx)*sx^2+1
> def s3=ComplexValue(s2);s3;
(a^2*dt^2-dx^2)/(dx^2)*sx^2+1
```

Der letzte Ausdruck wird nach Subtraktion mit 1 und der anschließenden Eliminierung der Nenner an QEPCAD geschickt, welches im Koordinatenring

Kapitel 4 Algebraische Methoden für Finite Differenzenschemata

$\mathbb{Q}[\cos_x, \sin_x, a, \Delta_t, \Delta_x, T_t]$ ein äquivalente Bedingung für

$$\forall T_t : a^2 \Delta_t^2 \sin_x^2 - \Delta_x^2 \sin_x^2 < 0$$

errechnet:

```
> qepcad(s3);  
sx /= 0 /\ [ [ dx + a dt > 0 /\ dx - a dt > 0 ]  
            \/ [ dx + a dt < 0 /\ dx - a dt < 0 ] ]
```

Die Bedingung $\sin_x \neq 0$ rührt daher, daß QEPCAD die Bedingung < 0 und nicht etwa ≤ 0 abschätzt, und kann ignoriert werden. Die anderen Bedingungen bedeuten im Supremum nichts anderes als die CFL-Bedingung aus Definition 2.18 und Abbildung 2.3:

$$\left| a \frac{\Delta t}{\Delta x} \right| \leq 1.$$

Kapitel 5

SINGULAR und QEPCAD

Im folgenden werden die beiden Systeme SINGULAR und QEPCAD in aller Kürze vorgestellt. Ziel ist, daß auch mit diesen Systemen nicht vertraute Anwender die Beispiele in dieser Arbeit eigenhändig nachvollziehen können und auch in der Lage sind, eigene Rechnungen anzustellen. Im Anschluß werden die wesentlichsten Routinen der SINGULAR-Bibliothek `findiff.lib` in kurzer Erklärung präsentiert und das Zusammenspiel zwischen SINGULAR und QEPCAD erläutert.

Es wird aufgrund der Programmvoraussetzungen von einem Linux-basierten Betriebssystem ausgegangen. Für nähere Informationen bezüglich der System- und Programmvoraussetzungen beider Systeme wird in den jeweiligen Abschnitten dieses Kapitels auf die Internetadressen der Entwickler hingewiesen.

5.0.1 SINGULAR

Für Details hinsichtlich Installation und neuester Version wird an dieser Stelle auf die Adresse www.singular.uni-kl.de verwiesen.

SINGULAR ist ein auf polynomiale Berechnungen in kommutativen und nicht-kommutativen Ringen ausgerichtetes Computeralgebrasystem mit Anwendungen in Algebraischer Geometrie und Singularitätentheorie.

Vor weitergehenden Berechnungen ist in SINGULAR zunächst die Definition des zugrundeliegenden Rings (beziehungsweise Algebra) wichtig. Dies geschieht beispielhaft durch folgende Eingabe

```
> ring R=(0,a,b),(x,y),(lp);
```

wobei bezüglich der Auswahl einer Charakteristik des Rings (im Beispiel 0), der Parameter (a,b) und der Monomordnung (lp) für die Variablen (x,y) viele

Kapitel 5 SINGULAR und QEPCAD

Möglichkeiten bestehen. Einen konkreten Überblick über die Definition und die Eigenschaften einer Variable des Datentyps `ring` gibt der Befehl

```
> help ring;
```

oder allgemeiner

```
> help;
```

für die allgemeine browsergestützte SINGULAR-Hilfe mit Suchfunktion für alle Themen. Um die Prozeduren der Bibliothek `findiff.lib` (analog für jede andere Bibliothek) zugänglich zu machen, muß diese von SINGULAR geladen werden:

```
> LIB "findiff.lib";
```

Dabei ist sicherzustellen, daß sich die Bibliothek in einem Verzeichnis befindet, das in der Umgebungsvariable `SearchPath` enthalten ist. Einen Überblick über `SearchPath` liefert der Befehl

```
> system("--version");
```

5.0.2 QEPCAD

Der Bezug des freien System QEPCAD ist über die Internet-Adresse <http://www.usna.edu/Users/cs/qepcad/B/QEPCAD.html> möglich. Hier finden sich auch Hinweise bezüglich der Installation.

QEPCAD ist ausgerichtet auf die Quantoreneliminierung mithilfe eines speziellen Algorithmus zur CAD-Berechnung. QEPCAD ist ein Kommandozeilenprogramm, welches in fester Reihenfolge Parameter abfragt und nach erfolgter Rechnung neu gestartet werden muß.

Da mit `findiff.lib` die Kommunikation zwischen SINGULAR und QEPCAD automatisiert ist, wird sich hier mit der Verführung eines einfachen Beispiels begnügt. Die Berechnung der p, q -Formel sieht in QEPCAD wie folgt aus, nachdem QEPCAD in Bash aufgerufen wurde:


```
=====
Quantifier Elimination
in
Elementary Algebra and Geometry
by
Partial Cylindrical Algebraic Decomposition
```

Version B 1.53, 16 Jul 2009

by
Hoon Hong
(hhong@math.ncsu.edu)

With contributions by: Christopher W. Brown, George E.
Collins, Mark J. Encarnacion, Jeremy R. Johnson
Werner Krandick, Richard Liska, Scott McCallum,
Nicolas Robidoux, and Stanly Steinberg

```
=====
Enter an informal description between '[' and ']':
```

[pq-Formel]

```
Enter a variable list:
```

(p,q,x)

```
Enter the number of free variables:
```

2

```
Enter a prenex formula:
```

(E x)[x^2 + p x + q =0].

```
=====
Before Normalization >
```

go

```
Before Projection (x) >
```

go

```
Before Choice >
```

go

```
Before Solution >
```

go

```
An equivalent quantifier-free formula:
```

4 q - p^2 <= 0

```
===== The End =====
```

Kapitel 5 SINGULAR und QEPCAD

In den Schritten `Before Normalization >`, `Before Projection (x) >`, `Before Choice >` und `Before Solution >` können erfahrene Anwender einige optionale Parameter eingeben. Die Voreinstellung `go` wählt den standardisierten Weg. Für Details schaue man in der offiziellen QEPCAD-Dokumentation nach, welche über die Internet-Adresse <http://www.usna.edu/Users/cs/qepcad/B/user/UsingQEPCAD.html> einzusehen ist.

5.0.3 findiff.lib

Die Bibliothek `findiff.lib` ist zu beziehen via

<http://www.mathematik.unikl.de/ftp/pub/Math/Singular/SOURCES/3-1-1/Singular-3-1-1-4.tar.gz>

Nach Laden der Bibliothek stehen die Tools zur Generierung von finiten Differenzenschemata für *lineare* Differentialgleichungen mit einer Fouriertransformation als Vorbereitung einer Stabilitätsanalyse gemäß der Von-Neumann-Methode im Abschnitt 2.2.3 zur Verfügung. Die Stabilitätsbedingung (falls Stabilität möglich ist) wird in einem zweiten Schritt durch das System QEPCAD errechnet.

Es werden im folgenden die wichtigsten Prozeduren sowohl für den Fall einer einzelnen Gleichung als auch für den Fall eines Systems von Gleichungen vorgestellt werden. Einzelheiten mit Blick auf den Syntax und die Parameter der Prozeduren findet man nach Laden der Bibliothek durch den Befehl `help findiff.lib`; und nach Auflistung aller Prozeduren durch `help procname`; für eine Prozedur des Namens `procname`.

Es besteht im Rahmen der Benutzung von QEPCAD und der Bibliothek `findiff.lib` die Möglichkeit, daß dies völlig automatisch und insbesondere innerhalb einer SINGULAR-Sitzung geschieht. Dazu müssen allerdings einige Bedingungen erfüllt sein:

1. QEPCAD muß abschließend installiert sein.
2. Es muß ein spezielles Perl script `qepcadfilter.pl` in einem Verzeichnis (zum Beispiel im Verzeichnis `/home/cad/qesource/bin`) eingefügt werden.

Der Quelltext von `qepcadfilter.pl` ist:

```
#!/usr/bin/perl
$flag = 0;
$res = "";
while(<>)
{
  if ($_ =~ /Warning|WARNING|warning|Error|error|ERROR/){print $_; }
  elsif ($_ =~ /An\ equivalent/) { $flag = 1; }
  elsif ($flag == 1 && $_ ne "\n") { print $_; $flag = 0; }
}
```

Dieses Perl script stellt sicher, daß QEPCAD in der Benutzung durch SINGULAR einen besser zu verarbeitenden Output liefert.

3. Damit SINGULAR dauerhaft während mehrerer Sitzungen über die Bash auf das Programm QEPCAD zugreifen kann, muß zum Aufruf via `$qepcad` in der Datei über die Eigenschaften der Bash `/etc/bash.bashrc` der Zusatz

```
export qe=$HOME/cad/qesource/

export qepcad=$qe/bin/qepcad
```

eingefügt werden. Dies erspart die Definition der Variable `$qepcad` in einer Bash-Sitzung, da SINGULAR in den Prozeduren `qepcad` sowie `qepcadsystem` der Bibliothek `findiff.lib` das Programm QEPCAD mittels `qepcad` anspricht. Die Abänderung in `/etc/bash.bashrc` erfordert gegebenenfalls root-Rechte und kann in einer Bash-Sitzung nach Navigation in den Ordner `/etc` beispielsweise mit den Kommandos

```
$ cd /etc
$ sudo gedit bash.bashrc
```

nach Eingabe des Passworts erfolgen.

4. Der Name für QEPCAD in der Bash und der Pfad des Verzeichnisses, in welchem sich `qepcadfilter.pl` befindet, müssen in der Bibliothek `findiff.lib` innerhalb der Prozeduren `qepcad` und `qepcadsystem` an der Stelle

```
system("sh", "qepcad < QEPCAD-in | qepcadfilter.pl > QEPCAD-out");
```

geändert werden.

Einfache Gleichung

Die Prozeduren, welche hier dargestellt werden, sollen anhand eines einfachen Beispiels nachvollzogen werden. Es soll die Advektionsgleichung

$$u_t + au_x = 0$$

für eine unbekannte Funktion u gelöst und für u_x und u_t die Approximationen des Lax-Friedrich-Verfahrens verwendet werden. Zunächst wird in mehreren Schritten der zugrundeliegende Ring definiert:

```
> list D="Ux","Ut","U";
> list P="a";
> list V="t","x";
> setinitials(V,D,P);
```

Die Prozedur `setinitials` mit den Parametern V,D,P definiert den aktuellen Ring, dessen Eigenschaften man mithilfe des Befehls `basing` erhält:

```
> basing;
// characteristic : 0
// 8 parameter    : I T Px Cx Sx a dt dx
// minpoly        : 0
// number of vars : 5
//      block   1 : ordering c
//      block   2 : ordering lp
//              : names   i t x cx sx
// quotient ring from ideal
_[1]=cx^2+sx^2-1
_[2]=i^2+1
```

Es werden also die Ringvariablen x, t (als shift-Operatoren für die Argumente x, t der Funktion u), die trigonometrischen Funktionen sx, cx (für \sin_x, \cos_x) der räumlichen Variablen x (nicht der zeitlichen) und die imaginäre Einheit i innerhalb des Ringes `basing` definiert. Man beachte hierbei, daß der Ring vom Datentyp `qring` ist, d.h. es sich um einen Quotientenring (ein Ring modulo einem Ideal) handelt, wobei man das Ideal

$$\langle \cos_x^2 + \sin_x^2 - 1, i^2 + 1 \rangle$$

zur Darstellung des *Trigonometrischen Pythagoras* und der Eigenschaften der imaginären Einheit i automatisch definiert bekommt.

Darüber hinaus werden mit dx, dt die Gitterabstände Δ_x, Δ_t definiert und ebenso wie die Konstante a als Datentyp `number` eingeführt. Die parallele Benennung der ersten Parameter und Variablen dient der einfachen Umwandlung von Ringvariablen in Parameter (und vice versa) innerhalb von Prozeduren, je nachdem ob Teilbarkeit erforderlich ist oder aber nach Monomen geordnet werden soll, und ist in diesem Zusammenhang von keiner Bedeutung.

Im nächsten Schritt werden die Definition möglicher Approximationen und die Generierung des finiten Differenzschemas vorgenommen. Für den Fall der Linearität, der hier grundsätzlich vorausgesetzt wird, können die partielle Differentialgleichung und die zugehörigen Approximationen in Vektorform gegeben werden (siehe 4.32 und 4.34). Hier bleibt zu vermerken, daß die Reihenfolge der vorkommenden Ableitungen die Belegung auf die Einheitsvektoren festlegt:

```
> list D="Ux","Ut","U";
```

und damit gilt, daß für Ux der erste, respektive für Ut und U jeweils der zweite und dritte Einheitsvektor reserviert ist. Diese Zuordnung erledigt die Prozedur `u`:

```
> u(Ux);u(Ut);u(U);
[1]
[0,1]
[0,0,1]
```

Nun gibt man gewisse Approximationen ein und errechnet daraus ein Schema. Das obige Beispiel mit Lax-Friedrich-Approximationen versehen ergibt:

```
> def v1=u(Ut)+a*u(Ux);def v2=laxfrX(Ux,U,x);def v3=laxfrT(Ut,U,x);
> v1;v2;v3;
[(a),1]
[(2*dx)*x,0,-x^2+1]
[0,(dt)*x,-t*x+1/2*x^2+1/2]
```

Die vordefinierten Approximationen durch die Prozeduren `laxfrX` und `laxfrT` können entweder durch andere vorgegebene Approximationen und Prozeduren in `findiff.lib` oder aber beliebig von Hand gesetzt werden. Die Prozedur

Kapitel 5 SINGULAR und QEPCAD

`scheme` berechnet mittels Modulelimination der entsprechenden Komponenten ein Schema und nach Ausführung ergibt sich schließlich das Lax-Friedrich-Schema:

```
> def s1=scheme(v1,v2,v3);s1;
(-a*dt+dx)/(2*dx)*x^2+(a*dt+dx)/(2*dx)
```

Der nächste logische Schritt ist die Fourieranalyse, die mit der Prozedur `errormap` durchgeführt wird. Input für `errormap` ist das zuvor berechnete Schema:

```
> def s2=errormap(s1);s2;
(-a*dt+dx)/(dx)*i*cx*sx+(a*dt-dx)/(dx)*sx^2+1
```

Zur weitergehenden Abschätzung wird in der Von-Neumann-Analyse der *quadratierte* komplexe Betrag des Symbols, also $|\rho(\xi)|^2$, gebraucht. Denn es gilt

$$|\rho(\xi)|^2 \leq 1 \quad \Leftrightarrow \quad |\rho(\xi)| \leq 1$$

Dieses errechnet `ComplexValue`:

```
> def s3=ComplexValue(s2);s3;
(a^2*dt^2-dx^2)/(dx^2)*sx^2+1
```

Nachdem dies geschehen ist, wird das so erreichte Zwischenergebnis zur CAD-Analyse an QEPCAD gesendet. Damit sollen im Falle der Existenz zur Gleichung

$$|\rho(\xi)|^2 \leq 1$$

äquivalente Bedingungen berechnet werden, welche die erwünschten Stabilitätskriterien darstellen.

Vorausgesetzt, daß das Zusammenspiel zwischen SINGULAR und QEPCAD unter Beachtung der oben skizzierten Aspekte sichergestellt ist, soll die Art dieses Zusammenspiels kurz erläutert werden:

die Prozedur `qepcad` hat die Aufgabe, den quadrierten komplexen Betrag

```
> s3;
(a^2*dt^2-dx^2)/(dx^2)*sx^2+1
```

mit den entsprechenden Bedingungen und Quantoren in eine für QEPCAD lesbare Form zu bringen und diesen Ausdruck vom Datentyp `string` in eine Datei namens `QEPCAD-in` im `$HOME`-Verzeichnis zu schreiben. Diese Datei wird von QEPCAD ausgeführt, das Ergebnis in die Datei `QEPCAD-out` im selben

Verzeichnis geschrieben und - in eine lesbare Form gebracht - von SINGULAR ausgegeben.

Die Ausführung von `qepcad` mittels

```
> qepcad(s3);
```

liefert das von QEPCAD - aufgrund des Perl scripts in Kurzform - ausgewiesene Ergebnis in `QEPCAD-out` und gibt die einzelnen dort auftretenden Polynome zusätzlich für weitergehende Berechnungen mit aus:

```
sx /= 0 /\ [ [ dx + a dt > 0 /\ dx - a dt > 0 ]
\ / [ dx + a dt < 0 /\ dx - a dt < 0 ] ]
```

```
[1]:
```

```
  sx
```

```
[2]:
```

```
  (a*dt+dx)
```

```
[3]:
```

```
  (-a*dt+dx)
```

```
[4]:
```

```
  (a*dt+dx)
```

```
[5]:
```

```
  (-a*dt+dx)
```

Dies kann als die bekannte CFL-Bedingung $|a \frac{\Delta t}{\Delta x}| \leq 1$ interpretiert werden. Konkret an diesem Beispiel ergibt sich für `QEPCAD-in`:

```
[trial]
```

```
(cx,sx,a ,dt ,dx ,t)
```

```
5
```

```
(A t)[ + a^2 dt^2 sx^2 - dx^2  sx^2<0].
```

```
go
```

```
go
```

```
go
```

```
go
```

Die Datei `QEPCAD-in` ist für den Anwender besonders dann nützlich, wenn die Berechnungen in einer separaten QEPCAD-Sitzung gesondert berechnet oder Änderungen jedweder Art vorgenommen werden sollen.

Systeme

Hier werden die wesentlichen Prozeduren vorgestellt, die bei der automatischen Stabilitätsanalyse von Systemen partieller Differentialgleichungen verwendet werden. Die Vorgehensweise in der Implementierung entspricht derjenigen in Abschnitt 2.2.3.

Es sei für das System

$$u_t + Au_x \quad \text{mit } A = \begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix}$$

das Lax-Friedrich-Verfahren betrachtet, welches nun erzeugt und auf Stabilität untersucht werden soll.

Wie für einzelne Gleichungen werden zuerst die Parameter eingegeben und der Grundring erzeugt:

```
> list D="Ut","Ux","U";
> list V="t","x";
> list P="a","b";
> setinitials(V,D,P);
```

Anschließend werden die Daten für die Matrix A sowie die Approximationen für t und x definiert:

```
> matrix A[2][2]=-a,0,0,-b;
> list Mat=unitmat(2),A;
> list Appr=laxfrT(Ut,U,x),laxfrX(Ux,U,x);
```

Die erste Matrix `unitmat(2)` beschreibt die Identität für die Zeitableitung u_t . Die Prozeduren `matrixsystem` und `fouriersystem` liefern:

```
> matrixsystem(Mat,Appr);
_[1,1]=-1/(dt)*t*x+(a*dt+dx)/(2*dt*dx)*x^2+(-a*dt+dx)/(2*dt*dx)
_[1,2]=0
_[2,1]=0
_[2,2]=-1/(dt)*t*x+(b*dt+dx)/(2*dt*dx)*x^2+(-b*dt+dx)/(2*dt*dx)
> fouriersystem(Mat,Appr);
[1]:
  [1]:
    [1]:
```



```

    _[1,1]=(I*T*Sx+T*Cx)
    _[1,2]=0
    _[2,1]=0
    _[2,2]=(I*T*Sx+T*Cx)
[2]:
    _[1,1]=(I*Cx*Sx*a*dt+I*Cx*Sx*dx-Sx^2*a*dt-Sx^2*dx+dx)/(dx)
    _[1,2]=0
    _[2,1]=0
    _[2,2]=(I*Cx*Sx*b*dt+I*Cx*Sx*dx-Sx^2*b*dt-Sx^2*dx+dx)/(dx)
[2]:
[1]:
    _[1]=(I*T*Sx+T*Cx)
[2]:
    _[1]=(I*Cx*Sx*a*dt+I*Cx*Sx*dx-Sx^2*a*dt-Sx^2*dx+dx)/(dx)
    _[2]=(I*Cx*Sx*b*dt+I*Cx*Sx*dx-Sx^2*b*dt-Sx^2*dx+dx)/(dx)
[2]:
[1]:
[1]:
    (T^2)
[2]:
[1]:
    (Sx^2*a^2*dt^2-Sx^2*dx^2+dx^2)/(dx^2)
[2]:
    (Sx^2*b^2*dt^2-Sx^2*dx^2+dx^2)/(dx^2)

```

`matrixsystem` errechnet aus den Approximationen und den vorher definierten Matrizen eine Matrix als Summe, die sich nach Einsetzungen der Approximationen in die Gleichung ergibt. Diese Summe enthält je nach Approximation unterschiedliche Zeitschritte oder, polynomial in den Differenzenoperatoren ausgedrückt, unterschiedliche Grade des Operators T_t . Diese Grade höchster Ordnung werden in `fouriersystem` im Sinne einer Gleichung auf die eine Seite, die niedrigeren Grades auf die andere Seite gebracht und werden einer diskreten Fouriertransformation unterworfen ([1] [1] [1..2] der Ausgabe). Die Eigenwerte dieser Matrizen werden in [1] [2] [1] [1] und [1] [2] [2] [1..2] symbolisch berechnet, um danach ihren quadrierten komplexen Betrag in [2] [1] [1] und [2] [2] [1..2] zurückzugeben.

Die Errechnung der Stabilitätsbedingung für die Eigenwerte der Matrix mit niedrigerer Ordnung in T_t übernimmt `qpcadsystem`:

Kapitel 5 SINGULAR und QEPCAD

```
> qepcadsystem(_[2]);
0
0
0
sx /= 0 /\ [ [ dx + a dt > 0 /\ dx - a dt > 0
              /\ dx + b dt > 0 /\ dx - b dt > 0 ]
            \/[ dx + a dt < 0 /\ dx - a dt < 0
              /\ dx + b dt < 0 /\ dx - b dt < 0 ] ]

[1]:
  sx
[2]:
  (a*dt+dx)
[3]:
  (-a*dt+dx)
[4]:
  (b*dt+dx)
[5]:
  (-b*dt+dx)
[6]:
  (a*dt+dx)
[7]:
  (-a*dt+dx)
[8]:
  (b*dt+dx)
[9]:
  (-b*dt+dx)
```

Das Resultat ist wieder die CFL-Bedingung, diesmal für a und b :

$$\left|a \frac{\Delta t}{\Delta x}\right| \leq 1 \text{ und } \left|b \frac{\Delta t}{\Delta x}\right| \leq 1.$$

Literaturverzeichnis

- [ACM98a] Dennis S. Arnon, George E. Collins, and Scott McCallum. Cylindrical algebraic decomposition. I. The basic algorithm. In *Quantifier elimination and cylindrical algebraic decomposition (Linz, 1993)*, Texts Monogr. Symbol. Comput., pages 136–151. Springer, Vienna, 1998.
- [ACM98b] Dennis S. Arnon, George E. Collins, and Scott McCallum. Cylindrical algebraic decomposition. II. An adjacency algorithm for the plane. In *Quantifier elimination and cylindrical algebraic decomposition (Linz, 1993)*, Texts Monogr. Symbol. Comput., pages 152–165. Springer, Vienna, 1998.
- [AM69] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969.
- [BPR03] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry*. Algorithms and Computation in Mathematics. 10. Springer, 2003.
- [Bro99] Christopher W. Brown. *Solution formula construction for truth invariant cads*. PhD thesis, Newark, DE, USA, 1999. Professor In Charge-Collins, George E.
- [Bro03] Christopher W. Brown. QEPCADB: A program for computing with semi-algebraic sets using cads. *SIGSAM Bull.*, 37(4):97–108, 2003.
- [Bro10] Christopher W. Brown. QEPCADB 1.54 — A program for quantifier elimination and formula simplification using cylindrical algebraic decomposition. 2010. <http://www.usna.edu/Users/cs/qepcad/B/QEPCAD.html>.

Literaturverzeichnis

- [Buc83] Bruno Buchberger. A note on the complexity of constructing gröbner-bases. In *EUROCAL '83: Proceedings of the European Computer Algebra Conference on Computer Algebra*, pages 137–145, London, UK, 1983. Springer-Verlag.
- [CJ98] B. F. Caviness and J. R. Johnson, editors. *Quantifier elimination and cylindrical algebraic decomposition*, Texts and Monographs in Symbolic Computation, Vienna, 1998. Springer-Verlag.
- [CLO97] David A. Cox, John Little, and Donal O’Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 2/e (Undergraduate Texts in Mathematics)*. Springer-Verlag New York, Inc., 1997.
- [Col75] George E. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. *Autom. Theor. form. Lang.*, 2nd GI Conf., Kaiserslautern 1975, Lect. Notes Comput. Sci. 33, 134-183 (1975)., 1975.
- [DGPS09] W. Decker, G.-M. Greuel, G. Pfister, and H. Schönemann. SINGULAR 3-1-0 — A computer algebra system for polynomial computations. 2009. <http://www.singular.uni-kl.de>.
- [DH88] James H. Davenport and Joos Heintz. Real quantifier elimination is doubly exponential. *J. Symbolic Comput.*, 5(1-2):29–35, 1988.
- [Din10] Christian Dingler. `findiff.lib` — A SINGULAR 3-1-1 library for computing finite difference schemes for linear differential equations. 2010. <http://www.singular.uni-kl.de>.
- [DL10] C. Dingler and V. Levandovskyy. Algebraic Foundations for Finite Difference Schemes. July 2010. <http://ftp.math.rwth-aachen.de/~Viktor.Levandovskyy/filez/algfds.pdf>.
- [DS97] Andreas Dolzmann and Thomas Sturm. Redlog: Computer algebra meets computer logic. *ACM SIGSAM Bulletin*, 31(2):2–9, 1997.
- [Eis95] David Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry.

- [For99] Otto Forster. *Analysis Band 2 Differential- und Integralrechnung im \mathbb{R}^n , Gewöhnliche Differentialgleichungen*. Vieweg-Studium 31 : Grundkurs Mathematik. Vieweg, Braunschweig, 5 edition, 1999.
- [For01] Otto Forster. *Analysis Band 1 Differential- und Integralrechnung einer Veränderlichen*. Vieweg-Studium: Grundkurs Mathematik. Vieweg, Braunschweig, 6 edition, 2001.
- [FR74] Michael J. Fischer and Michael O. Rabin. Super-exponential complexity of Presburger arithmetic. In *Complexity of computation (Proc. SIAM-AMS Sympos., New York, 1973)*, pages 27–41. SIAM-AMS Proc., Vol. VII. Amer. Math. Soc., Providence, R.I., 1974.
- [FW67] George E. Forsythe and Wolfgang R. Wasow. *Finite-difference methods for partial differential equations*. Wiley, New York, 1967.
- [GB05] Vladimir P. Gerdt and Yuri A. Blinkov. Janet-like Gröbner bases. In *Computer algebra in scientific computing*, volume 3718 of *Lecture Notes in Comput. Sci.*, pages 184–195. Springer, Berlin, 2005.
- [GB09] Vladimir Gerdt and Yuri Blinkov. Involution and difference schemes for the Navier-Stokes equations. *Proceedings CASC 2009, Kobe, Japan*, 2009.
- [GBM06] Vladimir P. Gerdt, Yuri A. Blinkov, and Vladimir V. Mozzhilkin. Gröbner bases and generation of difference schemes for partial differential equations. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 2:Paper 051, 26 pp. (electronic), 2006.
- [Ger05] V. Gerdt. Involution algorithms for computing Groebner bases. In G. Pfister, S. Cojocaru, and V. Ufnarovski, editors, *Computational Commutative and Non-Commutative Algebraic Geometry*. IOS Press, 2005.
- [GLS06] G.-M. Greuel, V. Levandovskyy, and H. Schönemann. PLURAL. A SINGULAR 3.0 Subsystem for Computations with Non-commutative Polynomial Algebras. Centre for Computer Algebra, University of Kaiserslautern, 2006.

Literaturverzeichnis

- [GM80] David F. Griffiths and Andrew R. Mitchell. *The finite difference method in partial differential equations*. Wiley, Chichester [u.a.], 1980.
- [GP08] Gert-Martin Greuel and Gerhard Pfister. *A **Singular** introduction to commutative algebra*. Springer, Berlin, extended edition, 2008. With contributions by Olaf Bachmann, Christoph Lossen and Hans Schönemann, With 1 CD-ROM (Windows, Macintosh and UNIX).
- [GR10] Vladimir Gerdt and Daniel Robertz. Consistency of finite difference approximations for linear pde systems and its algorithmic verification. In *Proc. of the International Symposium on Symbolic and Algebraic Computation (ISSAC'10)*. ACM Press, 25-28 July 2010.
- [Gre09] Gert-Martin Greuel. `teachstd.lib` — A SINGULAR 3-1-1 library for teaching standard bases. 2009. <http://www.singular.uni-kl.de>.
- [GV96] V. Ganzha and E. Vorozhtsov. *Computer-Aided Analysis of Difference Schemes for Partial Differential Equations*. Wiley Interscience, 1996.
- [Had53] Jacques Hadamard. *Lectures on Cauchy's problem in linear partial differential equations*. Dover Publications, New York, 1953.
- [HLS97] H. Hong, R. Liska, and S. Steinberg. Testing stability by quantifier elimination. *Journal of Symbolic Computation*, 24(2):161–187, 1997.
- [Hon98a] Hoon Hong. An improvement of the projection operator in cylindrical algebraic decomposition. In *Quantifier elimination and cylindrical algebraic decomposition (Linz, 1993)*, Texts Monogr. Symbol. Comput., pages 166–173. Springer, Vienna, 1998.
- [Hon98b] Hoon Hong. Simple solution formula construction in cylindrical algebraic decomposition based quantifier elimination. In *Quantifier elimination and cylindrical algebraic decomposition (Linz, 1993)*, Texts Monogr. Symbol. Comput., pages 201–219. Springer, Vienna, 1998.

- [JS06] Winfried Just and Brandilyn Stigler. Computing gröbner bases of ideals of few points in high dimensions. *ACM Commun. Comput. Algebra*, 40(3-4):67–78, 2006.
- [Kai08] AG Technomathematik (TU Kaiserslautern). Vorlesungsskript: Numerik partieller Differentialgleichungen II. WS 2007/2008. Download via <http://www.agtm.mathematik.uni-kl.de/agtm/home/teaching/alllectures.html>.
- [Kev00] Jirair Kevorkian. *Partial differential equations*. Springer, New York [u.a.], 2. ed. edition, 2000.
- [KS82] K. Keller and A. Schneider. Ein funktionalanalytischer Beweis des Satzes von Cauchy-Kowalewsky. *Manuscripta Math.*, 39(1):31–37, 1982.
- [LeV92] Randall J. LeVeque. *Numerical methods for conservation laws*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 1992.
- [Lev05] V. Levandovskyy. Non-commutative computer algebra for polynomial algebras: Gröbner bases, applications and implementation. *Ph.D. Thesis, Universität Kaiserslautern*, 2005.
- [Lev06] Levandovskyy, V. Intersection of ideals with non-commutative subalgebras. In *Proc. of the International Symposium on Symbolic and Algebraic Computation (ISSAC'06)*. ACM Press, 2006.
- [LM10] V. Levandovskyy and B. Martin. A Symbolic Approach to Generation and Analysis of Finite Difference Schemes of Partial Differential Equations. *ArXiv e-prints*, July 2010.
- [LSL09] Roberto La Scala and Viktor Levandovskyy. Letterplace ideals and non-commutative Gröbner bases. *Journal of Symbolic Computation*, 44(10):1374–1393, 2009.
- [McC98] Scott McCallum. An improved projection operation for cylindrical algebraic decomposition. In *Quantifier elimination and cylindrical algebraic decomposition (Linz, 1993)*, Texts Monogr. Symbol. Comput., pages 242–268. Springer, Vienna, 1998.

Literaturverzeichnis

- [Mor94] Mora, T. An introduction to commutative and non-commutative Groebner bases. *Theor. Comp. Sci.*, 134:131–173, 1994.
- [Sch03] R. Schneider. Explicit and implicit finite-difference methods for the diffusion equation in two dimensions. *Wissenschaftliche Berichte FZKA*, 6866:43, 2003.
- [Smi70] Gordon D. Smith. *Numerische Lösung von partiellen Differentialgleichungen*. Vieweg, Braunschweig, studienausgabe edition, 1970.
- [SST00] S. Saito, B. Sturmfels, and N. Takayama. *Gröbner Deformations of Hypergeometric Differential Equations*. Springer, 2000.
- [Tho95] J. W. Thomas. *Numerical partial differential equations: finite difference methods*, volume 22 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1995.
- [Tho99] J. W. Thomas. *Numerical partial differential equations*, volume 33 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999. Conservation laws and elliptic equations.
- [vdH96] P. J. van der Houwen. The development of runge-kutta methods for partial differential equations. *Appl. Numer. Math.*, 20(3):261–272, 1996.
- [Wal02] Wolfgang Walter. *Analysis 2*. Springer, Berlin [u.a.], 2002.
- [Wal04] Wolfgang Walter. *Analysis 1*. Springer, Berlin [u.a.], 2004.
- [Šo83] Jurij I. Šokin. *The method of differential approximation*. Springer, Berlin [u.a.], 1983.

Wissenschaftlicher Werdegang

Name	Christian Dingler
Geburtsdatum	7. August 1983
Geburtsort	Kaiserslautern (Deutschland)

1990-1994 Martin-Luther Schule Neuss (Grundschule)

1994-2003 Quirinus-Gymnasium Neuss

2003 Allgemeine Hochschulreife

2003-2007 Studium der Mathematik mit Nebenfach Wirtschaftswissenschaften an der Technischen Universität Kaiserslautern (Diplomstudiengang)

2007 Diplom in Mathematik

2007-2010 Doktorand am Fachbereich Mathematik (TU Kaiserslautern)

Scientific Career

Name	Christian Dingler
Date of Birth	August 7th, 1983
Place of Birth	Kaiserslautern, Germany

1990-1994 Martin-Luther Schule Neuss (primary school)

1994-2003 Quirinus-Gymnasium Neuss (grammar school)

2003 general qualification for university entrance

2003-2007 Study of Mathematics with Economic Sciences as secondary subject at the University of Kaiserslautern

2007 Diploma in Mathematics

2007-2010 doctoral student at the department of Mathematics of the University of Kaiserslautern