

Wortprobleme in Ringen (II)

Moritz Hartlieb

29. Mai 2020

Wir arbeiten im Folgenden in einer Logik erster Stufe ohne Gleichheit und mit der Sprache der Ringe $S_{\text{Ring}} = \{0, 1, +, -, \cdot, =\}$.

Definition 1. Sei $\Phi_{\text{Ring}} \subset L^{S_{\text{Ring}}}$ die Menge der folgenden (implizit universellquantifizierten) Axiome:

	$x + y = y + x$
	$x + (y + z) = (x + y) + z$
	$x + 0 = x$
	$x + (-x) = 0$
	$x \cdot y = y \cdot x$
	$x \cdot (y \cdot z) = (x \cdot y) \cdot z$
	$x \cdot 1 = x$
	$x \cdot (y + z) = (x \cdot y) + (x \cdot z)$
	$x = x$
$x = y \Rightarrow$	$y = x$
$x = y \Rightarrow$	$-x = -y$
$x = y \wedge y = z \Rightarrow$	$x = z$
$x = x' \wedge y = y' \Rightarrow$	$x + y = x' + y'$
$x = x' \wedge y = y' \Rightarrow$	$x \cdot y = x' \cdot y'$

Anmerkung. Da wir in einer Logik erster Stufe ohne Gleichheit arbeiten, fassen wir neben den klassischen Ringaxiomen auch einige Äquivalenzaxiome in unsere Axiomenmenge auf. Dadurch erreichen wir die gewünschte Eigenschaft, dass eine Formel $\varphi \in L^{S_{\text{Ring}}}$ genau dann in jedem Ring gilt, wenn $\Phi_{\text{Ring}} \models \varphi$.

1 Motivation

Ziel dieses Abschnittes ist es, folgenden algebraischen Satz zu beweisen:

Satz. $\Phi_{\text{Ring}} \models \forall \bar{x}. p_1(\bar{x}) = 0 \wedge \cdots \wedge p_n(\bar{x}) = 0 \Rightarrow q(\bar{x}) = 0$ gdw. $q \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$.

Anstatt einen algebraischen Beweis zu geben, werden wir im folgenden Beweis auf Aussagen aus der mathematischen Logik zurückgreifen.

2 Definite Klauseln

Sei S eine Sprache mit mindestens einem Konstantensymbol. Im Folgenden betrachten wir Klauseln der konjunktiven Normalform, das heißt eine Klausel $\{L_i | i \in I\}$ mit Literalen L_i entspricht der Formel $\forall (\bigvee_{i \in I} L_i)$.

Definition 2. Eine *definite Klausel* ist eine Klausel mit genau einem positiven Literal.

Anmerkung. Eine definite Klausel der Form $\{\neg\varphi_1, \dots, \neg\varphi_n, \psi\}$ lässt sich logisch äquivalent auch in der Form

$$\varphi_1 \wedge \cdots \wedge \varphi_n \Rightarrow \psi$$

auffassen. Insbesondere sind die zu Φ_{Ring} korrespondierenden Klauseln definite Klauseln.

Wir werden nun Ideen aus dem Satz von Herbrand benutzen, um einige interessante Eigenschaften von Mengen definiter Klauseln zu zeigen.

Definition 3. Eine *Herbrand-Interpretation* ist ein S -Struktur M mit Trägermenge $\{t \in T^S | \text{var}(t) = \emptyset\}$, sodass $f^M(t_1, \dots, t_n) = f(t_1, \dots, t_n)$ für alle Funktionssymbole $f \in S$ und $t_1, \dots, t_n \in |M|$.

Definition 4. Die Menge

$$\{R(t_1, \dots, t_n) | R \in S \text{ } n\text{-stelliges Relationsymbol, } t_1, \dots, t_n \in T^S \text{ variabelnfrei}\},$$

der variabelnfreien positiven atomaren Aussagen einer Sprache bezeichnen wir als *Herbrand-Basis*.

Beispiel 5. Für die Sprache der Ringe ist die Herbrand-Basis die Menge aller Gleichungen ohne Variablen ($\{0 = 1, 1 = 1, 1 + 1 = 1, \dots\}$)

Eine Herbrand-Interpretation ist bestimmt durch die Interpretation der Relationssymbole:

Lemma 6. *Es gibt eine Bijektion:*

$$\{\text{Teilmengen der Herbrand-Basis}\} \longleftrightarrow \{\text{Herbrand-Interpretationen}\}$$

Beweis. Auf der einen Seite ist eine Herbrand-Interpretation M per Definition eindeutig bestimmt durch die Interpretation der Relationssymbole. Also erhalten wir für jede Herbrand-Interpretation eine eindeutige Teilmenge

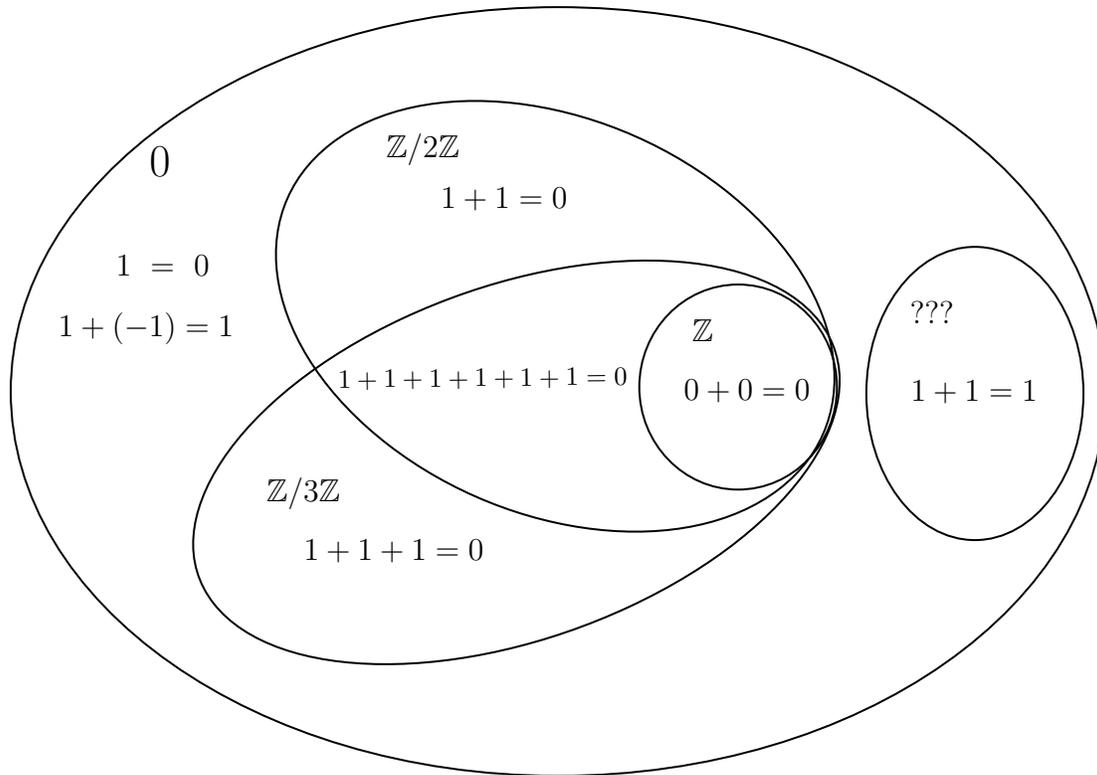
$$\{R(t_1, \dots, t_n) | M \models R(t_1, \dots, t_n), t_1, \dots, t_n \text{ variabelnfrei}\}$$

der Herbrand-Basis. Auf der anderen Seite lässt sich aus jeder Teilmenge U der Herbrand-Basis eine Herbrand-Interpretation M konstruieren, indem man

$$R^M := \{(t_1^M, \dots, t_n^M) | t_1, \dots, t_n \in |M|, R(t_1, \dots, t_n) \in U\}$$

für alle Relationssymbole $R \in S$ setzt. □

Beispiel 7. Anschaulich können wir uns die Herbrand-Basis für S_{Ring} wie folgt vorstellen:



Definition 8. Ein *Herbrand-Modell* einer Menge $\Psi \subset L_0^S$ ist eine Herbrand-Interpretation M , sodass $M \models \Psi$.

Lemma 9. Jede Menge $\Psi \subset L^S$ von definiten Klauseln besitzt ein Herbrand-Modell

Beweis. Die Herbrand-Interpretation, die jedes Element der Herbrand-Basis als Wahr interpretiert, ist ein Modell von Ψ , da aus $M \models \psi$ schon $M \models \chi \Rightarrow \psi$ für alle $\chi \in L^S$ folgt und alle Formeln in Ψ definit sind. \square

Insbesondere ist also jede Menge von definiten Klauseln erfüllbar.

Anmerkung. In der Sprache der Ringe ist das im obigen Beweis konstruierte Modell der Nullring.

Definition 10. Ein *minimales* Herbrand-Modell von $\Psi \subset L^S$ ist eine Herbrand-Modell M , sodass für jede atomare Formel ϕ in der Herbrand-Basis gilt: $M \models \phi$ gdw. $M' \models \phi$ für alle Herbrand-Modelle M' von Ψ .

Warnung. Eine in dieser Art konstruierte Herbrand-Interpretation muss kein Modell von Ψ sein: Sei S die Sprache bestehend aus einer Konstante 0 und zwei einstelligigen Relationssymbolen P, Q . Betrachte $\Psi = \{P(0) \vee Q(0)\} \subset L^S$. Dann gilt weder $P(0)$ noch $Q(0)$ in allen Herbrand-Modellen von Ψ und somit existiert kein minimales Herbrand-Modell.

Für definite Klauseln gilt dies jedoch:

Satz 11. Jede Menge Ψ von definiten Klauseln besitzt ein minimales Herbrand-Modell.

Beweis. Wir haben bereits gesehen, dass Ψ mindestens ein Herbrand-Modell besitzt. Es lässt sich nun prüfen, dass die in Frage kommende Herbrand-Interpretation auch wirklich ein Modell von Ψ ist: Sei M eine Herbrand-Interpretation, sodass $M \models \phi$ genau dann wenn $M' \models \phi$ für alle Herbrand-Modelle M' von Ψ und ϕ aus der Herbrand-Basis. Sei $c' = \phi_1 \wedge \dots \wedge \phi_n \Rightarrow \psi$ eine beliebige Grundinstanz einer Klausel c in Ψ . Falls ϕ_1, \dots, ϕ_n in allen Herbrand-Modellen von Ψ gelten, dann gilt schon ψ in allen Herbrand-Modellen von Ψ , also $M \models c'$. Aus der Definition der Herbrand-Interpretation und der Erfüllbarkeitsrelation folgt dann schon $M \models c$. \square

Aus der Existenz des minimalen Herbrand-Modells können wir nun die sogenannte "Konvexität" der definiten Klauseln folgern:

Korollar 12. Falls Ψ eine Menge von definiten Klauseln ist und A_1, \dots, A_n atomare Formeln sind, dann gilt $\Psi \models A_1 \vee \dots \vee A_n$ gdw. $\Psi \models A_i$ für ein $1 \leq i \leq n$.

Beweis. Die Rückrichtung ist klar. Für die Hinrichtung können wir annehmen, dass die A_i keine Variablen enthalten (Indem wir Variablen durch neue Konstantensymbole ersetzen). Ψ hat ein minimales Herbrand-Modell M und da $\Psi \models A_1 \vee \dots \vee A_n$ und die A_i keine Variablen enthalten, gilt $M \models A_k$ für ein $1 \leq k \leq n$. Dieses A_k gilt dann also in allen Herbrand-Modellen von Ψ .

Angenommen, es existiert ein S -Modell M' , sodass $M' \models \Psi$ und $M' \models \neg A_k$. Betrachte die Teilmenge

$$\{R(t_1, \dots, t_n) \mid M' \models R(t_1, \dots, t_n), t_1, \dots, t_n \in T^S \text{ variablenfrei}\}$$

der Herbrand-Basis. Diese definiert eine Herbrand-Interpretation H . Da die Formeln in Ψ universell sind und $\neg A_k$ keine Variablen enthält, folgt aus $M' \models \Psi$ und $M' \models \neg A_k$ schon $H \models \Psi$ und $H \models \neg A_k$. Das steht im Widerspruch zu $M \models A_k$.

Somit gilt $\Psi \models A_k$. □

Definition 13. Ein (*gerichteter*) Baum ist eine endliche Menge von Knoten $\emptyset \neq V$ zusammen mit einer Relation $R \subset V \times V$, sodass:

- (V, R) ist kreisfrei:

Seien $x, y \in V$. Falls $a_1, \dots, a_s \in V$ und $b_1, \dots, b_t \in V$ existieren, sodass

$$xRa_1, \dots, a_sRy \text{ und } xRb_1, b_1Rb_2, \dots, b_tRy,$$

dann gilt schon $s = t$ und $a_i = b_i$ für $1 \leq i \leq s$.

- (V, R) ist zusammenhängend:

Seien $x, y \in V$. Dann existiert $z \in V$ und $x_1, \dots, x_n, y_1, \dots, y_m \in V$, sodass

$$(zRx_1, \dots, x_nRx \text{ oder } z = x) \text{ und } (zRy_1, \dots, y_mRy \text{ oder } z = y).$$

Anmerkung. Die Relation R lässt sich als Menge von Pfeilen zwischen Knoten in V verstehen:

$$xRy \iff (x \rightarrow y)$$

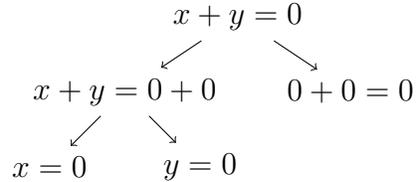
Definition 14. Sei (V, R) ein Baum und $x, y \in V$ Knoten, sodass xRy . Dann ist x der *Vorgänger* von y und y ein *Nachfolger* von x . Ein *Blatt* ist ein Knoten, der keine Nachfolger hat. Eine *Wurzel* ist ein Knoten ohne Vorgänger.

Lemma 15. Jeder Baum hat genau eine Wurzel.

Beweis. Angenommen, $a, b \in V$ sind Wurzeln von V . Dann gibt es $c \in V$ und $x_1, \dots, x_n, y_1, \dots, y_m \in V$, sodass $(cRx_1, \dots, x_nRa \text{ oder } c = a)$ und $(cRy_1, \dots, y_mRb \text{ oder } c = b)$. Da a und b Wurzeln sind, gilt dann aber schon $a = c = b$. □

Definition 16. Sei Ψ eine Menge von definiten Klauseln. Dann bezeichne $B_S(\Psi)$ die Menge aller Bäume deren Knoten variabelnfreie atomare Formeln sind, sodass für jeden Knoten P mit Nachfolgern Q_1, \dots, Q_n , die Formel $Q_1 \wedge \dots \wedge Q_n \Rightarrow P$ Grundinstanz einer Formel in Ψ ist.

Beispiel 17. Für die Sprache $S = S_{\text{Ring}} \cup \{x, y\}$ und die Menge $\Psi = \Phi_{\text{Ring}} \cup \{x = 0, y = 0\}$ ist folgender Baum ein Element von $B_S(\Psi)$:



Satz 18. Sei Ψ eine Menge von definiten Klauseln. Die Teilmenge X der Herbrand-Basis, der Formeln, die als Wurzel eines Baumes in $B_S(\Psi)$ auftreten, korrespondiert unter der obigen Bijektion mit dem minimalen Herbrand-Modell von Ψ .

Beweis. Sei M das Modell, das mit X korrespondiert. Dann gilt $M \models P_1 \wedge \dots \wedge P_n \Rightarrow Q$ für alle Grundinstanzen von Formeln in Ψ : Wenn P_i für alle $1 \leq i \leq n$ die Wurzel eines Baumes in $B_S(\Psi)$ ist, dann können wir einen Baum T konstruieren, dessen Wurzel Q ist und die Nachfolger P_1, \dots, P_n von Q Wurzeln der obigen Bäume als Unterbäume sind.

Sei M' ein beliebiges Herbrand-Modell der Grundinstanzen von Ψ . Dann enthält die Teilmenge der Herbrand-Basis, die mit M' korrespondiert, schon X : Wenn für eine Grundinstanz $P_1 \wedge \dots \wedge P_n \Rightarrow Q$ jedes der P_i in M' erfüllt sind, dann gilt auch $M' \models Q$.

Aus den Definitionen des All-Quantors und der Herbrand-Modelle folgt außerdem, dass die Herbrand-Modelle der Menge aller Grundinstanzen von Ψ genau die Herbrand-Modelle von Ψ sind.

Also ist M das minimale Herbrand-Modell von Ψ . □

Anmerkung. Die in diesem Abschnitt behandelten Aussagen lassen sich auch auf allgemeinere Klauseln, sog. Hornklauseln übertragen (Siehe Abschnitt 3.14 im Buch von Harrison).

3 Das Wortproblem für Ringe

Definition 19. Seien p_1, \dots, p_n Polynome in $R[x_1, \dots, x_k]$. Dann ist

$$\text{Id}_R\langle p_1, \dots, p_n \rangle := \left\{ \sum_{i=1}^n q_i p_i \mid n \in \mathbb{N}, q_i \in R[x_1, \dots, x_k] \right\}$$

das von p_1, \dots, p_n erzeugte Ideal.

Anmerkung. Für beliebige Terme p, p_1, \dots, p_n in der Sprache der Ringe erlauben über wir die Schreibweise $p \in \text{Id}_R\langle p_1, \dots, p_n \rangle$ für $\text{norm}(p) \in \text{Id}_R\langle \text{norm}(p_1), \dots, \text{norm}(p_n) \rangle$.

Aus der Algebra kennen wir folgendes Lemma:

Lemma 20. Ideale erfüllen folgende Abgeschlossenheitseigenschaften:

- $0 \in \text{Id}_R\langle p_1, \dots, p_n \rangle$.
- $p_i \in \text{Id}_R\langle p_1, \dots, p_n \rangle$ für $1 \leq i \leq n$.
- Falls $p \in \text{Id}_R\langle p_1, \dots, p_n \rangle$ und $q \in \text{Id}_R\langle p_1, \dots, p_n \rangle$ dann ist schon $p + q \in \text{Id}_R\langle p_1, \dots, p_n \rangle$
- Falls $p \in \text{Id}_R\langle p_1, \dots, p_n \rangle$ und $q \in R[x_1, \dots, x_k]$, dann ist schon $qp \in \text{Id}_R\langle p_1, \dots, p_n \rangle$.

Anmerkung. Wir schreiben \bar{x} für x_1, \dots, x_k .

Anmerkung. Aus der Algebra wissen wir, dass es für jeden Ring R einen eindeutigen Ringhomomorphismus $\mathbb{Z} \rightarrow R$ gibt. Dies lässt uns Elemente von \mathbb{Z} als Terme in der Sprache der Ringe auffassen.

Die Ideale ermöglichen uns nun, folgende Äquivalenz zu formulieren:

Satz 21. $\Phi_{\text{Ring}} \models \forall \bar{x}. p_1(\bar{x}) = 0 \wedge \dots \wedge p_n(\bar{x}) = 0 \Rightarrow q(\bar{x}) = 0$ gdw. $q \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$.

Beweis. Wir zeigen zunächst die Rückrichtung:

\Leftarrow Da $q \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$, existieren $c_i \in \mathbb{Z}[x_1, \dots, x_k]$, sodass

$$q(\bar{x}) = \sum_{i=1}^n c_i(\bar{x}) p_i(\bar{x}).$$

Sei R ein beliebiges Modell von Φ_{Ring} (also ein Ring). Falls für beliebige $a_1, \dots, a_n \in R$ schon $p_i(a_1, \dots, a_n) = 0$ für $1 \leq i \leq n$ gilt, dann gilt unter Verwendung der Ringaxiome auch

$$q(\bar{a}) = \sum_{i=1}^n c_i(\bar{a}) p_i(\bar{a}) = 0.$$

Da a, R beliebig gewählt wurden, folgt

$$\Phi_{\text{Ring}} \models \forall \bar{x}. p_1(\bar{x}) = 0 \wedge \dots \wedge p_n(\bar{x}) = 0 \Rightarrow q(\bar{x}) = 0$$

\Rightarrow Falls $\Phi_{\text{Ring}} \models \forall \bar{x}. p_1(\bar{x}) = 0 \wedge \dots \wedge p_n(\bar{x}) = 0 \Rightarrow q(\bar{x}) = 0$, dann gilt auch

$$\Phi := \Phi_{\text{Ring}} \cup \{p_1 = 0, \dots, p_n = 0\} \models q = 0,$$

wobei wir die Variablen \bar{x} durch neue Konstantensymbole ersetzt haben.

Sämtliche in Φ auftretenden Formeln entsprechen definiten Klauseln. Somit hat Φ ein minimales Herbrand-Modell. In diesem gilt $q = 0$. Also enthält $B(\Phi)$ einen Baum T mit Wurzel $q = 0$. Wir zeigen nun per Induktion über T , dass für jede Formel $s = t \in T$ schon $(s - t) \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$ gilt: (Siehe Definition 1)

- Für die auftretenden Blätter der Form $s = t$ gilt $s - t \approx 0 \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$.
- Für die restlichen Knoten müssen wir nur zeigen, dass die Eigenschaft unter Anwendung der Ring- und Äquivalenzaxiome erhalten bleibt. Dies folgt jedoch schon aus Lemma 19: Wenn ein Knoten $s = u$ beispielsweise die Nachfolger $s = t$ und $t = u$ hat (also eine Grundinstanz der Transitivität ist), dann folgt aus der Induktionshypothese $(s - t) \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$ und $(t - u) \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$ schon

$$(s - u) \approx ((s - t) + (t - u)) \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle.$$

Im Fall, dass der Knoten $x \cdot y = x' \cdot y'$ die Nachfolger $x = x'$ und $y = y'$ hat, dann ist nach Induktionshypothese $(x - x') \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$ und $(y - y') \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$. Es gilt nun aber schon

$$xy - x'y' \approx xy - xy' + xy' - x'y' \approx x(y - y') + y'(x - x') \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle.$$

Die restlichen Fälle lassen sich mit dem gleichen Prinzip zeigen.

Also gilt für die Wurzel $q = 0 \in T$ schon $q \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$. □

Im Spezialfall des freien Wortproblems sehen wir:

Korollar 22. $\Phi_{\text{Ring}} \models s = t$ gdw. $s \approx t$, s und t also das gleiche Polynom darstellen.

Beweis. Wende den obigen Satz mit $n = 0$ auf $q = s - t$ an. □

Satz 23. Die Frage der Gültigkeit einer universellen Formel in der Sprache der Ringe in allen Ringe kann auf das Wortproblem zurückgeführt werden.

Beweis. In der Logikvorlesung haben wir gesehen, dass jede quantorenfreie Formel äquivalent zu einer Formel in konjunktiver Normalform ist. Wir können also jede universelle Formel in der Sprache der Ringe ohne Beschränkung der Allgemeinheit als Formel der Form

$$\bigwedge_{k \in K} \left(\forall \bar{x}. \bigwedge_{i \in I} p_i^k(\bar{x}) = 0 \Rightarrow \bigvee_{j \in J} q_j^k(\bar{x}) = 0 \right). \quad (1)$$

auffassen.

Wir betrachten eine Formel der Form

$$\forall \bar{x}. \bigwedge_{i \in I} p_i(\bar{x}) = 0 \Rightarrow \bigvee_{j \in J} q_j(\bar{x}) = 0. \quad (2)$$

Falls kein q_j existiert, gilt die Formel (2) nicht in allen Ringen, da die Ringaxiome und $p_i(x) = 0$ definite Klauseln sind und somit mindestens ein Modell M existieren muss, sodass $M \models \Phi_{\text{Ring}}$ und $M \models p_i(x) = 0$ für alle $i \in I$.

Falls es mindestens ein q_j gibt, dann gilt

$$\Phi_{\text{Ring}} \models \forall \bar{x}. \bigwedge_{i \in I} p_i(\bar{x}) = 0 \Rightarrow \bigvee_{j \in J} q_j(\bar{x}) = 0.$$

genau dann, wenn

$$\Phi_{\text{Ring}} \cup \{p_1 = 0, \dots, p_m = 0\} \models \bigvee_{j \in J} q_j = 0, \quad (3)$$

wobei wir in (3) die Variablen \bar{x} durch neue Konstantensymbole ersetzen. Aus der Konvexität der definiten Klauseln folgt nun, dass ein $k \in J$ gibt, sodass (3) genau dann gilt, wenn

$$\Phi_{\text{Ring}} \cup \{p_1 = 0, \dots, p_m = 0\} \models q_k = 0. \quad \square$$

Wir können also die universelle Theorie der Ringe lösen, wenn wir das Wortproblem lösen können. Und das können wir, wenn wir die Idealzugehörigkeit entscheiden können.

4 Das Wortproblem für torsionsfreie Ringe

Definition 24. Ein Ring ist *torsionsfrei*, wenn er die folgende Axiomenmenge erfüllt:

$$\Phi_T = \{\forall x. nx = 0 \Rightarrow x = 0 \mid n \geq 1\}$$

Wir können eine ähnliche Äquivalenz wie oben für torsionsfreie Ringe zeigen:

Satz 25. $\Phi_{\text{Ring}} \cup \Phi_T \models \forall \bar{x}. p_1(\bar{x}) = 0 \wedge \dots \wedge p_n(\bar{x}) = 0 \Rightarrow q(\bar{x}) = 0$ gdw. es $0 \neq c \in \mathbb{Z}$ gibt, sodass $cq \in \text{Id}_{\mathbb{Z}}\langle p_1, \dots, p_n \rangle$.

Beweisidee. Der Beweis verläuft wie der obige Beweis für allgemeine Ringe per Induktion über den Beweisbaum, da auch die in Φ_T auftretenden Formeln definite Klauseln liefern.

Anmerkung. Jeder nicht-triviale torsionsfreie Ring hat Charakteristik 0. Die Umkehrung ist nicht immer wahr, gilt aber in Integritätsringen, die im nächsten Teil behandelt werden.