Algorithmische Aspekte der Quantorenelimination auf den reellen Zahlen

TECHNISCHE UNIVERSITÄT DARMSTADT

Ausarbeitung im Seminar Logik "Quant*", Prof. Dr. Martin Ziegler, WS 2010/2011
Felix Günther

1 Einleitung [12]

Definition 1.1 (Quantorenelimination)**.** Sei L eine Sprache und T eine Theorie. T hat *Quantorenelimination*, wenn es zu jeder L-Formel $\varphi(x_1,\dots,x_n)$ eine quantorenfreie Formel $\psi(x_1,\dots,x_n)$ gibt mit

$$T \vdash \forall x_1,\dots,x_n \left(\varphi(x_1,\dots,x_n) \leftrightarrow \psi(x_1,\dots,x_n)\right).$$

Als *Quantorenelimationsproblem* bezeichnen wir: Gegeben eine Formel φ, finde die unter T äquivalente quantorenfreie Formel ψ.

Definition 1.2 (Einfache und primitive Existenzformeln)**.** Eine *einfache Existenzformel* hat die Form

$$\varphi(\bar{x}) = \exists y\ \psi(\bar{x}, y)$$

für eine quantorenfreie Formel ψ. Wenn ψ eine Konjunktion von atomaren Formeln[1] oder negierten atomaren Formeln ist, dann heißt φ *primitiv*.

Lemma 1.3. *T hat genau dann Quantorenelimination, wenn jede primitive Existenzformel modulo T zu einer quantorenfreien Formel äquivalent ist.*

Beweis. $\Longrightarrow$: Trivial.

$\Longleftarrow$: Jede einfache Existenzformel $\exists y \bigvee_{i<n} \psi_i$ ist äquivalent zu einer Disjunktion $\bigvee_{i<n}(\exists y\ \psi_i)$ von primitiven Existenzformeln. Damit ist jede einfache Existenzformel modulo T äquivalent zu einer quantorenfreien Formel.

Wir betrachten eine beliebige Formel in pränexer Normalform:

$$Q_1 x_1 \dots Q_n x_n\ \psi$$

Wenn $Q_n = \exists$, dann wählen wir eine quantorenfreie Formel ψ', die modulo T zu $\exists x_n\ \psi$ äquivalent ist und fahren mit der Formel $Q_1 x_1 \dots Q_{n-1} x_{n-1} \psi'$ fort. Wenn $Q_n = \forall$, dann wählen wir eine quantorenfreie Formel ψ', die modulo T zur $\exists x_n \neg\psi$ äquivalent ist und fahren mit der Formel $Q_1 x_1 \dots Q_{n-1} x_{n-1} \neg\psi'$ fort. □

Beispiel 1.4. Wir betrachten folgende Beispiele für Quantorenelimination auf $\mathbb{R}$:

a) Die Formel

$$\varphi(x) = \exists y\ [x^2 + y = 4 \wedge y \geq 0]$$

ist äquivalent zu folgender quantorenfreien Formel:

$$\psi(x) = x \leq 2 \wedge x \geq -2$$

[1] Atomare Formeln sind $t_1 \doteq t_2$ und $Rt_1,\dots,t_n$ für L-Terme $t_1,\dots t_n$. Im Fall der Theorie der reellen Zahlen und der Sprache der geordneten Körper, die wir im Folgenden betrachten, haben atomare Formeln die Form $p_1 = p_2$ oder $p_1 < p_2$ für Polynome p_1, p_2.

b) Die Formel

$$\varphi = \forall x\, \exists y\, [x > y^2]$$

enthält keine freien Variablen; ihr kann also sogar ein konstanter Wahrheitswert (in diesem Fall *false* bzw. $\bot$) zugewiesen werden.

Eine geometrische Interpretation der Tatsache, dass die Theorie der reellen Zahlen in der Sprache der geordneten Körper Quantorenelimination hat[2], ist die folgende:

Theorem 1.5. *Die Projektion von semialgebraischen Mengen ist wiederum semialgebraisch.*

Beispiel 1.6. Ein einfaches Beispiel hierfür ist die Kugel im $\mathbb{R}^3$. Die Formel

$$\varphi(x, y) = \exists z\, [x^2 + y^2 + z^2 \leq 1]$$

ist äquivalent zur quantorenfreien Formel

$$\psi(x, y) = x^2 + y^2 \leq 1$$

die sich geometrisch als Projektion der Kugel auf die Ebene interpretieren lässt. Beide Formeln sind wahr für Punkte (x, y) auf dem oder im Einheitskreis.

2 Quantorenelimination nach Muchniks Beweis des Tarski-Seidenberg-Theorems [8]

Wir betrachten nun ein konkretes Verfahren für Quantorenelimination, das Andrej Muchnik als Beweis des Tarski-Seidenberg-Theorems konstruiert hat [8]. Hierzu führen wir zunächst die Notation von Vorzeichendiagrammen und Muchnik-Mengen ein, bevor wir im Anschluss an den Beweis des Theorems den zentralen Algorithmus detailliert beschreiben.

2.1 Vorzeichendiagramme

Sei $p(Y)$ ein Polynom in einer Variable Y mit den reellen Nullstellen $\xi_2 < \xi_4 < \cdots < \xi_{2d}$ (ohne Vielfachheit). Das Vorzeichen von p ist auf jedem Zwischenintervall (ξ_{2i}, ξ_{2i+2}) konstant. Sei ξ_{2i+1} $(1 \leq i \leq d-1)$ ein beliebiger Wert im Intervall (ξ_{2i}, ξ_{2i+2}) und ξ_1 bzw. ξ_{2d+1} ein beliebiger Wert kleiner als ξ_2 bzw. größer als ξ_{2d}. Die ξ_i bezeichnen wir als *Testpunkte*.

Wir definieren

$$\operatorname{sgn} a = \begin{cases} 0 & \text{für } a = 0 \\ 1 & \text{für } a > 0 \\ -1 & \text{für } a < 0 \end{cases}$$

Definition 2.1. Die *Vorzeichenentwicklung* oder *Vorzeichenreihe* von p ist

$$(\operatorname{sgn} p(\xi_1), \operatorname{sgn} p(\xi_2), \ldots, \operatorname{sgn} p(\xi_{2d+1})).$$

Im Folgenden arbeiten wir mit mehreren Polynomen gleichzeitig, daher müssen ab jetzt Testpunkte mit geradem Index nicht zwingend Nullstellen sein.

Sei P eine endliche Liste von Polynomen in einer Variablen Y mit reellen Koeffizienten und D eine $m \times n$-Matrix mit Elementen aus $\{-1, 0, 1\}$, deren i-te Zeile die Vorzeichenentwicklung des i-ten Polynoms sein soll. Wir beschriften Zeilen der Matrix mit dem jeweiligen Polynom und Spalten mit dem jeweiligen Testpunkt, d. h. $D(p, \xi)$ bezeichnet den Eintrag von D in der mit p beschrifteten Zeile und mit ξ beschrifteten Spalte.

Definition 2.2 (Vorzeichendiagramm)**.** Sei D eine derartige Matrix. D ist ein *Vorzeichendiagramm* von P, wenn es ein n-Tupel reeller Zahlen $\xi = (\xi_1, \ldots, \xi_n)$ mit $\xi_1 < \xi_2 < \cdots < \xi_n$ gibt, so dass gilt:

[2] Siehe Tarski-Seidenberg-Theorem, 2.11.

- Jede Nullstelle eines Polynoms $p \in P$ (mit $p(Y) \neq 0$) ist unter den ξ_i.
- Zwischen je zwei solcher Nullstellen liegt mindestens ein weiteres ξ_j, das keine Nullstelle irgendeines Polynoms $p \in P$ (mit $p(Y) \neq 0$) ist.
- $D(p, \xi_i) = \operatorname{sgn} p(\xi_i)$

Beispiel 2.3. Sei P die Liste der Polynome $p_1(Y) = 2Y^2$, $p_2(Y) = Y + 1$, $p_3(Y) = 3$, $p_4(Y) = 0$.

Ein gültiges Vorzeichendiagramm hierfür wäre

$p \backslash \xi_i$	-2	-1	$0,5$	0	1
$2Y^2$	1	1	1	0	1
$Y+1$	-1	0	1	1	1
3	1	1	1	1	1
0	0	0	0	0	0

Kein gültiges Vorzeichendiagramm für P wäre

$p \backslash \xi_i$	-2	-1	0	1
$2Y^2$	1	1	0	1
$Y+1$	-1	0	1	1
3	1	1	1	1
0	0	0	0	0

da hier kein Testpunkt zwischen den beiden Nullstellen von p_1 und p_2 liegt.

In einem Vorzeichendiagramm D für eine Liste P von Polynomen gilt:

a) Die mit p beschriftete Zeile enthält die Vorzeichenreihe von p sowie eventuell die Vorzeichen an weiteren Punkten.

b) Wenn $p = 0$, dann sind auch alle Einträge in der Zeile p in D gleich 0. Wir bezeichnen eine solche Zeile als *Null-Zeile*.

c) Zwei benachbarte Spalten können nicht gleich sein, außer ihre einzigen 0-Einträge liegen in Null-Zeilen. Lässt man eine der beiden Spalten weg, ist D immer noch ein Vorzeichendiagramm. Lässt man alle solche redundanten Spalten weg, erhält man ein eindeutiges, minimales Vorzeichendiagramm für P, das *reduzierte Vorzeichendiagramm* für P, bezeichnet mit $\operatorname{diag}(P)$.

d) Sind alle $p \in P$ konstant, so enthält $\operatorname{diag}(P)$ nur eine einzelne Spalte die die Vorzeichen der jeweiligen Konstanten enthält.

e) Zu jedem $y \in \mathbb{R}$ gibt es eine Spalte ξ, so dass $D(p, \xi) = \operatorname{sgn} p(y)$ für alle $p \in P$.

f) Zwei benachbarte Einträge in einer Zeile können weder beide 0 sein (abgesehen von Null-Zeilen), noch können sie entgegengesetzte Vorzeichen haben.

2.2 Muchnik-Mengen

Sei A ein Integritätsbereich (d. h. ein nullteilerfreier kommutativer Ring mit einem Einselement, dass von Null verschieden ist), Y eine einzelne Variable und

$$p := a_d Y^d + a_{d-1} Y^{d-1} + \cdots + a_1 Y + a_0$$

ein Polynom in $A[Y]$ vom Grad d mit $a_d \neq 0$. Den Grad des Nullpolynoms $p(Y) = 0$ definieren wir als $-\infty$.

Wir bezeichnen mit p° das Abschneiden des höchsten Terms von p, d. h.

$$p^\circ := a_{d-1} Y^{d-1} + \cdots + a_1 Y + a_0$$

Sei $q = b_e Y^e + \cdots + b_1 Y^1 + b_0$ ein Polynom in $A[Y]$ ungleich 0 mit $e \leq d$.

Definition 2.4 (Pseudorest)**.** Mittels Polynomdivision von p durch q erhält man eindeutige Polynome h und r, mit Grad von r kleiner e, so dass

$$b_e^{d-e+1} p = hq + r.$$

Wir nennen r den *Pseudorest* von p und q und bezeichnen ihn mit $\mathrm{rem}(p,q)$.[3]

Definition 2.5 (Muchnik-Menge)**.** Sei L eine Menge von Polynomen (einschließlich Konstanten) in $A[Y]$. L ist eine *Muchnik-Menge*, wenn die folgenden drei Eigenschaften gelten:

(**T**) $p \in L \implies p^\circ \in L$

(**D**) $p \in L \implies p' \in L$

(**R**) $p, q \in L \implies \mathrm{rem}(p,q) \in L$

Es gilt:

a) Jede Muchnik-Menge enthält das 0-Polynom.

b) Jede endliche Menge von Konstanten, die 0 enthält, ist eine Muchnik-Menge.

Wir bezeichnen die Menge aller Konstanten eine Muchnik-Menge L mit L_0.

Anmerkung 2.6. Die obigen drei Operationen sind Hüllenoperationen, d. h. wir können für jede Menge L' von Polynomen die kleinste Muchnik-Menge L angeben, die L' enthält. Wir bezeichnen diese als *Muchnik-Hülle*. Da die Anwendung jeder der drei Operationen jeweils ein Polynom von kleinerem Grad liefert, ist die Muchnik-Hülle einer endlichen Menge endlich.

Beispiel 2.7. Sei L' die bekannte Liste von Polynome $p_1(Y) = 2Y^2$, $p_2(Y) = Y + 1$, $p_3(Y) = 3$, $p_4(Y) = 0$. Durch wiederholte Anwendung der Regeln (**T**), (**D**), (**R**) können wir nun die Muchnik-Hülle L von L' bilden:

- Die Muchnik-Hülle muss natürlich die Polynome p_1, p_2, p_3 und p_4 enthalten: $L \supseteq \{2Y^2, Y+1, 3, 0\}$.
- Regelanwendung:

  (**T**) $(2Y^2)^\circ = 0$ (bereits enthalten), $(Y+1)^\circ = 1$ (1 kommt hinzu), $(3)^\circ = (0)^\circ = 0$ (bereits enthalten)

  (**D**) $(2Y^2)' = 4Y$ ($4Y$ kommt hinzu), $(Y+1)' = 1$ (bereits enthalten), $(3)' = (0)' = 0$ (bereits enthalten)

  (**R**) $\mathrm{rem}(2Y^2, Y+1) = 2$ (2 kommt hinzu)[4], weitere Pseudorest-Bildungen sind nicht möglich
- Zur Muchnik-Hülle sind 1, $4Y$ und 2 hinzugekommen: $L \supseteq \{2Y^2, Y+1, 3, 0, 1, 4Y, 2\}$.
- Regelanwendung:

  (**T**) $(4Y)^\circ = (2)^\circ = (1)^\circ = 0$ (bereits enthalten)

  (**D**) $(4Y)' = 4$ (4 kommt hinzu), $(2)' = (1)' = 0$ (bereits enthalten)

  (**R**) $\mathrm{rem}(2Y^2, 4Y) = 0$ (bereits enthalten), $\mathrm{rem}(4Y, Y+1) = -4$ (-4 kommt hinzu)
- Zur Muchnik-Hülle sind 4 und -4 hinzugekommen: $L \supseteq \{2Y^2, Y+1, 3, 0, 1, 4Y, 2, 4, -4\}$.
- Weitere Regelanwendungen fügen keine neuen Elemente hinzu, also ist die Muchnik-Hülle von L':

$$L = \{2Y^2, Y+1, 3, 0, 1, 4Y, 2, 4, -4\}$$

[3] Der Faktor b_e^{d-e+1} vor p rührt daher, dass der Leitkoeffizient von q nicht zwingend eine Einheit auf dem Integritätsbereich ist (d. h. nicht immer invertierbar ist).

[4] Wir suchen h, r mit Grad von r kleiner 1 ($e = 1$, $b_e = 1$), so dass $p = hq + r$, also $2Y^2 = h \cdot (Y+1) + r$. Dies gilt offensichtlich für $h = 2Y - 2$ und $r = 2$.

Beispiele für Mengen von Polynomen, die *keine* Muchnik-Mengen sind, wären

$\{1\}$	$(1)' = 0$ nicht enthalten
$\{Y+2, 1, 0\}$	$(Y+2)^\circ = 2$ nicht enthalten
$\{Y^2-2, 2Y, Y, 2, -2, 1, 0\}$	$\mathrm{rem}(Y^2-2, Y) = -8$ nicht enthalten

Definition 2.8 (Muchnik-Liste)**.** Sei L eine Muchnik-Menge. Jede Aufzählung der Elemente von L, die diese nicht-absteigend nach ihrem Grad anordnet, bezeichnen wir als *Muchnik-Liste*.

Eine Muchnik-Liste L beginnt also mit 0, gefolgt von den verbleibenden Elementen in L_0 in beliebiger Reihenfolge und schließt dann die Elemente höheren Grades von L nach ansteigendem Grad sortiert an.

Beispiel 2.9. Die in Beispiel 2.7 hergeleitete Muchnik-Hülle $L = \{2Y^2, Y+1, 3, 0, 1, 4Y, 2, 4, -4\}$ der Polynommenge L' lautet als Muchnik-Liste notiert:

$$\{0, 1, 2, 3, 4, -4, Y+1, 4Y, 2Y^2\}$$

Lemma 2.10. *Sei L eine Muchnik-Liste, dann ist jedes Anfangssegment von L wiederum eine Muchnik-Liste.*

Beweis. Wir führen Induktion über die Länge des Anfangssegments M von L.

Wenn M die Länge 1 hat, dann gilt per Definition $M = \{0\}$, also ist M trivialerweise wieder eine Muchnik-Liste.

Sei nun M das Anfangssegment von L und p sein letztes Element. Wir zeigen, dass M eine Muchnik-Liste ist, vorausgesetzt, $M \setminus \{p\}$ ist bereits eine Muchnik-Liste. Nach Regel (**T**) und (**D**) gilt $p^\circ, p' \in L$. Da p° und p' geringeren Grad als p haben, müssen sie in L vor p aufgezählt werden und damit auch in M vorkommen. Wenn $q \in M$, dann haben sowohl $\mathrm{rem}(p, q)$ als auch $\mathrm{rem}(q, p)$ geringeren Grad als p (da q, wenn es in M liegt, vom Grad kleiner/gleich dem von p sein muss). Nach Regel (**R**) liegen beide Pseudoreste in L, also auch in M. □

Im Folgenden betrachten wir eine Muchnik-Liste immer als Spalte, deren Einträge als Bezeichner für die Zeilen von Matrizen (Vorzeichendiagrammen) verwendet werden.

2.3 Quantorenelimination

Sei L eine Spalte der Länge m von Elementen in $\mathbb{R}[X][Y]$ mit $X = (X_1, \ldots, X_N)$ und einzelner Variable Y. Für ein $\vec{x} \in \mathbb{R}^N$ bezeichnen wir mit $L(\vec{x})$ die Spalte $(p_1(\vec{x}, Y), \ldots, p_m(\vec{x}, Y))$ von Polynomen in $\mathbb{R}[Y]$. Dies ist im Besonderen auch auf eine Muchnik-Liste L und ihre Teilliste L_0 von Konstanten (bezogen auf Y) anwendbar; wir bezeichnen die Länge von L_0 mit m_0.

Im nächsten Abschnitt wir ein Algorithmus vorgestellt, der, gegeben L und eine Spalte C, beschriftet mit den Einträgen aus L_0 (d. h. eine $m_0 \times 1$-Matrix) mit Einträgen aus $\{-1, 0, 1\}$, eine Matrix $D \in \{-1, 0, 1\}^{m \times n}$ berechnet (mit Zeilen beschriftet mit den Einträgen aus L), so dass für jedes $\vec{x} \in \mathbb{R}^N$ gilt

$$C = \mathrm{diag}(L_0(\vec{x})) \implies D = \mathrm{diag}(L(\vec{x})).$$

Der Algorithmus geht hierbei schrittweise vor, indem er zu einem gegebenen Vorzeichendiagramm ein einzelnes Polynom hinzufügt.

Da L_0 nur Konstanten enthält ist $\mathrm{diag}(L_0(\vec{x}))$ einfach die Spalte die die Vorzeichen $\mathrm{sgn}\, p$ der Konstanten $p \in L_0$ enthält. Insbesondere gilt, dass die ersten m_0 Zeilen von D (also die mit Elementen aus L_0 beschrifteten) konstant sind, genauer immer den entsprechenden Eintrag aus C enthalten.

Unter der Annahme, dass ein solcher Algorithmus existiert, können wir nun das Tarski-Seidenberg-Theorem beweisen.

Theorem 2.11 (Tarski-Seidenberg)**.** *Der Körper der reellen Zahlen erlaubt Quantorenelimination in der Sprache der geordneten Körper.*

Beweis. Nach Lemma 1.3 reicht es zu zeigen, dass eine primitive Existenzformel $\exists y\ \varphi(\vec{x}, y)$ mit quantorenfreiem φ, $\vec{x} = (x_1, \ldots, x_N)$ und einzelner Variable y äquivalent zu einer quantorenfreien Formel ist.

O.b.d.A. können wir annehmen, dass φ eine Konjunktion von Formeln der Form $\operatorname{sgn} p(\vec{x}, y) = \varepsilon_p$ mit reellem Polynom p und $\varepsilon_p \in \{-1, 0, 1\}$ ist (vgl. Anmerkung 2.12). Sei L' die Menge aller Polynome die so in φ vorkommen und L die Muchnik-Hülle von L', angeordnet als Muchnik-Liste der Länge m.

Wir bezeichnen eine Matrix $D \in \{-1, 0, 1\}^{m \times n}$ als *φ-kompatibel*, wenn es eine Spalte ξ gibt mit $D(p, \xi) = \varepsilon_p$ für alle $p \in L'$. Für ein $\vec{x} \in \mathbb{R}^N$ gilt $\exists y\ \varphi(\vec{x}, y)$ genau dann, wenn $\operatorname{diag}(L(\vec{x}))$ φ-kompatibel ist. Dies gilt, da, wenn $\operatorname{diag}(L(\vec{x}))$ φ-kompatibel ist eine Spalte ξ existiert mit $D(p, \xi) = \varepsilon_p$, d. h. $\operatorname{sgn} p(\vec{x}, \xi) = \varepsilon_p$, also eine Belegung $y = \xi$, die alle konjugierten Formeln $\operatorname{sgn} p(\vec{x}, y) = \varepsilon_p$ in φ war macht, also ist $\exists y\ \varphi(\vec{x}, y)$ wahr. Die Rückrichtung gilt ebenfalls.

Sei $C_1, \ldots, C_l$ eine Aufzählung aller möglichen Spalten der Höhe m_0 mit Einträgen in $\{-1, 0, 1\}$ und $\psi_i(\vec{x})$ die quantorenfreie Formel die

$$\operatorname{diag}(L_0(\vec{x})) = C_i$$

ausdrückt. Sei D_i die Matrix, die man durch Anwendung des nachfolgend in Abschnitt 2.4 beschriebenen Algorithmus auf C_i erhält. Sei $I \subset \{1, \ldots, l\}$ die Menge der Indizes, für die D_i φ-kompatibel ist. Es folgt, dass

$$\exists y\ \varphi(\vec{x}, y) \leftrightarrow \bigvee_{i \in I} \psi_i(\vec{x}).$$

Die rechte Seite ist quantorenfrei. □

Anmerkung 2.12. Da $\exists y\ \varphi(\vec{x}, y)$ eine primitive Existenzformel ist, ist φ eine Konjunktion über atomare Formeln und negierte atomare Formeln (vgl. Definition 1.2). Als Beispiel könnte φ also wie folgt aussehen:

$$(p_1 = p_2 \wedge p_2 < p_3 \wedge \neg p_3 < p_4 \wedge \ldots)$$

Dieses als Beispiel gegebene φ ist äquivalent zu

$$(\operatorname{sgn} p_1 - p_2 = 0 \wedge \operatorname{sgn} p_2 - p_3 = -1 \wedge \neg \operatorname{sgn} p_3 - p_4 = -1 \wedge \ldots).$$

Enthält φ keine negierte atomaren Formeln ist damit die Transformation zu einer Konjunktion von Formeln der Form $\operatorname{sgn} p(\vec{x}, y) = \varepsilon_p$ mit reellem Polynom p und $\varepsilon_p \in \{-1, 0, 1\}$, wie im Beweis benötigt, bereits fertig.

Enthält φ negierte atomare Formeln wie im Beispiel, lässt es sich durch Distribuieren in folgende Form (hier am Beispiel) bringen:

$$\begin{aligned}
\exists y\ \varphi &\equiv \exists y\ (p_1 = p_2 \wedge p_2 < p_3 \wedge \neg p_3 < p_4 \wedge \ldots) \\
&\equiv \exists y\ (p_1 = p_2 \wedge p_2 < p_3 \wedge (p_3 = p_4 \vee p_4 < p_3) \wedge \ldots) \\
&\equiv \exists y\ ((p_1 = p_2 \wedge p_2 < p_3 \wedge p_3 = p_4 \wedge \ldots) \vee (p_1 = p_2 \wedge p_2 < p_3 \wedge p_4 < p_3 \wedge \ldots)) \\
&\equiv (\exists y\ (p_1 = p_2 \wedge p_2 < p_3 \wedge p_3 = p_4 \wedge \ldots)) \vee (\exists y\ (p_1 = p_2 \wedge p_2 < p_3 \wedge p_4 < p_3 \wedge \ldots))
\end{aligned}$$

Durch ggf. wiederholte Anwendung lässt sich $\exists y\ \varphi$ so in eine Menge von Formeln $\exists y\ \psi_i$ transformieren, wobei ψ_i jeweils eine Konjunktion von Formeln der Form $\operatorname{sgn} p(\vec{x}, y) = \varepsilon_p$ ist. Auf die ψ_i lässt sich dann der Beweis entsprechend anwenden.

2.4 Algorithmus zur Berechnung eines Vorzeichendiagramms

Wir betrachten wieder eine Menge L' von Polynomen in $\mathbb{R}[X][Y]$ und ihre Muchnik-Hülle L (in Bezug auf Y). Ziel ist es, einen Algorithmus anzugeben, der, gegeben $\mathrm{diag}(L_0(\vec{x}))$ für ein $\vec{x} \in \mathbb{R}^N$, $\mathrm{diag}(L(\vec{x}))$ berechnet.

Nach Lemma 2.10 ist jedes Anfangssegment M von L wiederum eine Muchnik-Liste. Wir nutzen dies, um induktiv ein Vorzeichendiagramm für jedes $M(\vec{x})$ zu erstellen, im Allgemeinen durch Hinzufügen von einer weiteren Zeile und mehreren Spalten (aufgrund hinzukommender Nullstellen) zu einem bestehenden Vorzeichendiagramm. Daher reicht es aus, das folgende Lemma zu beweisen.

Lemma 2.13. *Sei $A = \mathbb{R}[X]$ mit $X = (X_1, \ldots, X_N)$. Sei L eine endliche Muchnik-Liste in $A[Y]$ der Länge m und p ein nicht-konstantes Polynom in $A[Y]$. Sei $L^+ = L \cup \{p\}$ wiederum eine Muchnik-Liste. Es existiert ein Algorithmus $C \mapsto C^+$ der jeder Matrix $C \in \{-1, 0, 1\}^{m \times n}$ eine Matrix $C^+ \in \{-1, 0, 1\}^{(m+1) \times n'}$ zuordnet (wobei $n' \leq n$), so dass für alle $\vec{x} \in \mathbb{R}^N$ gilt:*

$$C = \mathrm{diag}(L(\vec{x})) \implies C^+ = \mathrm{diag}(L^+(\vec{x}))$$

Beweis. Wir konstruieren nun den beschriebenen Algorithmus, der zwei Aufgaben zu erfüllen hat. Zum einen soll, basierend auf der Matrix C eine Matrix C^+ generiert werden, die eine zusätzliche Zeile (beschriftet mit p) sowie möglicherweise zusätzliche Spalten enthält. Zum anderen muss verifiziert werden, dass, wenn $\vec{x} \in \mathbb{R}^N$ und $C = \mathrm{diag}(L(\vec{x}))$, die generierte Matrix C^+ ein Vorzeichendiagramm von $L^+(\vec{x})$ ist (anschließend kann C^+ ggf. auf das reduzierte Vorzeichendiagramm gekürzt werden). Beim Durchführen dieser Schritte kann es sein, dass Inkonsistenzen in C zutage treten, die implizieren, dass $C \neq \mathrm{diag}(L(\vec{x}))$ für ein $\vec{x}$ war; die Matrix C wird dann abgewiesen.

Angenommen p hat Grad $d \geq 1$ in Y und $a \neq 0$ ist sein höchster Koeffizient ($a \in R[X]$). Dann sind $d!a$, p° und p' per Definition der Muchnik-Liste in L enthalten. Wir fixieren $\vec{x} \in \mathbb{R}^N$ und ξ sei ein Spaltenbezeichner von C. Wir definieren $C^+(p, \xi)$ abhängig von den nachfolgenden, unterschiedlichen Fällen.

Fall 1. Die Zeile $d!a$ in C ist nicht konstant. Dies ist unmöglich, wenn C von der Form $\mathrm{diag}(L(\vec{x}))$ ist, also weisen wir diese Matrix ab.

Fall 2. Die Zeile $d!a$ in C ist eine Null-Zeile. Wenn C von der Form $\mathrm{diag}(L(\vec{x}))$ ist, dann muss $a(\vec{x}) = 0$ sein und $p(\vec{x}, Y)$ und $p^\circ(\vec{x}, Y)$ müssen die gleiche Vorzeichenentwicklung haben. Also kopieren wir in diesem Fall einfach die Zeile p° als letzte Zeile und sind fertig.

Im Folgenden können wir also davon ausgehen, dass die Zeile $d!a$ in C konstanten Wert $\alpha \neq 0$ hat. Wenn C nur eine einzelne Spalte hat, verdoppeln wir diese an diesem Punkt.

Fall 3. Die Spalte ξ ist entweder die erste oder die letzte Spalte. Wenn C von der Form $\mathrm{diag}(L(\vec{x}))$ ist, dann muss $C^+(p, Y)$ das Vorzeichen von $p(\vec{x}, Y)$ an $-\infty$ bzw. $+\infty$ sein. Wir setzen es daher auf $(-1)^d \alpha$ bzw. α und sind fertig.

Im Folgenden können wir also zusätzlich annehmen, dass ξ eine interne Spalte ist.

Fall 4. Es gibt ein $q \in L$ mit $C(q, \xi) = 0$, aber die Zeile q ist keine Null-Zeile. Wenn C von der Form $\mathrm{diag}(L(\vec{x}))$ ist, dann bedeutet das, dass $q(\vec{x}, \xi) = 0$, aber $q(\vec{x}, Y)$ ist nicht konstant 0. Wir wählen ein $q \in L$ mit minimalem Grad, das diese Eigenschaften hat. Sei e sein Grad und b sein höchster Koeffizient, so dass $e!b \in L$. Wenn die mit $e!b$ beschriftete Zeile in C nicht konstant ist, weisen wir die Matrix ab.

Im Folgenden können wir also zusätzlich annehmen, dass die Zeile $e!b$ konstanten Wert β hat.

Fall 5. Angenommen $\beta = 0$. Wenn C von der Form $\mathrm{diag}(L(\vec{x}))$ ist, dann muss $b(\vec{x}) = 0$ sein. Da $q = bY^e + q^\circ$ (und $q(\vec{x}, \xi) = 0$, s. o.) muss $q^\circ(\vec{x}, \xi)$ ebenfalls 0 sein. Da $q^\circ \in L$ geringeren Grad als e hat, impliziert die Auswahl von q als Polynom mit minimalem Grad dieser Eigenschaft in Fall 4, dass die mit q° beschriftete Zeile in C eine Null-Zeile sein muss, d. h. $q^\circ(\vec{x}, Y) = 0$. Dann ist aber auch $q(\vec{x}, Y) = 0$, was einen Widerspruch darstellt. Also weisen wir die Matrix ab.

Im Folgenden können wir also zusätzlich $\beta \neq 0$ annehmen.

Sei r der Pseudorest von $\text{rem}(p,q)$, so dass $b^{d-e+1}p = hq + r$ für Polynome $h, r \in \mathbb{R}[X,Y]$ mit r vom Grad höchstens $e-1$. Wenn C von der Form $\text{diag}(L(\vec{x}))$ ist, dann gilt $b(\vec{x})^{d-e+1}p(\vec{x},\xi) = r(\vec{x},\xi)$ (da $q(\vec{x},\xi) = 0$). Wir setzen $C^+(p,\xi) := \beta^{d-e+1} \cdot C(r,\xi)$ und sind fertig.

Im Folgenden können wir also davon ausgehen, dass die einzigen Nullen in der Spalte ξ von C von Null-Zeilen stammen.

Seien ξ_- und ξ_+ die Bezeichner der Spalten direkt links und rechts der mit ξ beschrifteten Spalten. Angenommen, für ξ_- trifft auch der gerade behandelte Fall zu, so dass $C^+(p,\xi_-)$ zu diesem Zeitpunkt noch nicht durch vorangegangene Fälle definiert ist. Da aber in einem reduzierten Vorzeichendiagramm keine benachbarten Spalten existieren können, deren Bezeichner nicht Nullstellen eines eine Zeile bezeichnenden Polynoms sind, können wir eine solche Matrix abweisen. Für ξ_+ gilt dies analog.

Im Folgenden können wir also davon ausgehen, dass $C^+(p,\xi_-) := \varepsilon_-$ und $C^+(p,\xi_+) := \varepsilon_+$ bereits definiert sind.

Fall 6. Angenommen $\varepsilon_- = \varepsilon_+ = 0$. Wenn C von der Form $\text{diag}(L(\vec{x}))$ ist, muss $p(\vec{x},\xi_-) = p(\vec{x},\xi_+) = 0$ gelten. Nach Satz von Rolle gibt es ein $\eta \in (\xi_-,\xi_+)$ mit $p'(\vec{x},\eta) = 0$. Da per Definition $p' \in L$ müssen seine Nullstellen als Spaltenbezeichner auftreten, also gilt $\eta = \xi$. Nach Annahme über ξ muss die Zeile p' in C eine Null-Zeile sein, also muss auch $p'(\vec{x},Y) = 0$ sein. Das bedeutet wiederum, dass $p(\vec{x},Y)$ konstant, also vom Grad 0 ist, was ein Widerspruch zur initialen Annahme über p ist. Daher weisen wir die Matrix in diesem Fall zurück.

Fall 7. Angenommen ε_- und ε_+ haben entgegengesetzte Vorzeichen und sind nicht 0. Wenn C von der Form $\text{diag}(L(\vec{x}))$ ist, dann haben $p(\vec{x},\xi_-)$ und $p(\vec{x},\xi_+)$ entgegengesetzte Vorzeichen. Nach dem Zwischenwertsatz gibt es ein $\eta \in (\xi_-,\xi_+)$ mit $p(\vec{x},\eta) = 0$. Da wir nicht davon ausgehen können, dass $\xi = \eta$, haben wir eine neue Nullstelle. Wir ersetzen die Spalte ξ durch drei Kopien von ihr und setzen in der untersten Zeile von C^+ die Werte $(\varepsilon_-, 0, \varepsilon_+)$ um das Vorzeichenverhalten widerzuspiegeln. Damit sind wir in diesem Fall fertig.

Im Folgenden können wir also davon ausgehen, dass entweder $\varepsilon_- = 0$ oder $\varepsilon_+ = 0$, oder dass beide das gleiche Vorzeichen haben.

Wir behaupten, dass wenn C von der Form $\text{diag}(L(\vec{x}))$ ist, $p(\vec{x},Y)$ keine Nullstelle im offenen Intervall (ξ_-,ξ_+) hat. Ausgehend von dieser Behauptung (die wir gleich beweisen), muss das Vorzeichen auf dem gesamten Intervall gleich dem Vorzeichen des Endpunkts sein, der keine Nullstelle ist. Wir setzen also $C^+(p,\xi)$ auf denjenigen Wert ε_- oder ε_+, der nicht 0 ist.

Abschließend beweisen wir die Behauptung. Angenommen, es gäbe ein $\eta \in (\xi_-,\xi_+)$ mit $p(\vec{x},\eta) = 0$. Wenn einer der beiden Endpunkte eine Nullstelle ist, dann gäbe es nach Satz von Rolle eine Nullstelle von p' in diesem Intervall, was wir aber bereits ausgeschlossen haben. Also muss $p(\vec{x},Y)$ das gleiche Vorzeichen an beiden Endpunkten des Intervalls haben. O.b.d.A. seien beiden Endpunkte positiv. Damit $p(\vec{x},Y)$ auf dem Intervall 0 werden kann müsste es zunächst abfallen und dann wieder ansteigen, also müsste p' das Vorzeichen auf dem Intervall wechseln und demnach eine Nullstelle haben, was ebenfalls ein Widerspruch ist. □

Beispiel 2.14. Wir betrachten für $A = \mathbb{R}[X]$ (mit einzelner Variable X) die Liste $L = \{0, 2X, 4X, 4X^3, -4X, -4X^2, (Y+2)X, (Y-2)X, 2XY^2\}$ von Polynomen in $A[Y]$ sowie das einzelne Polynom $p = (Y^2-2)X$. Der Einfachheit und Verständlichkeit halber behandeln wir L und $L^+ = L \cup \{p\}$ im Folgenden, als seien es Muchnik-Listen[5]. Ausgehend von L wollen wir nun beispielhaft das Einfügen des Polynoms p in ein gegebenes Vorzeichendiagramm C, also die Konstruktion von C^+ durchführen.

Sei C als Eingabe des Algorithmus wie in Abbildung 2.14 gegeben. Zum Hinzufügen von $p = (Y^2-2)X = XY^2 - 2X$ müssen wir nun sämtliche Zellen $C^+(p,\xi_i)$ berechnen (entsprechend der Algorithmusbeschreibung ist im Folgenden $d = 2$, $a = X$:

- ξ_1 und ξ_7 sind die erste und letzte Spalte und werden mit Fall 3 abgehandelt:
  - $C^+(p,\xi_1) := (-1)^d\alpha = 1$

[5] In Wirklichkeit sind es keine Muchniklisten sondern Teilmengen der Muchnik-Hülle von L^+. Die vollständigen Listen würden den Rahmen des Beispiels sprengen und das Beispiel erlaubt auch mit dieser reduzierten Liste einen gleichwertigen Einblick in den Algorithmus.

	ξ_1	$\xi_2 = -2$	ξ_3	$\xi_4 = 0$	ξ_5	$\xi_6 = 2$	ξ_7
0	0	0	0	0	0	0	0
$2X$	1	1	1	1	1	1	1
$4X$	1	1	1	1	1	1	1
$4X^3$	1	1	1	1	1	1	1
$-4X$	-1	-1	-1	-1	-1	-1	-1
$-4X^2$	-1	-1	-1	-1	-1	-1	-1
$(Y+2)X$	-1	0	1	1	1	1	1
$(Y-2)X$	-1	-1	-1	-1	-1	0	1
$2XY^2$	1	1	1	0	1	1	1

Abbildung 1: Beispielhafte Eingabe C des Algorithmus.

  - $C^+(p, \xi_7) := \alpha = 1$

- Für ξ_4 kommt Fall 5 zur Anwendung, da es eine 0 in dieser Spalte gibt, die nicht zu einer Nullzeile gehört. Entsprechend Fall 4 ist $q = 2XY^2$ (das Polynom, für das die 0 in der Spalte steht), $e = 2$, $b = 2X$, $e!b = 4X$, $\beta = 1$. Wir bilden den Pseudorest[6] $\mathrm{rem}(p, q) = -4X^2$ und setzen
  - $C^+(p, \xi_4) := \beta^1 \cdot C(r, \xi_4) = 1 \cdot (-1) = -1$
- Bei ξ_2 und ξ_6 wird analog zu ξ_4 nach Fall 5 verfahren und es ergibt sich mit $\mathrm{rem}(XY^2 - 2X, XY \pm 2X) = 4X^3$:
  - $C^+(p, \xi_2) := 1$
  - $C^+(p, \xi_6) := 1$

  Damit sieht die Zeile p von C^+ bisher so aus:

	ξ_1	$\xi_2 = -2$	ξ_3	$\xi_4 = 0$	ξ_5	$\xi_6 = 2$	ξ_7
$(Y^2-2)X$	1	1		-1		1	1

- Wir betrachten nun ξ_3, wo augenscheinlich die Nullstelle $Y = -\sqrt{2}$ von p untergebracht werden muss. Da Nullen in ξ_3 nur in Nullzeilen auftreten kommen wir nun zu Fall 7, wobei $\xi_- = \xi_2$, $\xi_+ = \xi_4$, $\varepsilon_- = 1$ und $\varepsilon_+ = -1$ gilt. Da ϵ_- und ε_+ entgegengesetztes Vorzeichen haben gibt es nach dem Zwischenwertsatz eine Nullstelle $\eta \in (\xi_-, \xi_+)$ (nämlich $-\sqrt{2}$). Wir ersetzen also die Spalte ξ_3 durch drei Kopien und für die Zeile p die Werte $(1, 0, -1)$ ein.
- Bei ξ_5 wird analog zu ξ_3 verfahren, auch hier wird die Spalte durch drei Kopien ersetzt die in der Spalte p die Werte $(-1, 0, 1)$ erhalten.

Damit ergibt sich als Ausgabe C^+ des Algorithmus das in Abbildung 2.14 dargestellte Diagramm.

An dieser Stelle sei noch angemerkt, dass das betrachtete C nur für positive X ein gültiges Vorzeichendiagramm war, d. h. im Beweis von Theorem 2.11 die zu C gehörende Formel ψ_i nur für positive X wahr wird.

2.4.1 Anmerkungen zum Algorithmus und seiner Komplexität

Der Beweis gilt nicht nur, wie hier durchgeführt, über den reellen Zahlen, sondern auch über reell abgeschlossenen Körpern, da die einzigen nicht-trivialen Eigenschaften die Verwendung des Satzes von Rolle sowie des Zwischenwertsatzes sind, die auch über reell abgeschlossenen Körpern gelten.

Der Algorithmus hat ungefähr folgende Komplexität: Für eine Formel $\Phi(x)$ mit einer existenziell quantifizierten Variable, die aus l Polynomungleichungen jeweils vom Grad d besteht (d. h. die Kardinalität der Menge von Polynomen L' ist l) ergibt sich eine quantorenfreie Formel $\Psi(x)$ mit $O(l^{2^{d-1}})$ Ungleichungen. Präziser ist die Kardinalität

[6] $2X(Y^2 - 2X) = 2XY^2 \cdot \underbrace{X}_{h} \underbrace{-4X^2}_{r}$

	ξ_1	$\xi_2 = -2$	(neu)	ξ_3	(neu)	$\xi_4 = 0$	(neu)	ξ_5	(neu)	$\xi_6 = 2$	ξ_7
0	0	0	0	0	0	0	0	0	0	0	0
$2X$	1	1	1	1	1	1	1	1	1	1	1
$4X$	1	1	1	1	1	1	1	1	1	1	1
$4X^3$	1	1	1	1	1	1	1	1	1	1	1
$-4X$	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
$-4X^2$	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
$(Y+2)X$	-1	0	1	1	1	1	1	1	1	1	1
$(Y-2)X$	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	1
$2XY^2$	1	1	1	1	1	0	1	1	1	1	1
$(Y^2-2)X$	1	1	1	0	-1	-1	-1	0	1	1	1

Abbildung 2: Beispielhafte Ausgabe C^+ des Algorithmus.

von L (die Hülle von L') und L_0 in $O(m)$ wobei $m = l^{2^{d-1}}$; daraus resultieren 3^m mögliche Vorzeichenzuweisungen, von denen manche zu einer kompatiblen Formel führen. Die Reduktion von Φ nach Ψ ist also in der Größenordnung eines Polynoms in l mal 3^m.

Anmerkung des Autors: Bei der Vorbereitung des Seminarvortrags zu dieser Ausarbeitung ist mir aufgefallen, dass die Beschreibung des Algorithmus in [8] unvollständig oder fehlerhaft ist, da je nach Eingabe der Algorithmus kein gültiges Ergebnis liefert. Als konkrete Beispiele seien hier kurz folgende genannt:

a) $L = \{0, 1, 2, 3, 4, -4\}$, $p = Y + 1$, C die $\{-1, 0, 1\}^{m\times 1}$-Matrix für L. Der Algorithmus verdoppelt die Spalte in Fall 2 und fügt dann zweimal mit Fall 3 die Werte -1 und 1 für p ein. Die Nullstelle von p wird nicht berücksichtigt.

b) $L = \{0, 1, 2, 3, 4, -4, Y + 1\}$, $p = 4Y$, C wie folgt:

	ξ_1	$\xi_2 = -1$	ξ_3
0	0	0	0
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
-4	-1	-1	-1
$Y+1$	-1	0	1

Die Spalten ξ_1 und ξ_3 werden mit Fall 3 abgehandelt, die Spalte ξ_2 mit Fall 5 auf -1 gesetzt; hier müsste aber eigentlich eine weitere Nullstelle eingefügt werden.

3 Reelle Quantorenelimination ist im Worst Case mindestens doppelt exponentiell [4]

Wir skizzieren im Folgenden ein Beispiel das zeigt, dass Quantorenelimination über reell abgeschlossenen Körpern doppelt exponentiellen Platz (und damit auch Laufzeit) erfordern kann. Hierfür konstruieren wir eine Sequenz von Formeln, deren Länge linear in der Anzahl von Quantoren ist und deren Länge bei Quantorenelimination mindestens doppelt exponentiell wird. Konkret handelt es sich um eine Sequenz von Ausdrücken der Form $x_1^{2^{2^{k+1}}} = x_2$, deren quantorenfreie Äquivalente mindestens 2^{2^k} Symbole benötigen.

3.1 Konstruktion der Formel

Wir betrachten eine Sequenz von Formeln $\phi_0, \phi_1, \ldots, \phi_k$ mit jeweils vier freien Variablen, nämlich $x_{1R}, x_{1I}, x_{2R}, x_{2I}$ bei geradem und $z_{1R}, z_{1I}, z_{2R}, z_{2I}$ bei ungeradem Formelindex. Die Grundformel ϕ_0 lautet

$$\left(x_{1R}^4 - 6x_{1R}^2x_{1I}^2 + x_{1I}^4 = x_{2R}\right) \wedge \left(4x_{1R}^3x_{1I} - 4x_{1R}x_{1I}^3 = x_{2I}\right)$$

und drückt die reellen und imaginären Teile der Formel

$$(x_{1R} + ix_{1I})^4 = x_{2R} + ix_{2I}$$

aus, welche wiederum als komplexe Gleichung

$$x_1^4 = x_2 \tag{1}$$

betrachtet werden kann.

Ausgehend einer Formel ϕ_j wird nun die Formel ϕ_{j+1} abhängig von der Parität des Index j berechnet. Aus der folgenden Berechnungsregel für ungerades j erhält man die Regel für gerades j durch Vertauschen der Rollen von x und z.

$$\begin{aligned}\phi_{j+1}(z_{1R}, z_{1I}, z_{2R}, z_{2I}) = \exists y_R \exists y_I \forall x_{1R} \forall x_{1I} \forall x_{2R} \forall x_{2I} \big(((x_{1R} = z_{1R} \wedge x_{1I} = z_{1I} \wedge x_{2R} = y_R \wedge x_{2I} = y_I) \\ \vee (x_{1R} = y_R \wedge x_{1I} = y_I \wedge x_{2R} = z_{2R} \wedge x_{2I} = z_{2I})) \implies \phi_j(x_{1R}, x_{1I}, x_{2R}, x_{2I})\big)\end{aligned}$$

In ihrer komplexen Form reduziert sich diese Gleichung zu

$$\phi_{j+1}(z_1, z_2) = \exists y \forall x_1 \forall x_2 \big(((x_1 = z_1 \wedge x_2 = y) \vee (x_1 = y \wedge x_2 = z_2)) \implies \phi_j(x_1, x_2)\big).$$

Diese Formel ist logisch äquivalent zu

$$\exists y \forall x_1 \forall x_2 \big(((x_1 = z_1 \wedge x_2 = y) \implies \phi_j(x_1, x_2)) \wedge ((x_1 = y \wedge x_2 = z_2) \implies \phi_j(x_1, x_2))\big)$$

Die Implikationen sind trivialerweise wahr, außer x_1 und x_2 haben die in der Hypothese angegebenen Werte, also ist $\phi_{j+1}(z_1, z_2)$ logisch äquivalent zu

$$\exists y(\phi_j(z_1, y) \wedge \phi_j(y, z_2)). \tag{2}$$

Die durchgeführten Transformationen können natürlich auch mit der reellen Version der Formel gemacht werden.

Satz 3.1. *Die komplexe Version von ϕ_j ist äquivalent zur Gleichung*

$$x_1^{2^{2^{j+1}}} = x_2$$

für j gerade bzw. der gleichen Gleichung mit z_1 und z_2 für j ungerade.

Beweis. Der Satz lässt sich durch einfache Induktion beweisen.

Für $j = 0$ erhalten wir

$$x_1^{2^{2^{j+1}}} = x_1^{2^{2^1}} = x_1^4 = x_2$$

was gerade der Grundformel ϕ_0 entspricht.

Für $j + 1$ können wir als Hypothese annehmen, dass ϕ_j äquivalent zur Gleichung $x_1^{2^{2^{j+1}}} = x_2$ ist (wir betrachten nur den Fall, dass j gerade ist; der Fall j ungerade erfolgt analog). Mit (2) und der Hypothese gilt

$$\begin{aligned}\phi_{j+1} &\equiv \exists y(\phi_j(z_1, y) \wedge \phi_j(y, z_2)) \\ &\Leftrightarrow \exists y(z_1^{2^{2^{j+1}}} = y \wedge y^{2^{2^{j+1}}} = z_2) \\ &\Leftrightarrow \left(z_1^{2^{2^{j+1}}}\right)^{2^{2^{j+1}}} = z_2 \\ &\Leftrightarrow z_1^{2^{2^{j+1}} \cdot 2^{2^{j+1}}} = z_2 \\ &\Leftrightarrow z_1^{2^{2^{j+1} + 2^{j+1}}} = z_2 \\ &\Leftrightarrow z_1^{2^{2^{j+2}}} = z_2\end{aligned}$$

□

Korollar 3.2. *ϕ_j ist äquivalent zu folgender logischen Formel (wobei $\mathcal{R}$ und $\mathcal{I}$ für den Real- bzw. Imaginärteil stehen)*

$$\mathcal{R}\left((x_{1R}+ix_{1I})^{2^{2^{j+1}}}\right)=x_{2R}\wedge\mathcal{I}\left((x_{1R}+ix_{1I})^{2^{2^{j+1}}}\right)=x_{2I}$$

für j gerade bzw. der gleichen Formel mit z_{1R}, z_{1I}, z_{2R} und z_{2I} für j ungerade.

Satz 3.3. *$\phi_k(x_{1R},x_{1I},1,0)$ (bzw. $\phi_k(z_{1R},z_{1I},1,0)$ für k ungerade) definiert eine semialgebraische Menge in $\mathbb{R}^2$, die aus $2^{2^{k+1}}$ isolierten Punkten besteht.*

Jeder Punkt der durch ϕ_k definierten semialgebraischen Menge ist eine der $2^{2^{k+1}}$-ten Einheitswurzeln, also ein isolierter Punkt auf dem Einheitskreis.

Wir halten an dieser Stelle fest, dass das Alphabet, das benötigt wird um ϕ_k zu definieren, unabhängig von k ist und die Länge von ϕ_k damit linear in k.

3.2 Die Länge einer quantorenfreien Formel

Seien $p(x,y)$ und $q(x,y)$ zwei Polynome in $\mathbb{R}[x,y]$. Wir definieren $D(p)$ als die Menge der isolierten Punkte von $p=0$ und $D(p,q)$ als die Menge der isolierten Punkte der Schnittmenge der Kurven $p=0$ und $q=0$, die keine isolierten Punkte der Kurven sind, wenn man diese separat betrachtet. Man beachte, dass $D(p_1p_2)\subset D(p_1)\cup D(p_2)$ und dass $D(p_1p_2,q)\subset D(p_1,q)\cup D(p_2,q)$.

Nachfolgend geben wir die beiden Sätze wieder, die zum Beweis des abschließenden Theorems 3.6 benötigt werden. Für die Beweise dieser beiden Sätze siehe [4].

Satz 3.4. *Gegeben ein quantorenfreier, polynomieller Ausdruck in zwei reellen Variablen x und y, der Polynome $p_1(x,y),\ldots,p_n(x,y)$ beinhaltet, ist die Menge der isolierten Punkte der durch die Formel definierten Teilmenge von $\mathbb{R}^2$ eine Teilmenge von*

$$\left(\bigcup_{i=1}^{m}D(q_i)\right)\cup\left(\bigcup_{i=1}^{m}\bigcup_{j=i+1}^{m}D(q_i,q_j)\right)$$

wobei q_i die irreduziblen Faktoren von p_i sind.

Satz 3.5. *Wenn d_i der totale Grad des Polynoms p_i ist, dann ist die Gesamtanzahl der isolierten Punkte höchstens*

$$\left(\sum_{i=1}^{n}d_i\right)^2.$$

Unter der Anmerkung, dass – bei dichter Repräsentation von Polynomen – das Aufschreiben eines Polynoms in zwei Variablen vom totalen Grad n mindestens n Symbole erfordert, können wir die Sätze 3.4 und 3.5 kombinieren um folgendes Theorem abzuleiten:

Theorem 3.6. *Es existieren Formeln mit Länge linear in k, $6k$ Quantoren und zwei freien Variablen, so dass zum Aufschreiben ihrer reellen, quantorenfreien Darstellung mindestens 2^{2^k} Symbole nötig sind.*

Da jedes Symbol geschrieben werden muss, folgt sofort:

Korollar 3.7. *Die Elimination von n Quantoren über $\mathbb{R}$ ist im Worst Case mindestens doppelt exponentiell in n.*

4 Praktische Verfahren zur Quantorenelimination [7]

Das erste Verfahren zur reellen Quantorenelimination wurde von Tarski 1948 veröffentlicht [9], war aber sehr ineffizient. Präziser gesagt war der Algorithmus nicht mal primitiv-rekursiv; seine Laufzeit ist durch keinen Turm von Exponenten beschränkt.

Im Folgenden benennen wir abschließend beispielhaft drei praktische Verfahren zur Quantorenelimination auf Basis des Surveys von Dolzmann et al. [7]:

- **Partial Cylindrical Algebraic Decomposition** (partial CAD): Basierend auf der 1975 von Collins veröffentlichten CAD-Methode [2] veröffentlichten Collins et. al [3] 1991 nach diversen Verbesserungen partial CAD, das im Paket QEPCAD implementiert wurde. CAD (und partial CAD) haben doppelt-exponentielle Laufzeit in der Anzahl der Variablen und sind die effizientesten Algorithmen, die beliebige Quantoreneliminationsprobleme lösen (vgl. [1]). QEPCAD[7] wird aktiv weiterentwickelt [1].
- **Virtual Substitution**: Eine von Weispfenning 1988 [10] initiierte Methode, die Formeln fokussiert, bei denen quantifizierte Variablen nur mit geringem Grad auftreten. Ihre Worst-Case-Komplexität ist nur doppelt-exponentiell in der Anzahl der Quantorenblöcke der Eingabeformel, wodurch dieser Ansatz interessant für Probleme mit vielen Parametern ist. Virtual Substitution ist im REDLOG-Paket[8] implementiert [6].
- **Comprehensive Gröbner Bases and Multivariate Real Root Counting**: Ein ebenfalls von Weispfenning 1993 [11] vorgeschlagenes Verfahren, das auf dem Zählen reeller Nullstellen nulldimensionaler Ideale und der Berechnung von umfassenden Gröbnerbasen basiert. Es ist vollständig, aber insbesondere für Probleme mit sehr vielen Gleichungen geeignet. Der Algorithmus wurde im Paket QERRC implementiert [5].

Literatur

[1] C. W. Brown. Qepcad b: A program for computing with semi-algebraic sets using cads. *ACM SIGSAM Bulletin*, 37(4):97–108, 2003.

[2] G. E. Collins. Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. In *Automata Theory and Formal Languages. 2nd GI Conference*, pages 134–183, 1975.

[3] G. E. Collins and H. Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *Journal of Symbolic Computation*, 12(3):299–328, 1991.

[4] J. H. Davenport and J. Heintz. Real quantifier elimination is doubly exponential. *Journal of Symbolic Computation*, 5(1/2):29–35, 1988.

[5] A. Dolzmann. Reelle Quantorenelimination durch parametrisches Zählen von Nullstellen. Diploma thesis, Universität Passau, 1994.

[6] A. Dolzmann and T. Sturm. Redlog: Computer algebra meets computer logic. *ACM SIGSAM Bulletin*, 31(2):2–9, 1997.

[7] A. Dolzmann, T. Sturm, and V. Weispfenning. Real quantifier elimination in practice. In *Algorithmic Algebra and Number Theory*, pages 221–247. Springer, Berlin, 1998.

[8] H. Schoutens. Muchnik's proof of Tarski-Seidenberg. `http://websupport1.citytech.cuny.edu/faculty/hschoutens/PDF/Muchnik.pdf`, 2001.

[9] A. Tarski. A decision method for elementary algebra and geometry. Technical report, RAND, Santa Monica, CA, 1948.

[10] V. Weispfenning. The complexity of linear problems in fields. *Journal of Symbolic Computation*, 5(1-2):3–27, 1988.

[11] V. Weispfenning. A new approach to quantifier elimination for real algebra. Technical Report MIP-9305, FMI, Universität Passau, 1993.

[7] http://www.usna.edu/Users/cs/qepcad/B/QEPCAD.html

[8] http://redlog.dolzmann.de

[12] M. Ziegler. Skript zur Vorlesung über Modelltheorie. http://home.mathematik.uni-freiburg.de/ziegler/skripte/modell1.pdf, 1997. Uni Freiburg, Wintersemester 1997/1998.