Russell, Gödel und der Lügner

Ontologische und semantische Antinomien und Gödels Unvollständigkeitssatz im Vergleich

Inauguraldissertation
zur Erlangung des Grades eines Doktors der Philosophie
im Fachbereich Philosophie und Geschichtswissenschaften
der Johann-Wolfgang-Goethe-Universität
zu Frankfurt am Main

vorgelegt von
Stefanie Ucsnay
aus Bonn

2008

Gutachter:
Prof. Dr. Wilhelm K. Essler
Prof. Dr. Ulrich Kohlenbach
Prof. Dr. Peter Gold
Tag der mündlichen Prüfung: 2. Juli 2008

Für die Unterstützung bei dieser Arbeit danke ich besonders
Professor Wilhelm K. Essler und Professor Ulrich Kohlenbach.

Inhaltsverzeichnis

Einleitung

“Dieser Satz ist falsch” – eine schnell aufgestellte Behauptung, über deren Status schon viel gegrübelt worden ist. Die Menge aller Mengen, die sich nicht selbst enthalten – enthält sie sich selbst?

Antinomien sind einfach zu schildern und üben auf viele Menschen einen besonderen Reiz aus. Sie fordern uns heraus, über unsere intuitiven Annahmen bezüglich Sprache, Semantik, Logik und Ontologie nachzudenken. Die Auseinandersetzung mit Antinomien hat der Philosophie und der mathematischen Grundlagenforschung immer wieder wichtige Impulse vermittelt.[1] Philosophen und Mathematiker haben vielfältige Lösungsansätze zur Antinomienproblematik entwickelt – ohne dass es Einigkeit über *die* Lösung gibt.

Ein Aspekt der Beschäftigung mit Antinomien besteht in der Untersuchung verschiedener Typen von Antinomien auf Gemeinsamkeiten. Auf welche Weise lassen sich intuitiv vermutete Ähnlichkeiten zwischen der Aussage “Dieser Satz ist falsch.” und der Bildung der Menge aller Mengen, die sich nicht selbst enthalten, präzisieren? Oder erweisen sie sich am Ende als trügerisch, weil sie einer genaueren Untersuchung nicht standhalten? Dieser Aspekt steht hier im Vordergrund. Und daran schließt sich auch die Frage nach der Beziehung zu einem bedeutenden Satz der Logik an, nämlich Gödels erstem Unvollständigkeitssatz, in dem ein mathematisches Pendant des Satzes “Dieser Satz ist nicht beweisbar.” gebildet und als unentscheidbar nachgewiesen wird.

Es gibt eine Vielzahl unterschiedlicher Arten von Paradoxien und Antinomien.[2] Die in dieser Arbeit behandelten Antinomien werden gewöhnlich in die

[1] Hier sind z. B. die Bildung verschiedener Mengenlehren, die Entstehung verschiedener Wahrheitstheorien, aber auch die Gödelschen Unvollständigkeitssätze zu nennen.

[2] Antinomien werden hier als bestimmte Arten von Paradoxien betrachtet. Während in Paradoxien allgemein aus intuitiven Annahmen widersprüchliche Ergebnisse resultieren, soll von Antinomien gesprochen werden, wenn die Ergebnisse innerhalb eines logischen Systems aus für den jeweiligen Anwendungsbereich gültigen Annahmen folgen und einen Widerspruch innerhalb des logischen Systems darstellen. Die Unterscheidung zwischen Paradoxien und echten Antinomien hängt dabei natürlich davon ab, was für Annahmen für einen Anwendungsbereich vorausgesetzt werden und auf welches logische System man sich bezieht.

zwei Typen der semantischen und der ontologischen Antinomien eingeteilt.[3] Unberücksichtigt bleiben hier also beispielsweise Zenons Paradoxien des Raums, der Zeit und der Bewegung, Sorites Paradoxien, Paradoxien der Bestätigung sowie das Gefangenendilemma.[4] Auch Kants Antinomien der reinen Vernunft, die dem Thema der Arbeit näher stehen als die zuletzt genannten, werden nicht eingehend behandelt.

Die Ursprünge der Unterscheidung zwischen semantischen und ontologischen Antinomien finden sich bei Peano und Ramsey.[5] Semantische Antinomien sind dadurch ausgezeichnet, dass in ihnen semantische Begriffe wie "Wahrheit", "Zutreffen auf", "Definition" usw. eine wesentliche Rolle spielen. Zu den bekanntesten semantischen Antinomien zählen die Lügner-Antinomie, die Antinomien von Curry, Grelling, Richard, Berry und König, die auch in dieser Arbeit behandelt werden. Ontologische Antinomien nehmen keinen Bezug auf semantische Begriffe, sondern letztlich nur auf den Begriff der Menge oder Klasse und den Begriff "Element von". Die Antinomien von Cantor, Burali-Forti und Russell gehören zu den ontonlogischen Antinomien.

Die Unterscheidung nach semantischem Bezug und reinem Mengen- bzw. Klassenbezug scheint auf den ersten Blick klar und berechtigt. Die bekannten Antinomien lassen sich diesen Kategorien leicht zuordnen. Zunächst hat diese Unterscheidung sicher auch viel zu einer klareren Sicht der Antinomien beigetragen, denn der Blick wurde gezielt in Richtung Sprache und Semantik bzw. Logik und mengentheoretische Voraussetzungen gelenkt. Wie jedoch die Beziehung zwischen semantischen und ontologischen Antinomien genau aussieht und inwieweit ein philosophisch relevanter Unterschied besteht, wird in der Geschichte der Antinomiendiskussion sehr unterschiedlich beurteilt. So betonen beispielsweise Peano[6] und Ramsey[7] einen wesentlichen Unterschied; Russell[8], Priest[9] und Champlin[10] äußern sich im Sinne der wesentlichen Ähnlichkeit der Typen.

[3]Semantische Antinomien werden auch als sprachliche Antinomien, ontologische Antinomien auch als logische, mengentheoretische oder mathematische Antinomien bezeichnet. Bezieht man sich mit der Bezeichnung auf die Art der involvierten Objekte, so wären die passendsten Bezeichnungen "semantische Antinomie" und "mengentheoretische Antinomie".

[4]Einen Überblick über diese und andere Paradoxien bietet z. B. Sainsbury (2001).

[5]Vgl. Peano (1906), Ramsey (1925).

[6]Vgl. Peano (1906) und van Heijenoort (1967), S. 142.

[7]Vgl. Ramsey (1925).

[8]Vgl. Russell (1905) und Russell (1908).

[9]Vgl. Priest (2002).

[10]Vgl. Champlin (1988).

In direktem Zusammenhang damit steht die Frage, welche Auswirkungen diese Beurteilung auf mögliche Lösungen für die Antinomien hat. Kann oder muss es zum Beispiel eine einheitliche Lösung für ontologische und semantische Antinomien geben? Die ganz strenge Trennung, die bewirkt hat, dass semantische Antinomien als außermathematisch und damit für die Mathematik unbedeutsam eingestuft wurden, ist spätestens seit dem ersten Unvollständigkeitssatz von Gödel und der dort verwendeten Methode, auf semantische Begriffe über natürliche Zahlen Bezug zu nehmen, hinfällig geworden.

Das Hauptziel dieser Arbeit ist es, zu zeigen, dass und in welcher Form eine enge strukturelle Beziehung zwischen den beiden Typen von Antinomien besteht. Dies geschieht durch die Angabe eines mengentheoretischen Diagonalschemas, auf das die Antinomien bezogen werden. Durch die Gegenüberstellung der beiden Typen von Antinomien auf dieser Basis soll aber auch deutlich gemacht werden, dass trotz der strukturellen Ähnlichkeiten Unterschiede aus philosophischer Perspektive bestehen. Hieran schließt sich die Frage nach der Bedeutung dieser Ergebnisse für Lösungen von Antinomien an, insbesondere die Frage, wie einheitliche Lösungen aussehen können.

Ein weiterer Schwerpunkt wird auf die Einbeziehung logischer bzw. metamathematischer Sätze in die Analyse gelegt, speziell des Satzes von Tarski, der Gödelschen Unvollständigkeitssätze[11] und des Unentscheidbarkeitssatzes. Die Beweise dieser Sätze ähneln in entscheidenden Aspekten ihres Aufbaus den semantischen Antinomien. Diese erstaunliche Tatsache, die vielfach in der Literatur erwähnt wird,[12] soll hier in Form von abstrakten Sätzen genau ausgeführt werden.

In Kapitel 1 geht es zunächst um die Darstellung der Antinomien. Die intuitiven Voraussetzungen, die zum jeweiligen Widerspruch führen, werden herausgearbeitet und formalsprachlich präzisiert. In den meisten Fällen handelt es sich um bekannte Rekonstruktionen der Antinomien. Eine Besonderheit ist jedoch die Hervorhebung zweier sprachlicher Feinunterschiede in den Formulierungen der semantischen Antinomien. Zum einen wird ein Schwerpunkt auf die Unterscheidung zwischen Objekt- und Metasprache gelegt, zum anderen wird, im Zusammenhang hiermit, differenziert zwischen Ansätzen, die von einem Wahrheitsprädikat ausgehen, und solchen, die von einem Falschheitsprädikat ausgehen.

[11] Im Wesentlichen interessiert hierbei der erste Unvollständigkeitssatz. Es wird jedoch auch kurz auf den zweiten Unvollständigkeitssatz eingegangen.

[12] Zum Beispiel von Gödel selbst; vgl. Gödel (1931), S. 175.

Im Anschluss folgen einige mit den semantischen Antinomien verwandte Widerspruchsbeweise der Logik bzw. Metamathematik. Die Darstellung orientiert sich dabei an dem Ziel, Parallelen zu den semantischen Antinomien besonders deutlich werden zu lassen. So werden die Beweise des Satzes von Tarski, der Unvollständigkeitssätze und des Unentscheidbarkeitssatzes nur so weit ausgeführt, wie dies für einen Vergleich mit den semantischen Antinomien sinnvoll ist, und in einer den Antinomien entsprechenden Form angegeben. Der erste Unvollständigkeitssatz wird in einer semantischen Form nach Tarski, in einer semantischen und einer syntaktischen Form nach Gödel und nach dem Ansatz von Rosser bewiesen. Der Beweis des Unentscheidbarkeitssatzes wird nach Church über die Unentscheidbarkeit des Halteproblems und in der Form von Rosser dargestellt.

Für den ontologischen Teil schließt sich ein erster Vergleich an. Hierbei zeigen sich Ähnlichkeiten sowohl zwischen den Antinomien von Russell und Cantor als auch zwischen den Antinomien von Burali-Forti und Cantor, wobei in Bezug auf die Antinomie von Cantor jeweils andere Gesichtspunkte in den Vordergrund treten. Diese drei ontologischen Antinomien lassen sich jedoch alle so darstellen, dass eine gemeinsame Struktur zu Tage tritt, die sich am klarsten in der Russellschen Antinomie zeigt.

In Kapitel 2 werden abstrakte Fassungen der semantischen Antinomien und mit ihnen verwandter Widerspruchsbeweise entwickelt, die auf den Darstellungen des vorangegangenen Kapitels beruhen. Hierzu wird zunächst eine abstrakte Sprache eingeführt. Diese entspricht in ihren wesentlichen Aspekten der von Smullyan[13] und Serény[14] verwendeten Sprache. Die abstrakten Fassungen haben die Form von mathematischen Sätzen mit Beweisen. Der eigentliche Widerspruch der Antinomien tritt entweder in Form einer Äquivalenzaussage als spezieller Einsetzungsfall der Satzaussage oder im indirekten Beweis auf.

Zunächst werden zwei Gruppen von abstrakten Sätzen unterschieden:

- Abstrakte Fassungen der Antinomien von Grelling und Richard: Hier lassen sich die Beweise des Satzes von Tarski und des Unvollständigkeitssatzes zuordnen, die nicht explizit auf den Fixpunktsatz zurückgreifen. Auch der Beweis des Unentscheidbarkeitssatzes von Church fällt in seinem Kern, der sich in der hier gegebenen Darstellung auf Registermaschinen bezieht, darunter.

[13] Vgl. Smullyan (1992).
[14] Vgl. Serény (2003).

- Abstrakte Fassungen der Lügner-Antinomie, die nach Stegmüller in Anlehnung an den mathematischen Fixpunktsatz dargestellt wird: Hier lassen sich Beweise des Satzes von Tarski und des Unvollständigkeitssatzes zuordnen, die auf den Fixpunktsatz zurückgreifen.

Jede dieser beiden Gruppen besteht aus vier abstrakten Sätzen. Diese vier Sätze geben die Unterscheidung zwischen rein objektsprachlichen und teils metasprachlichen Ansätzen sowie zwischen solchen, die von einem Wahrheitsprädikat ausgehen, und solchen, die von einem Falschheitsprädikat ausgehen, wieder. Bei den zum zweiten Punkt gehörenden Sätzen spielt der hier ebenfalls für eine abstrakte Sprache formulierte Fixpunktsatz eine Rolle. Das Verhältnis der Sätze der beiden Gruppen und die Bedeutung des Fixpunktsatzes wird untersucht, wobei sich herausstellt, dass die Sätze der zweiten Gruppe im Wesentlichen Spezialfälle der Sätze der ersten sind.

Ähnliche Ansätze zu Abstraktionen auf diesem Gebiet finden sich bei Smullyan[15] und Serény[16]. In diesem Kapitel wird immer wieder auf Gemeinsamkeiten und Unterschiede zu jenen beiden Ansätzen eingegangen. Insbesondere der Ansatz von Serény, der eine allgemeinere Form der Darstellung wählt, in deren Folge jedoch die Lügner-Antinomie und der Unvollständigkeitssatz nicht präzise interpretiert werden können, wird einem ausführlichen Vergleich mit dem hier vorgestellten Ansatz unterzogen.

Die bisher untersuchten semantischen Antinomien und Widerspruchsbeweise beziehen sich auf Formeln. Parallel können sie aber auch für Terme betrachtet werden. So ergibt sich beispielsweise ein Satz von Tarski für Terme, in dem die Frage nach der objektsprachlichen Ausdrückbarkeit der semantischen Zuordnung eines Terms zu dem von ihm bezeichneten Objekt negativ beantwortet wird. Kapitel 2 endet mit der Übertragung der abstrakten Sätze für Formeln auf abstrakte Sätze für Terme und der Anwendung in Form eines Unvollständigkeitssatzes für Terme.

In Kapitel 3 werden die ontologischen und semantischen Antinomien und Widerspruchsbeweise gemeinsam betrachtet. Auf der Basis von Cantors Diagonalverfahren wird ein mengentheoretisches Schema entwickelt, dessen Einsetzungsfälle Diagonalstrukturen genannt werden. Das eigentliche Diagonalargument findet

[15] Vgl. Smullyan (1992).
[16] Vgl. Serény (2003).

sich in einem auf den Begriff der Diagonalstruktur bezogenen Satz wieder. Das Schema enthält zwei Arten von Diagonalstrukturen, die der Unterscheidung nach Bezug auf ein Wahrheitsprädikat und Bezug auf ein Falschheitsprädikat entsprechen. Im Anschluss an die Entwicklung dieses Schemas wird durch eine semantische Interpretation über einer Struktur gezeigt, wie sich die Antinomien und Widerspruchsbeweise als Diagonalargumente auffassen lassen. Alle behandelten Antiomien und Widerspruchsbeweise mit Ausnahme der Antinomien von Berry und König finden sich hier wieder.

Vor diesem Hintergrund kann nach Analogien in den Antinomietypen gefragt werden und die Voraussetzungen für das Zustandekommen ontologischer und semantischer Antinomien können gegenübergestellt werden. Sie erweisen sich als Voraussetzungen bezüglich Logik und Semantik der Objektsprache und bezüglich der Ontologie – für die ontologischen Antinomien – und der Ausdrucksfähigkeit – für die semantischen Antinomien – der Objektsprache. Obwohl sich die Voraussetzungen auf verschiedene Aspekte der Sprache beziehen, zeigen sie strukturelle Parallelen in ihrer Wirkung bei der Erzeugung einer Diagonalstruktur und der Bildung eines Diagonalarguments. Grob gesagt entspricht eine Klasse als Objekt aus Sicht der ontologischen Antinomien einer einstelligen Formel als Objekt aus Sicht der semantischen Antinomien. Das Enthaltensein in einer Klasse entspricht dem Zutreffen einer einstelligen Formel. Die naive Annahme von beliebigen in der Sprache der Klassentheorie beschreibbaren Zusammenfassungen als Klassen entspricht so der Annahme von einstelligen Formeln, die beliebige Formelschemata erfüllen, in denen auch die metasprachliche Abbildung, die einer Formel ihren Standardnamen in der Objektsprache zuordnet, verwendet werden kann. Das Abstraktionsschema der naiven Klassentheorie entspricht in diesem Sinn der Wahrheitskonvention von Tarski für einen objektsprachlich ausgedrückten Wahrheitsbegriff.

Die meisten bisher bekannten Versuche, strukturelle Ähnlichkeiten verschiedener Antinomien aufzuzeigen, beziehen sich in irgendeiner Art auf die Idee der Diagonalisierung. Die wohl erste Darstellung dieser Art stammt von Russell[17] selbst. Allerdings scheitert sein Schema an der Erfassung der semantischen Antinomien. An dieser Stelle setzt die Arbeit von Priest[18] an, der Russells Schema so erweitert, dass es auch den semantischen Antinomien gerecht wird. In

[17] Vgl. Russell (1905).
[18] Vgl. Priest (2002).

der neueren Zeit ist es besonders er, der sich mit strukturellen Ähnlichkeiten von ontologischen und semantischen Antinomien beschäftigt. Priests Schema unterscheidet sich wesentlich von dem der vorliegenden Arbeit, da es seinen Schwerpunkt gerade nicht auf das Diagonalverfahren legt, das jedoch als möglicher Bestandteil darin integriert sein kann. Das Schema von Priest wird einer ausführlichen Kritik unterzogen, wobei sich herausstellt, dass es keine Klarheit über die Frage bringt, welches die Voraussetzungen für das Zustandekommen der Widersprüche sind und welche dieser Voraussetzungen sich in den einzelnen Antinomien entsprechen.

Ein weiterer früher Ansatz ist neben dem Russellschen der von Nelson und Grelling[19]. Er enthält ein Schema, das dem Schema der vorliegenden Arbeit, genauer einem Teil hiervon, gleicht. Es wird von Nelson und Grelling jedoch nicht auf alle Antinomien angewendet, und der sich daraus ergebende Vergleich ontologischer und semantischer Antinomien wird nicht durchgeführt. Den Abschluss dieses Abschnitts bildet ein Vergleich des Begriffs einer Diagonalstruktur mit dem von Essler[20] präzisierten Begriff des Cantorschen Diagonalverfahrens.

Aufbauend auf die vorangegangene Untersuchung der Voraussetzungen in den einzelnen Antinomien wird im Anschluss eine Übersicht und Klassifizierung von Lösungsansätzen zur Antinomienproblematik gegeben und die Frage nach der Einheitlichkeit von Lösungen diskutiert. Es zeigt sich, dass durch Antinomien Unvollständigkeiten bezüglich der Klassenbildung in der Objektsprache bzw. der Ausdrückbarkeit semantischer Begriffe für die Objektsprache in der Objektsprache deutlich werden – egal für welchen der angesprochenen Lösungswege man sich entscheidet. Denn selbst im Fall von Lösungen, die zunächst den Anschein einer Vollständigkeit bezüglich Klassenbildung und Ausdrückbarkeit semantischer Begriffe erwecken, kehrt die Unvollständigkeit in Form einer verstärkten Antinomie zurück.

[19] Vgl. Nelson u. Grelling (1908).
[20] Vgl. Essler (1964).

Kapitel 1

Antinomien und Sätze der Logik

1.1 Ontologische Antinomien

Ontologische Antinomien basieren auf intuitiven ontologischen Annahmen, die die Typen von Objekten und die Arten von Zusammenfassungen von ihnen betreffen. Diese gilt es jeweils zunächst zu beschreiben. Dabei sollen die Voraussetzungen, die zur Darstellung einer Antinomie, d. h. zur Herleitung des Widerspruchs, gemacht werden, möglichst gering gehalten werden. In den Formulierungen der ontologischen Antinomien wird eine Klassensprechweise und für die Elementbeziehung das Symbol $\in$ verwendet. Die Klassensprechweise wird in Abgrenzung zur metasprachlichen Mengensprechweise in späteren Abschnitten verwendet. Mit ihr wird das logische Konzept von Zusammenfassungen als Extensionen von Begriffen verbunden. Dies entspricht dem naiven Vorgehen, das für die Entstehung der ontologischen Antinomien charakteristisch ist. Im Gegensatz dazu steht das eher mit der Mengensprechweise verbundene iterative Konzept, nach dem die Elemente von Zusammenfassungen kombinatorisch aus schon gebildeten Zusammenfassungen ausgewählt werden.

Im letzten Teil dieses Abschnitts werden die geschilderten ontologischen Antinomien im Zusammenhang betrachtet. Dabei nimmt die Antinomie von Burali-Forti zwar eine Sonderstellung ein, da der zusätzliche Begriff der Wohlordnung eine wesentliche Rolle spielt, sie kann jedoch als Parallele zur Antinomie von Cantor auf der Grundlage von Ordinalzahlen gedeutet werden.

1.1.1 Die Antinomie von Cantor und der Satz von Cantor

Der Satz von Cantor besagt, dass die Mächtigkeit der Potenzklasse einer Klasse echt größer ist als die Mächtigkeit der Klasse selbst. In einer allgemeineren

Form kann der Satz für Abbildungen in eine beliebige Klasse mit mindestens zwei Elementen formuliert werden.

Satz 1.1. *a, a' seien Klassen, wobei $|a'| \succeq 2$.*[1] *b sei die Klasse aller Abbildungen von a nach a'. Dann gilt $|a| \prec |b|$.*

Beweis. $id^+ : a' \to a'$ sei eine Abbildung, so dass für alle $x \in a'$ gilt: $id^+(x) \neq x$. Angenommen $|a| \succeq |b|$. Dann gibt es eine surjektive Abbildung $f : a \to b$. $c \in b$ sei definiert durch $c(x) := id^+(f(x)(x))$ für alle $x \in a$. x_c sei ein Urbild von c unter f. Dann gilt:

$$\begin{aligned} c(x_c) &= id^+(f(x_c)(x_c)) \\ &\neq f(x_c)(x_c) \\ &= c(x_c). \end{aligned}$$

Widerspruch.[2] □

Der Satz von Cantor ist der Spezialfall von Satz 1.1 für $|a'| = 2$. Der Beweis wird für diesen Spezialfall noch einmal formuliert, analog zum allgemeinen Fall, da er in dieser Schreibweise leichter lesbar ist und auch später auf diese Formulierung Bezug genommen wird.

Satz 1.2 (Satz von Cantor). *a sei eine Klasse und b sei die Potenzklasse $P(a)$ von a. Dann gilt $|a| \prec |b|$.*

Beweis. Angenommen $|a| \succeq |b|$. Dann gibt es eine surjektive Abbildung $f : a \to b$. $c \in b$ sei definiert als $c := \{x \in a | x \notin f(x)\}$. x_c sei ein Urbild von c.

[1] Für Klassen a, b soll $|a| \succeq |b|$ bedeuten, dass die Mächtigkeit von a größer oder gleich der Mächtigkeit von b ist. $|a| \prec |b|$ soll bedeuten, dass die Mächtigkeit von a echt kleiner als die Mächtigkeit von b ist.

[2] Eine direkte Anwendung einer Verallgemeinerung dieses Satzes ist der weiter unten folgende Beweis für die Überabzählbarkeit von $\mathbb{R}$. b sei jetzt *eine* Klasse von Abbildungen von a nach a'. $f : a \to b$ sei eine Abbildung. Dann liegt die Abbildung c, die definiert ist durch $c(x) := id^+(f(x)(x))$ für alle $x \in a$, nicht im Bild von f.

Nun wählt man $a = a' = \mathbb{N}$. Sind die reellen Zahlen des offenen Intervalls $[0, 1[$ abzählbar, so ergibt sich aus der Dezimalbruch-Entwicklung in natürlicher Weise eine Abbildung $f : a \to b$. $id^+ : a' \to a'$ bilde $x \in \mathbb{N}$ auf 0 ab, falls $x \neq 0$, und auf 1, falls $x = 0$. Dann liegt die reelle Zahl, die durch $c(x) := id^+(f(x)(x))$ für alle $x \in a$ definiert ist, nicht in der Abzählung, was ein Widerspruch ist. Vgl. Cantor (1890).

Dann gilt:[3]

$$\begin{aligned} x_c \in c \quad & \text{gdw.} \quad x_c \in \{x \in a | x \notin f(x)\} \\ & \text{gdw.} \quad x_c \notin f(x_c) \\ & \text{gdw.} \quad x_c \notin c. \end{aligned}$$

Angenommen $x_c \in c$, dann gilt auch $x_c \notin c$. Es gilt also $x_c \notin c$ und damit auch $x_c \in c$. Widerspruch. □

Die Antinomie von Cantor
a sei die Klasse aller Klassen. Nach Satz 1.2 ist die Potenzklasse $P(a)$ von a von echt größerer Mächtigkeit als a. Andererseits enthält die Klasse aller Klassen alle Klassen als Teilklassen. Damit gibt es keine Klasse, die von echt größerer Mächtigkeit als a ist, insbesondere ist auch $P(a)$ nicht von echt größerer Mächtigkeit als a.

1.1.2 Die Antinomie von Burali-Forti

Definition 1.3. Eine Relation $<$ ist eine *Wohlordnung* auf einer Klasse a gdw. $<$ strikte Ordnung[4] auf a ist und jede nicht-leere Teilklasse von a ein bezüglich $<$ minimales Element[5] hat.

Definition 1.4. Eine Klasse a ist eine *Ordinalzahl* gdw. a transitiv[6] ist und die Elementrelation $\in$ eine Wohlordnung auf der Klasse a ist.

Aus der Definition folgt, dass für alle $\alpha \in \Omega$ mit $\alpha \neq \emptyset$: $\emptyset \in \alpha$. Für alle $x \in \alpha$: $x = \{y | y \in x \wedge y \in \alpha\}$. x ist die Klasse aller Elemente aus α, die bezüglich der Ordnung $\in$ kleiner als x sind.

Ω sei die Klasse aller Ordinalzahlen.

Satz 1.5. *a) Ω ist transitiv.*
b) Ω wird durch $\in$ strikt geordnet.
c) $\in$ ist eine Wohlordnung auf Ω.

[3] "gdw." steht im Folgenden immer für "genau dann, wenn".
[4] Eine Relation ist eine *strikte Ordnung* auf einer Klasse gdw. sie auf der Klasse asymmetrisch, transitiv und total ist.
[5] D. h. zu $b \subseteq a$, $b \neq \emptyset$ gibt es $x \in b$ mit $x < y$ für alle $y \in b, y \neq x$.
[6] Eine Klasse a ist *transitiv* gdw. für alle $x \in a, y \in x$ auch $y \in a$ gilt.

Beweis.
Zu a) Sei $x \in \alpha \in \Omega$. x wird durch $\in$ strikt geordnet. x ist transitiv: Sei $z \in y \in x$. Dann ist $y \in \alpha$ und damit $z \in \alpha$. Es gilt also $z \in x$. Damit ist x Ordinalzahl.
Zu b) Für alle $\alpha, \beta \in \Omega$ mit $\alpha \neq \beta$ gilt $\alpha \in \beta$ oder $\beta \in \alpha$: $\alpha \cap \beta$ ist eine Ordinalzahl und es gilt $\alpha \cap \beta \subseteq \alpha$ und $\alpha \cap \beta \subseteq \beta$. Da α, β Ordinalzahlen sind, gibt es jeweils ein minimales Element in $\alpha \backslash (\alpha \cap \beta), \beta \backslash (\alpha \cap \beta)$, falls $\alpha \backslash (\alpha \cap \beta) \neq \emptyset$, $\beta \backslash (\alpha \cap \beta) \neq \emptyset$. Dieses minimale Element muss in beiden Fällen $\alpha \cap \beta$ sein. Dann liegt es jedoch in $\alpha \cap \beta$, was der Voraussetzung widerspricht. Es gilt also $\alpha = \alpha \cap \beta$ oder $\beta = \alpha \cap \beta$ und damit $\alpha \in \beta$ oder $\beta \in \alpha$.
Zu c) Es sei $c \subseteq \Omega, \alpha \in c$. Wenn $\alpha \cap c = \emptyset$, dann ist α minimales Element in c. Wenn $\alpha \cap c \neq \emptyset$, gibt es ein minimales Element in $\alpha \cap c$, da $\alpha \cap c \subseteq \alpha$ und α Ordinalzahl ist. Dieses Element ist auch minimales Element in c. □

Für alle $\alpha \in \Omega$ gilt: $\alpha = \{y | y \in \alpha \wedge y \in \Omega\}$. α ist also die Klasse aller Ordinalzahlen, die bezüglich der Ordnung $\in$ kleiner als α sind. Die Klasse $\alpha \cup \{\alpha\}$ ist eine Ordinalzahl und α ist kleiner als $\alpha \cup \{\alpha\}$, genauer: $\alpha \cup \{\alpha\}$ ist der Nachfolger von α.

Die Antinomie von Burali-Forti[7]

Nach Satz 1.5 ist Ω eine Ordinalzahl. Damit ist auch $\Omega \cup \{\Omega\}$ eine Ordinalzahl. Ω ist kleiner als $\Omega \cup \{\Omega\}$. Ω ist aber die Ordinalzahl aller Ordinalzahlen und enthält damit alle Ordinalzahlen; insbesondere ist $\Omega \cup \{\Omega\}$ kleiner als Ω. Dies widerspricht den Eigenschaften der Ordnung $\in$.

[7] Vgl. Burali-Forti (1897). Ordinalzahlen werden in der Originalfassung nicht als spezielle wohlgeordnete Klassen, sondern als Äquivalenzklassen ordnungsisomorpher wohlgeordneter Klassen aufgefasst. In diesem Fall ist zunächst zu zeigen, dass es für zwei wohlgeordnete Klassen a, b immer entweder einen Ordnungsisomorphismus von a nach b oder von einem echten Anfangsstück von a nach b oder von einem echten Anfangsstück von b nach a gibt. Zusätzlich ist das jeweilige Anfangsstück eindeutig. Daraus ergibt sich eine strikte Ordnung auf der Klasse aller Ordinalzahlen. Nun kann gezeigt werden, dass jede wohlgeordnete Klasse ordnungsisomorph zu einem eindeutigen Anfangsstück dieser Ordnung ist; und zwar zu dem Anfangsstück, das aus allen echt kleineren Ordinalzahlen besteht. Satz 1.5 c) kann nun analog bewiesen werden.

Geht man, um Antinomien zu vermeiden, z. B. von der Mengenlehre ZF aus, treten Probleme schon an der Stelle der Äquivalenzklassenbildung auf. Die gebildeten Äquivalenzklassen sind keine Mengen mehr. Dies kann durch einen typentheoretischen Ansatz vermieden werden. Äquivalenzklassen von wohlgeordneten Klassen werden dann als Objekte einer höheren Stufe angesehen. Die Klasse aller Ordinalzahlen kann dann ebenfalls als Objekt einer wiederum höheren Stufe aufgefasst werden. Die Antinomie von Burali-Forti zeigt sich, wenn man diese Stufen zusammenfallen lässt. Vgl. Essler u. Brendel (1993), S. 298 ff.

1.1.3 Die Antinomie von Russell

r sei die Klasse aller Klassen, die sich nicht selbst enthalten ($r := \{x|x \notin x\}$). Dann gilt:

$$\begin{aligned} r \in r \quad & \text{gdw.} \quad r \in \{x|x \notin x\} \\ & \text{gdw.} \quad r \notin r. \end{aligned}$$

Angenommen $r \in r$, dann gilt auch $r \notin r$. Es gilt also $r \notin r$ und damit auch $r \in r$.

1.1.4 Erster Vergleich ontologischer Antinomien

Die Antinomie von Cantor und der Satz von Cantor
Die Antinomie von Cantor lässt sich so formulieren, dass sich der Widerspruch der Antinomie als spezielle Anwendung des Satzes von Cantor bzw. des oben formulierten Beweises hierzu ergibt:
a sei die Klasse aller Klassen. Dann ist a von mindestens so großer Mächtigkeit wie $P(a)$; also gibt es eine surjektive Abbildung $f : a \to P(a)$. Damit gilt für $c := \{x \in a|x \notin f(x)\}$ und ein Urbild x_c von c: $x_c \in c$ und $x_c \notin c$.

Diese Darstellung ist die Grundlage der folgenden Herleitung der Russellschen Antinomie aus der Antinomie von Cantor.

Die Antinomie von Cantor und die Antinomie von Russell
Die Klasse aller Klassen ist nicht nur Oberklasse ihrer Potenzklasse, sondern sogar mit dieser identisch: Jede Klasse ist eine Klasse von Klassen, also Element der Potenzklasse der Klasse aller Klassen. Eine spezielle surjektive Abbildung von der Klasse aller Klassen a in ihre Potenzklasse $P(a)$ ist also die identische Abbildung. Verfolgt man die Argumentation, die im Beweis des Satzes von Cantor zum Widerspruch führt, für diesen Spezialfall, erhält man die Antinomie von Russell in der folgenden Form: $\{x \in a|x \notin x\} \in \{x \in a|x \notin x\}$ und $\{x \in a|x \notin x\} \notin \{x \in a|x \notin x\}$. Die Antinomie von Russell ergibt sich dieser Sichtweise nach also aus der Antinomie von Cantor, wenn als surjektive Abbildung $f : a \to P(a)$ speziell die Identität gewählt wird. [8]

[8] Diese Darstellung der Antinomie von Russell als Spezialfall der Antinomie von Cantor findet sich z. B. in Essler (1964). Russell selbst beschreibt dies als den Weg, wie er die nach

Der sich so ergebende Widerspruch – $\{x \in a|x \notin x\} \in \{x \in a|x \notin x\}$ und $\{x \in a|x \notin x\} \notin \{x \in a|x \notin x\}$ – unterscheidet sich jedoch von dem der Antinomie von Russell, wie sie in Abschnitt 1.1.3 dargestellt wird. In der Formulierung der Antinomie in Abschnitt 1.1.3 wird kein Bezug auf die Allklasse a genommen. Von einer Metaebene aus betrachtet lässt sich die Antinomie von Russell dennoch als Spezialfall der Cantorschen auffassen. Die Allklasse a kann dann als Metaklasse eingeführt werden. Sie ist identisch mit dem Teil der Potenzklasse der Allklasse, der nur die Teilklassen der Allklasse enthält, die Klassen im Sinn der Objektsprache sind. $P'(a)$ sei dieser Teil der Potenzklasse $P(a)$. Eine spezielle surjektive Abbildung von der Allklasse a nach $P'(a)$ ist also die identische Abbildung. Verfolgt man wieder die Argumentation, die im Beweis des Satzes von Cantor zum Widerspruch führt, jetzt angewandt auf die Metaklassen a und $P'(a)$, erhält man die Antinomie von Russell, da zu den Voraussetzungen der Antinomie gehört, dass die für den Widerspruch relevante Zusammenfassung $r := \{x|x \notin x\}$ eine Klasse ist und damit in $P'(a)$ liegt.

Die Antinomie von Cantor und die Antinomie von Burali-Forti

Die Antinomie von Burali-Forti ist die Parallele zur Antinomie von Cantor auf der Basis von Ordinalzahlen. Statt allgemeinen Klassen werden nur bestimmte Klassen, nämlich transitive, durch $\in$ wohlgeordnete betrachtet. Einer Klasse von Klassen entspricht im Fall von Ordinalzahlen eine transivite, durch $\in$ wohlgeordnete Klasse von Ordinalzahlen, also ein Anfangsstück von Ω. Einer Teilklasse einer Klasse entspricht ein Anfangsstück einer Ordinalzahl. Der Potenzklasse einer Klasse entspricht die wohlgeordnete Klasse aller Anfangsstücke einer Ordinalzahl α, d. h. $\alpha \cup \{\alpha\}$, also der Nachfolger von α. Einer bijektiven Abbildung zwischen zwei Klassen entspricht eine bijektive, ordnungserhaltende Abbildung zwischen zwei Ordinalzahlen.

Der Satz von Cantor lautet also übertragen auf Ordinalzahlen: α sei eine Ordinalzahl; dann gibt es keine bijektive, ordnungserhaltende Abbildung von einem Anfangsstück von α nach $\alpha \cup \{\alpha\}$, dem Nachfolger von α. Angenommen es gäbe eine bijektive, ordnungserhaltende Abbildung $f : \alpha' \to \alpha \cup \{\alpha\}$, wobei α' ein Anfangsstück von α sei. Es gibt höchstens eine bijektive, ordnungserhaltende Abbildung zwischen zwei Ordinalzahlen: die Identität. Da keine Ordinalzahl sich selbst als Element enthält, ist α' – es gilt $\alpha' \in \alpha \cup \{\alpha\}$ – die Ordinalzahl

ihm benannte Antinomie gefunden hat. Siehe Russell (1903), Abschnitte 346-9 und Russell (1905). Priest (2002), S. 129 verweist hierauf.

aller Ordinalzahlen aus α', die nicht in ihrem Bild enthalten sind. Es gilt also analog zur Antinomie von Cantor $\alpha' \in \alpha'$ und $\alpha' \notin \alpha'$, also ein Widerspruch.[9]

Überträgt man die Antinomie von Cantor von allgemeinen Klassen auf Ordinalzahlen, ergibt sich folgendermaßen die Antinomie von Burali-Forti: Ω sei die Ordinalzahl aller Ordinalzahlen. Wie oben gezeigt, gibt es keine bijektive, ordnungserhaltende Abbildung von einem Anfangsstück von Ω in den Nachfolger $\Omega \cup \{\Omega\}$. Andererseits enthält Ω alle Ordinalzahlen als Anfangsstücke. Damit gibt es zu jeder Ordinalzahl, insbesondere auch zu $\Omega \cup \{\Omega\}$, eine bijektive, ordnungserhaltende Abbildung von einem Anfangsstück von Ω in die Ordinalzahl.

Auch die Antinomie von Russell lässt sich in dieser Weise auf Ordinalzahlen übertragen: α sei die Klasse aller Ordinalzahlen, die sich nicht selbst enthalten. Da keine Ordinalzahl sich selbst enthält, ist α die Klasse aller Ordinalzahlen und damit selbst eine Ordinalzahl. Es gilt also analog zur Antinomie von Russell $\alpha \in \alpha$ und $\alpha \notin \alpha$.

1.2 Semantische Antinomien

Semantische Antinomien basieren auf der intuitiven Verwendung von semantischen Begriffen, die sich in den hier angegebenen Darstellungen auf syntaktische Entitäten beziehen. Die Voraussetzungen, die gemacht werden, sind nicht mehr von allgemein ontologischer Art, also Typen von Objekten und Arten von Zusammenfassungen hiervon betreffend, sondern sie beziehen sich auf die Ausdrückbarkeit semantischer Aspekte einer Objektsprache in der Objektsprache.

[9] Werden stattdessen Äquivalenzklassen ordnungsisomorpher wohlgeordneter Klassen als Ordinalzahlen betrachtet, so ist eine Ordinalzahl α nach Fußnote 7 eindeutig durch die Klasse von Ordinalzahlen bestimmt, die kleiner sind als α. Das Anfangsstück aus diesen Ordinalzahlen ist ein Repräsentant von α.

Wird nun die Elementrelation durch die Kleinerrelation ersetzt, ergibt sich folgende analoge Übertragung des Satzes von Cantor: α sei eine Ordinalzahl; der Potenzklasse einer Klasse entspricht dann die Ordinalzahl β, die durch $\bigwedge x \in \Omega : x < \beta \leftrightarrow (x < \alpha \vee x = \alpha)$ bestimmt ist. Es gibt es keine bijektive, ordnungserhaltende Abbildung von einem Anfangsstück $\{x|x < \alpha'\}$ von $\{x|x < \alpha\}$ in die Klasse $\{x|x < \beta\}$. Beweis: Angenommen es gäbe eine solche bijektive, ordnungserhaltende Abbildung $f : \{x|x < \alpha'\} \to \{x|x < \alpha \vee x = \alpha\}$. Es gibt höchstens eine bijektive, ordnungserhaltende Abbildung zwischen zwei Anfangsstücken der Ordinalzahlordnung: die Identität. Da keine Ordinalzahl kleiner als sie selbst ist, ist α' – es gilt $\alpha' \in \{x|x < \alpha \vee x = \alpha\}$ – die Ordinalzahl, die größer ist als alle Ordinalzahlen aus $\{x|x < \alpha'\}$, die nicht größer als sie selbst sind. Es gilt also analog zum Satz von Cantor $\alpha' < \alpha'$ und $\alpha' \not< \alpha'$.

Diese Voraussetzungen werden nicht in der Sprache der Antinomie, der Objektsprache, sondern in einer Metasprache hierzu als Schemata für die Sprache der Antinomie formuliert. In einer genauen Rekonstruktion muss also zunächst die Sprache der Antinomie, auf die sich die Voraussetzungen ja beziehen, angegeben werden. In der Metasprache werden in dieser Arbeit die Begriffe "Menge" und "$\in$" verwendet. Die verwendete Logik der Metasprache ist hier klassisch, d. h. die Junktoren und Quantoren werden in Argumentationen nach der klassischen Logik verwendet.

Die folgenden Antinomien sind alle seit langem bekannt und werden in zahlreichen Veröffentlichungen, in teilweise recht unterschiedlichen Fassungen, wiedergegeben. In der vorliegenden Darstellung wird ein besonderer Schwerpunkt auf die Unterscheidung von Objekt- und Metasprache in ihrer Bedeutung für die Wiedergabe der Antinomien gelegt. Ferner wird der Unterschied zwischen der Verwendung eines Wahrheitsprädikats und der Verwendung eines Falschheitsprädikats als Ausgangspunkt der Formulierung untersucht. Stellen, an denen Voraussetzungen über die Negation der Objektsprache gemacht werden, treten klar hervor.

Bei den zur Formulierung der Antinomien und der metamathematischen Beweise verwendeten und im Folgenden näher charakterisierten formalen Sprachen handelt es sich um konkrete Sprachen, die jedoch möglichst allgemein gehalten sind.

Allgemeine Voraussetzungen und Bezeichnungsweisen

Für die Objektsprachen, ihre Logik und ihre Semantik sollen für die gesamte Arbeit die folgenden Voraussetzungen gelten, soweit an den jeweiligen Stellen keine abweichenden Voraussetzungen getroffen werden.

$\alpha \geq 1$ sei eine natürliche Zahl. Eine Symbolmenge der Stufe α bestehe aus Objekt-, Relations- und Funktionskonstanten beliebiger Typen mit einer Stufe $\leq \alpha$.[10]

Das Alphabet einer Sprache der Stufe α enthalte an Objektvariablen zu jedem Typ α' einer Stufe $\leq (\alpha - 1)$ genau die abzählbar unendlich vielen, paarweise verschiedenen (auch außerhalb einzelner Typen verschiedenen) Objektvariablen $v_0^{\alpha'}, v_1^{\alpha'}, v_2^{\alpha'}, \ldots$ Handelt es sich um Variablen der Stufe 0, so wird der Typ-Index auch weggelassen.

[10] Genauer sei eine Symbolmenge eine Abbildung, die jedem Typ eine Menge zuordnet.

Für jede Symbolmenge S einer Stufe $\leq \alpha$ seien die Mengen der S^{α}-Terme/Formeln/Sätze die in üblicher Weise definierten Mengen der Terme/Formeln/Sätze der Quantorenlogik α-ter Stufe mit Identität zur Symbolmenge S. Dabei ist Quantifikation über Variablen bis Stufe $(\alpha - 1)$ zugelassen. An den meisten Stellen gelten die Aussagen für alle Stufen. Daher sei α fest, aber beliebig gewählt, falls nichts anderes bestimmt wird. Der Index wird dann weggelassen.

Für $m \in \mathbb{N}$ seien $S^{\alpha}_{[v_0,v_1,...,v_m]}$-Terme und $S^{\alpha}_{[v_0,v_1,...,v_m]}$-Formeln die S^{α}-Terme bzw. -Formeln, die genau die Objektvariablen $v_0, v_1, ..., v_m$ frei enthalten, und $S^{\alpha}_{[v_0,v_1,...,v_m]\subseteq}$-Terme und $S^{\alpha}_{[v_0,v_1,...,v_m]\subseteq}$-Formeln bzw. $S^{\alpha}_{[v_0,v_1,...,v_m]\supseteq}$-Terme und $S^{\alpha}_{[v_0,v_1,...,v_m]\supseteq}$-Formeln die S^{α}-Terme bzw. -Formeln, die höchstens bzw. mindestens die Objektvariablen $v_0, v_1, ..., v_m$ frei enthalten.

Für einen $S^{\alpha}_{[v_0,v_1,...,v_m]}$-Term $t(x_0, ..., x_m)$ oder eine $S^{\alpha}_{[v_0,v_1,...,v_m]}$-Formel $\varphi(x_0, ..., x_m)$ und Terme $a_0, ..., a_m$ von jeweils denselben Typen wie $v_0, ..., v_m$ sei $t(a_0, ..., a_m)$ der Term, der durch simultane Substitution von $v_0, ..., v_m$ durch $a_0, ..., a_m$ aus $t(x_0, ..., x_m)$ hervorgeht, und $\varphi(a_0, ..., a_m)$ die Formel, die durch simultane Substitution von $v_0, ..., v_m$ durch $a_0, ..., a_m$ aus $\varphi(x_0, ..., x_m)$ hervorgeht.

Die Semantik der Objektsprache sei klassisch. Ein Satz $\neg\varphi$ sei also beispielsweise genau dann wahr in einer vorgegebenen Struktur, wenn der Satz φ in dieser Struktur nicht wahr ist.

Die Ableitbarkeitsrelation $\vdash$ sei (semantisch) korrekt. Bezogen auf die erste Stufe einer Sprache sei $\vdash$ (semantisch) vollständig, stelle also die klassische Quantorenlogik erster Stufe dar. Die Menge der ableitbaren Sätze sei aufzählbar.

Zu einer Symbolmenge S einer Stufe $\leq \alpha$ und einer S-Struktur $\mathcal{A}$ sei $\mathrm{Th}^{\alpha}(\mathcal{A})$ die Menge der in $\mathcal{A}$ wahren S^{α}-Sätze.

Zu einer Symbolmenge S einer Stufe $\leq \alpha$ und einer Menge von S^{α}-Sätzen Φ sei ${\Phi^{\vdash}}_{\alpha}$ die Menge der S^{α}-Sätze, die aus Φ ableitbar sind.

1.2.1 Die Lügner-Antinomie

Es wird die folgende, häufig verwendete Formulierung der Lügner-Antinomie zu Grunde gelegt.[11]

[11] Vgl. z. B. Brendel (1992), S. 4.

Es wird ausgegangen von dem Satz χ,
$\chi :=$ "Dieser Satz ist falsch."
(oder, in alternativer Formulierung, $\chi : \chi$ ist falsch.).
Es ergibt sich: Wenn der Satz χ wahr ist, stimmt, was er ausdrückt, d. h. er ist falsch. Wenn der Satz χ falsch ist, stimmt, was er ausdrückt, d. h. er ist wahr.

Rekonstruktion der Lügner-Antinomie

$S_{\text{Lü}}$ sei die Symbolmenge der Sprache der Lügner-Antinomie. Um die Antinomie formulieren zu können, sei zunächst vorausgesetzt, dass die Sprache Standardterme für alle $S_{\text{Lü}}$-Sätze enthält. Dies soll zunächst nur heißen, dass jedem $S_{\text{Lü}}$-Satz φ ein fester $S^1_{\text{Lü}}$-Term $\ulcorner\varphi\urcorner$ zugeordnet ist. Die leicht fettgedruckten Haken mögen im Folgenden immer den Standardterm einer Zeichenkette andeuten.

Die Argumentation der Lügner-Antinomie spielt sich im Rahmen der klassischen Logik ab. Sie kann also als logische Folgerung in Bezug auf den klassischen Wahrheitsbegriff der Wahrheit in einer $S_{\text{Lü}}$-Struktur nach der klassischen Semantik dargestellt werden. Interpretiert man die Sprache in einer $S_{\text{Lü}}$-Struktur $\mathcal{A}_{\text{Lü}} = (U_{\text{Lü}}, j_{\text{Lü}})$, muss zunächst vorausgesetzt werden, dass alle $S_{\text{Lü}}$-Sätze Elemente von $U_{\text{Lü}}$ sind.[12] $j_{\text{Lü}}$ bilde den Standardterm einer Formel auf die Formel ab.

In der Formulierung der Antinomie wird nun von der Wahrheit von Sätzen gesprochen. Um die Argumentation in der formalen Sprache nachvollziehen zu können, wird eine $S_{\text{Lü}[v_0]}$-Formel $w(x)$ angenommen, so dass für alle $S_{\text{Lü}}$-Sätze φ gilt

$$\mathcal{A}_{\text{Lü}} \models \varphi \leftrightarrow w(\ulcorner\varphi\urcorner).$$

Der Ausdruck "x ist falsch" wird durch die $S_{\text{Lü}[v_0]}$-Formel $\neg w(x)$ wiedergegeben. Zur Formulierung des Lügner-Satzes "Dieser Satz ist falsch." wird nun vorausgesetzt, dass es einen $S_{\text{Lü}}$-Satz χ gibt mit

$$\mathcal{A}_{\text{Lü}} \models \chi \leftrightarrow \neg w(\ulcorner\chi\urcorner).$$

χ entspricht dann dem Lügner-Satz.

[12] Diese Voraussetzung ist allerdings aus formaler Sicht unnötig, da die Standardterme der $S_{\text{Lü}}$-Formeln beliebig auf paarweise verschiedene Elemente abgebildet werden können. Die Voraussetzung entspricht aber der Sprechweise der Antinomie.

Die Argumentation der Lügner-Antinomie lässt sich, da sie die logischen Zeichen in klassischem Sinn verwendet, in Form des KNS[13] für die klassische Logik formulieren. Hierbei werden nur auch intuitionistisch gültige Regeln verwendet. Sie lautet:

1	$\chi \leftrightarrow w(\ulcorner\chi\urcorner)$	Voraussetzung
2	$\chi \leftrightarrow \neg w(\ulcorner\chi\urcorner)$	Voraussetzung
3(3)	$w(\ulcorner\chi\urcorner)$	AE
4(3)	χ	$\leftrightarrow$ B, $\rightarrow$ B,1,3
5(3)	$\neg w(\ulcorner\chi\urcorner)$	$\leftrightarrow$ B, $\rightarrow$ B,2,4
6	$w(\ulcorner\chi\urcorner) \rightarrow \neg w(\ulcorner\chi\urcorner)$	$\rightarrow$ E,*3*,5
7(7)	$\neg w(\ulcorner\chi\urcorner)$	AE
8(7)	χ	$\leftrightarrow$ B, $\rightarrow$ B,2,7
9(7)	$w(\ulcorner\chi\urcorner)$	$\leftrightarrow$ B, $\rightarrow$ B,1,7
10	$\neg w(\ulcorner\chi\urcorner) \rightarrow w(\ulcorner\chi\urcorner)$	$\rightarrow$ E,*7*,9
11	$\neg w(\ulcorner\chi\urcorner) \leftrightarrow w(\ulcorner\chi\urcorner)$	$\leftrightarrow$ E,6,10

Folgende Überlegung, in Form des KNS formuliert, führt zu einem expliziten Widerspruch:

1	$\neg w(\ulcorner\chi\urcorner) \leftrightarrow w(\ulcorner\chi\urcorner)$	Voraussetzung
2(2)	$w(\ulcorner\chi\urcorner)$	AE
3(2)	$\neg w(\ulcorner\chi\urcorner)$	$\leftrightarrow$ B, $\rightarrow$ B,1,2
4	$\neg w(\ulcorner\chi\urcorner)$	WS,*2*,2,3
5	$w(\ulcorner\chi\urcorner) \wedge \neg w(\ulcorner\chi\urcorner)$	$\leftrightarrow$ B, $\rightarrow$ B, $\wedge$ E,1,4

Die letzte Voraussetzung der Existenz eines $S_{\text{Lü}}$-Satzes χ, so dass $\mathcal{A}_{\text{Lü}} \models \chi \leftrightarrow \neg w(\ulcorner\chi\urcorner)$, lässt sich erzwingen, indem vorausgesetzt wird, dass es einen $S_{\text{Lü}[v_0]}$-Term $\alpha(x)$ gibt, der semantisch der Selbsteinsetzung von $S_{\text{Lü}[v_0]}$-Formeln in sich entspricht. Dazu wird also zunächst angenommen, dass auch für alle $S_{\text{Lü}[v_0]}$-Formeln Standardterme in der Sprache enthalten sind. Es gelte hierzu ferner für alle $S_{\text{Lü}[v_0]}$-Formeln $\varphi(x)$: $\mathcal{A}_{\text{Lü}} \models \alpha(\ulcorner\varphi(x)\urcorner) \equiv \ulcorner\varphi(\ulcorner\varphi(x)\urcorner)\urcorner$. Dann ist $\neg w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner))$ ein Satz der gesuchten Art, denn es folgt

$$\mathcal{A}_{\text{Lü}} \models \neg w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner)) \leftrightarrow \neg w(\ulcorner\neg w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner))\urcorner).$$

[13]Kalkül des Natürlichen Schließens. Vgl. Essler u. a. (2001).

$\neg w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner))$ entspricht damit dem Lügner-Satz “Dieser Satz ist falsch.”[14]

Eine aus metasprachlicher Sicht rekonstruierte Variante ergibt sich folgendermaßen:
Es wird vorausgesetzt, dass es ein Wahrheitsprädikat in der Sprache gibt, d. h. eine $S_{\text{Lü}[v_0]}$-Formel $w(x)$, so dass für alle $S_{\text{Lü}}$-Sätze φ gilt:

$$\mathcal{A}_{\text{Lü}} \models \varphi \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models w(\ulcorner\varphi\urcorner).$$

Zur Formulierung des Lügner-Satzes wird nun ein $S_{\text{Lü}}$-Satz χ vorausgesetzt mit:

$$\mathcal{A}_{\text{Lü}} \models \chi \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models \neg w(\ulcorner\chi\urcorner).$$

Die Existenz eines $S_{\text{Lü}}$-Satzes χ mit der obigen Eigenschaft folgt wiederum aus der Existenz eines $S_{\text{Lü}[v_0]}$-Terms $\alpha(x)$, so dass für alle $S_{\text{Lü}[v_0]}$-Formeln $\varphi(x), \psi(x)$ gilt:

$$\mathcal{A}_{\text{Lü}} \models \psi(\alpha(\ulcorner\varphi(x)\urcorner)) \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models \psi(\ulcorner\varphi(\ulcorner\varphi(x)\urcorner)\urcorner).$$

Dann ist $\neg w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner))$ ein Satz der gesuchten Art. Die Argumentation der Rekonstruktion verläuft analog zu obiger:

$$\begin{aligned} \mathcal{A}_{\text{Lü}} \models \neg w(\ulcorner\chi\urcorner) \quad &\text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models w(\ulcorner\chi\urcorner), \text{ und damit} \\ \mathcal{A}_{\text{Lü}} \models w(\ulcorner\chi\urcorner) \quad &\text{und} \quad \mathcal{A}_{\text{Lü}} \models \neg w(\ulcorner\chi\urcorner). \end{aligned}$$

Der Unterschied dieser Rekonstruktionen besteht in der einmal objekt- und einmal metasprachlichen Wiedergabe der Äquivalenz. Die zweite, metasprachliche Variante ist sinnvoll für spätere Vergleiche mit Versionen der Lügner-Antinomie innerhalb einer dreiwertigen Semantik – in Zusammenhang mit der verstärkten Lügner-Antinomie – und mit dem Gödelschen Unvollständigkeitssatz.

Die Antinomie von Curry[15]

Die Antinomie von Curry stellt eine Variante der Lügner-Antinomie dar. Der Widerspruch wird hier nicht durch einen Satz und seine Negation

[14] In einem umgangssprachlichen Pendant zu dieser Rekonstruktion wird der Lügner-Satz von Stegmüller wiedergegeben. Siehe Stegmüller (1957), S. 31 f. Stegmüller orientiert sich hier an dem Satz, den Gödel für seinen Unvollständigkeitssatz verwendet hat.

[15] Vgl. Curry (1942).

$(w(\ulcorner\chi\urcorner) \wedge \neg w(\ulcorner\chi\urcorner))$, sondern durch einen beliebigen anderen Satz, der einen Widerspruch ausdrückt, dargestellt. Wird als ein solcher Satz $0 \equiv 1$ gewählt, baut die Antinomie von Curry statt auf den Lügner-Satz auf den Satz "Wenn dieser Satz wahr ist, dann ist $0 \equiv 1$." auf. Nach dem Prinzip der Rekonstruktion des Lügner-Satzes lässt sich dieser Satz rekonstruieren als $w(\alpha(\ulcorner w(\alpha(x)) \to 0 \equiv 1\urcorner)) \to 0 \equiv 1$. Wird dieser Satz mit χ bezeichnet, gilt nämlich $\mathcal{A}_{\text{Lü}} \models \chi \leftrightarrow (w(\chi) \to 0 \equiv 1)$. Analog zum Fall der Lügner-Antinomie erhält man $\mathcal{A}_{\text{Lü}} \models w(\ulcorner\chi\urcorner) \leftrightarrow (w(\ulcorner\chi\urcorner) \to 0 \equiv 1)$. Die weitere Argumentation lautet:[16]

1	$(w(\ulcorner\chi\urcorner) \to 0 \equiv 1) \leftrightarrow w(\ulcorner\chi\urcorner)$	Voraussetzung
2(2)	$w(\ulcorner\chi\urcorner)$	AE
3(2)	$w(\ulcorner\chi\urcorner) \to 0 \equiv 1$	$\leftrightarrow$ B, $\to$ B, 1,2
4(2)	$0 \equiv 1$	$\to$ B, 2,3
5	$w(\ulcorner\chi\urcorner) \to 0 \equiv 1$	$\to$ E, 2,4
6	$w(\ulcorner\chi\urcorner)$	$\leftrightarrow$ B, $\to$ B, 1,5
7	$0 \equiv 1$	$\to$ B, 5,6

Die verstärkte Lügner-Antinomie

Die verstärkte Lügner-Antinomie ist eine Variante der Lügner-Antinomie innerhalb einer mehrwertigen Semantik.[17] Bezogen auf eine dreiwertige Semantik kann diese Variante etwa so formuliert werden:

> Es wird ausgegangen von dem Satz χ,
> $\chi :=$ "Dieser Satz ist falsch oder unbestimmt."
> Es ergibt sich: Wenn der Satz χ wahr ist, stimmt, was er ausdrückt, d. h. er ist falsch oder unbestimmt. Wenn der Satz χ falsch oder unbestimmt ist, stimmt ja, was er ausdrückt, d. h. er ist wahr.

Zur Rekonstruktion werden zunächst die folgenden Abkürzungen eingeführt: "$\models_{\mathrm{W}}$", "$\models_{\mathrm{F}}$", "$\models_{\mathrm{U}}$" für "ist wahr in", "ist falsch in", "ist unbestimmt in". Diese Begriffe werden jeweils auf eine Struktur im Sinne eines dreiwertigen Logiksystems bezogen. Anders als in einem zweiwertigen Logiksystem wird in

[16] Die Argumentation kann für jeden beliebigen Satz anstelle von $0 \equiv 1$ durchgeführt werden. Da die verwendeten Regeln alle schon in der Minimalen Logik von Prawitz gelten, in der allerdings nicht die Regel "ex falso quodlibet" – wie noch in der intuitionistischen Logik – gilt, kann auf diesem Weg auch in der Minimalen Logik jeder Satz abgeleitet werden. Vgl. Feferman (1984), S. 80.

[17] Vgl. Brendel (1992).

diesem Fall einer einstelligen Formel – und das Gleiche gilt für mehrstellige Formeln analog – nicht nur eine Teilmenge des Universums als Extension und das Komplement hierzu als Antiextension zugordnet, sondern das Universum wird in drei disjunkte Teile zerlegt: die Extension des Begriffs, die Antiextension und den verbleibenden dritten Teil, auf dem der Begriff nicht anwendbar ist. Wenn in diesem Abschnitt von Strukturen die Rede ist, insbesondere von $S_{\text{Lü}}$-Strukturen, so sollen immer Strukturen im Sinn eines dreiwertigen Logiksystems gemeint sein. Die Zuordnung der Wahrheitswerte nicht-atomarer Sätze sei durch die starken oder schwachen Kleeneschen Wahrheitsregeln gegeben.[18]

Zunächst wird eine Rekonstruktion der einfachen Lügner-Antinomie innerhalb einer dreiwertigen Semantik angegeben:
Wählt man eine $S_{\text{Lü}}$-Struktur im Sinn eines dreiwertigen Logiksystems und verwendet die starken oder schwachen Kleeneschen Wahrheitsregeln, so folgt aus den Voraussetzungen über $w(x)$ – $\mathcal{A}_{\text{Lü}} \models \varphi \leftrightarrow w(\ulcorner\varphi\urcorner)$ – und $\alpha(x)$ – $\mathcal{A}_{\text{Lü}} \models \alpha(\ulcorner\varphi(x)\urcorner) \equiv \ulcorner\varphi(\ulcorner\varphi(x)\urcorner)\urcorner$ – weiterhin logisch $w(\chi) \wedge \neg w(\chi)$, woraus auch nach den Kleeneschen Wahrheitsregeln jeder Satz logisch folgt. Die Ursache hierfür ist, dass eine Äquivalenz innerhalb der dreiwertigen Semantik nach Kleene nur dann wahr ist, wenn die äquivalenten Sätze den Wahrheitswert "wahr" oder "falsch" haben, also nicht unbestimmt sind. Die Antinomie scheint also zunächst auch innerhalb der dreiwertigen Semantik zustande zu kommen. Bei näherer Betrachtung der Voraussetzungen wird jedoch klar, dass sie, sollen sie als eine Rekonstruktion der Lügner-Antinomie dienen, nicht unabhängig von dem metasprachlichen Wahrheitsbegriff für die Objektsprache gebildet werden dürfen. In einer $S_{\text{Lü}}$-Struktur, in der der Satz $\varphi \leftrightarrow w(\ulcorner\varphi\urcorner)$ nach den Kleeneschen Regeln wahr ist, ist der Satz φ wahr oder falsch. Die Voraussetzung: $\varphi \leftrightarrow w(\ulcorner\varphi\urcorner)$ ist wahr für alle $S_{\text{Lü}}$-Sätze φ, würde also implizieren, dass in den betrachteten Strukturen kein Satz den Wahrheitswert "unbestimmt" hat, was nicht die Intention einer Formulierung der Lügner-Antinomie, insbesondere eines objektsprachlichen Wahrheitsbegriffs, innerhalb eines dreiwertigen Logiksystems sein kann.

Stattdessen lässt sich die metasprachliche Variante für eine dreiwertige Semantik wie folgt wiedergeben: Vorausgesetzt wird, dass es ein Wahrheitsprädikat in der

[18] Vgl. z. B. Brendel (1992).

Sprache gibt, d. h. eine $S_{\text{Lü}[v_0]}$-Formel $w(x)$, so dass für alle $S_{\text{Lü}}$-Sätze φ gilt:

$$\mathcal{A}_{\text{Lü}} \models_{\text{W}} \varphi \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models_{\text{W}} w(\ulcorner\varphi\urcorner).$$

Zur Formulierung des Lügner-Satzes "Dieser Satz ist falsch." wird nun vorausgesetzt, dass es einen $S_{\text{Lü}}$-Satz χ gibt mit:

$$\mathcal{A}_{\text{Lü}} \models_{\text{W}} \chi \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models_{\text{W}} \neg w(\ulcorner\chi\urcorner).$$

Die letzte Voraussetzung wird wieder durch den Satz $\chi := \neg w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner))$ erfüllt, wobei $\alpha(x)$ ein $S_{\text{Lü}[v_0]}$-Term sei, so dass für alle $S_{\text{Lü}[v_0]}$-Formeln $\varphi(x), \psi(x)$ gilt:

$$\mathcal{A}_{\text{Lü}} \models_{\text{W}} \psi(\alpha(\ulcorner\varphi(x)\urcorner)) \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models_{\text{W}} \psi(\ulcorner\varphi(\ulcorner\varphi(x)\urcorner)\urcorner).$$

Analog zum zweiwertigen Fall gilt mit diesen Voraussetzungen:

$$\mathcal{A}_{\text{Lü}} \models_{\text{W}} \neg w(\ulcorner\chi\urcorner) \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models_{\text{W}} w(\ulcorner\chi\urcorner).$$

Ein Widerspruch lässt sich jetzt jedoch nicht mehr so erzeugen, wie das für den Fall der zweiwertigen Semantik möglich war: Angenommen $\mathcal{A}_{\text{Lü}} \models_{\text{W}} w(\ulcorner\chi\urcorner)$. Dann gilt auch $\mathcal{A}_{\text{Lü}} \models_{\text{W}} \neg w(\ulcorner\chi\urcorner)$. Daraus folgt jedoch nicht $\mathcal{A}_{\text{Lü}} \models_{\text{W}} \neg w(\ulcorner\chi\urcorner)$, also $\mathcal{A}_{\text{Lü}} \models_{\text{F}} w(\ulcorner\chi\urcorner)$, denn es gibt ja jetzt die dritte Möglichkeit, dass $\mathcal{A}_{\text{Lü}} \models_{\text{U}} w(\ulcorner\chi\urcorner)$.

Insofern erweist sich der (einfache) Lügner-Satz als nicht antinomisch innerhalb der dreiwertigen Semantik.

Rekonstruktion der verstärkten Lügner-Antinomie:
Im Fall der einfachen Lügner-Antinomie wurde "falsch" in der rekonstruierten Sprache durch "nicht wahr" ausgedrückt. Um im Fall des verstärkten Lügners, dann den Begriff "falsch oder unbestimmt" betreffend, analog vorgehen zu können, müsste die rekonstruierte Sprache um eine weitere Negation erweitert werden, die semantisch der metasprachlichen Negation entspricht. Eine andere Möglichkeit besteht darin, diese Negation nur in Zusammenhang mit dem Wahrheitsprädikat einzuführen. Sie entspricht der Einführung eines Prädikats "falsch oder unbestimmt" und wird hier gewählt. $\neg^{\text{m}} w(x)$ sei diese $S_{\text{Lü}[v_0]}$-Formel; es

gelte also für alle $S_{\text{Lü}}$-Sätze φ:

$$(\mathcal{A}_{\text{Lü}} \models_{\text{F}} w(\varphi) \text{ oder } \mathcal{A}_{\text{Lü}} \models_{\text{U}} w(\varphi)) \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models_{\text{W}} \neg^{\text{m}} w(\ulcorner \varphi \urcorner).$$

Zur Formulierung des Lügner-Satzes "Dieser Satz ist falsch oder unbestimmt." wird nun vorausgesetzt, dass es einen $S_{\text{Lü}}$-Satz χ gibt mit:

$$\mathcal{A}_{\text{Lü}} \models_{\text{W}} \chi \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models_{\text{W}} \neg^{\text{m}} w(\ulcorner \chi \urcorner).$$

$\chi := \neg^{\text{m}} w(\alpha(\ulcorner \neg^{\text{m}} w(\alpha(x)) \urcorner))$ erfüllt diese Voraussetzung. Damit gilt:

$$\mathcal{A}_{\text{Lü}} \models_{\text{W}} \neg^{\text{m}} w(\ulcorner \chi \urcorner) \quad \text{gdw.} \quad \mathcal{A}_{\text{Lü}} \models_{\text{W}} w(\ulcorner \chi \urcorner).$$

Folgende Überlegung führt zu einem expliziten Widerspruch:
Angenommen $\mathcal{A}_{\text{Lü}} \models_{\text{W}} w(\ulcorner \chi \urcorner)$. Dann gilt auch $\mathcal{A}_{\text{Lü}} \models_{\text{W}} \neg^{\text{m}} w(\ulcorner \chi \urcorner)$, also $\mathcal{A}_{\text{Lü}} \models_{\text{F}} w(\ulcorner \chi \urcorner)$ oder $\mathcal{A}_{\text{Lü}} \models_{\text{U}} w(\ulcorner \chi \urcorner)$, was ein Widerspruch ist. Es folgt also ($\mathcal{A}_{\text{Lü}} \models_{\text{F}} w(\ulcorner \chi \urcorner)$ oder $\mathcal{A}_{\text{Lü}} \models_{\text{U}} w(\ulcorner \chi \urcorner)$) und damit $\mathcal{A}_{\text{Lü}} \models_{\text{W}} \neg^{\text{m}} w(\ulcorner \chi \urcorner)$. Dies impliziert jedoch auch $\mathcal{A}_{\text{Lü}} \models_{\text{W}} w(\ulcorner \chi \urcorner)$.

Die verstärkte Lügner-Antinomie innerhalb eines dreiwertigen Logiksystems entspricht der einfachen Lügner-Antinomie innerhalb eines zweiwertigen Logiksystems, da in beiden Fällen die Negation – im dreiwertigen Fall $\neg^{\text{m}} w$ – der metasprachlichen Negation entsprechend interpretiert wird.

1.2.2 Die Antinomie von Grelling

Die folgende Formulierung der Antinomie von Grelling wird zugrunde gelegt.

> Es wird ausgegangen von dem Begriff "heterologisch". Ein einstelliger Begriff heißt *heterologisch* gdw. er nicht auf sich selbst zutrifft. Es ergibt sich: Wenn der Begriff "heterologisch" heterologisch ist, trifft er nicht auf sich selbst zu, d. h. er ist nicht heterologisch. Wenn der Begriff "heterologisch" nicht heterologisch ist, trifft er nicht auf sich selbst zu, d. h. er ist heterologisch.

Rekonstruktion der Antinomie von Grelling:
S_{Gr} sei die Symbolmenge der Sprache der Antinomie von Grelling. Die Sprache

enthalte Standardterme für alle $S_{\mathrm{Gr}[v_0]}$-Formeln. Die Sprache sei interpretiert über einer S_{Gr}-Struktur $\mathcal{A}_{\mathrm{Gr}} = (U_{\mathrm{Gr}}, j_{\mathrm{Gr}})$, deren Universum alle $S_{\mathrm{Gr}[v_0]}$-Formeln enthalte und deren Interpretation den Standardterm einer Formel auf sich selbst abbilde.

Entscheidend ist die Bildung des Begriffs "heterologisch". Es wird also vorausgesetzt, dass es eine $S_{\mathrm{Gr}[v_0]}$-Formel $het(x)$ gibt, so dass für jede $S_{\mathrm{Gr}[v_0]}$-Formel $\varphi(x)$ gilt

$$\mathcal{A}_{\mathrm{Gr}} \models \neg\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow het(\ulcorner\varphi(x)\urcorner).$$

Nach Einsetzen von $het(x)$ ergibt sich

$$\mathcal{A}_{\mathrm{Gr}} \models \neg het(\ulcorner het(x)\urcorner) \leftrightarrow het(\ulcorner het(x)\urcorner).$$

Und damit gilt auch

$$\mathcal{A}_{\mathrm{Gr}} \models het(\ulcorner het(x)\urcorner) \wedge \neg het(\ulcorner het(x)\urcorner).$$

Eine metasprachliche Variante ergibt sich folgendermaßen: Es wird vorausgesetzt, dass es eine $S_{\mathrm{Gr}[v_0]}$-Formel $het(x)$ gibt, so dass für jede $S_{\mathrm{Gr}[v_0]}$-Formel $\varphi(x)$ gilt:

$$\mathcal{A}_{\mathrm{Gr}} \not\models \varphi(\ulcorner\varphi(x)\urcorner) \quad \text{gdw.} \quad \mathcal{A}_{\mathrm{Gr}} \models het(\ulcorner\varphi(x)\urcorner).$$

Es ergibt sich:

$$\begin{aligned}&\mathcal{A}_{\mathrm{Gr}} \not\models het(\ulcorner het(x)\urcorner) \quad \text{gdw.} \quad \mathcal{A}_{\mathrm{Gr}} \models het(\ulcorner het(x)\urcorner), \text{ und damit:}\\ &\mathcal{A}_{\mathrm{Gr}} \models het(\ulcorner het(x)\urcorner) \quad \text{und} \quad \mathcal{A}_{\mathrm{Gr}} \not\models het(\ulcorner het(x)\urcorner).\end{aligned}$$

1.2.3 Die Antinomie von Richard

> Es wird ausgegangen von einer injektiven Zuordnung der einstelligen Begriffe zu den natürlichen Zahlen. Ferner wird von dem Begriff "richardisch" ausgegangen. Das Bild eines einstelligen Begriffs heißt *richardisch* gdw. der Begriff nicht auf es zutrifft.
> Es ergibt sich: Wenn das Bild des Begriffs "richardisch" richardisch ist, trifft der Begriff nicht auf es zu, d. h. es ist nicht richardisch. Wenn das Bild des Begriffs "richardisch" nicht richardisch ist, trifft

der Begriff nicht auf es zu, d. h. es ist richardisch.[19]

Rekonstruktion der Antinomie von Richard:
Die Symbolmenge S_{N} sei definiert als $S_{\mathrm{N}} := \{0, \mathit{nf}\}$, wobei 0 eine Objektkonstante sei und nf eine einstellige Funktionskonstante sei. $\mathcal{A}_{\mathrm{N}}$ sei die S_{N}-Struktur $(\mathbb{N}, 0^{\mathbb{N}}, (+1)^{\mathbb{N}})$. Für $n \in \mathbb{N}$ sei $\boldsymbol{n}$ der Standardterm für n aus den S^1_{N}-Termen.

S_{Ri} sei die Symbolmenge der Sprache der Antinomie von Richard. Die Symbolmenge S_{Ri} enthalte die Symbolmenge S_{N} und sei höchstens abzählbar. Die Sprache sei interpretiert über einer S_{Ri}-Struktur $\mathcal{A}_{\mathrm{Ri}} = (U_{\mathrm{Ri}}, j_{\mathrm{Ri}})$, deren Universum $\mathbb{N}$ enthalte. $\mathcal{A}_{\mathrm{Ri}}$ enthalte $\mathcal{A}_{\mathrm{N}}$, nach Einschränkung auf die Symbolmenge S_{N}, als Substruktur.

Da die Menge der $S_{\mathrm{Ri}[v_0]}$-Formeln abzählbar ist, gibt es eine injektive Abbildung $\ulcorner\cdot\urcorner : \{\varphi(x) | \varphi(x) \text{ ist } S_{\mathrm{Ri}[v_0]}\text{-Formel}\} \to \mathbb{N}$.
Entscheidend ist die Bildung des Begriffs "richardisch". Es wird also vorausgesetzt, dass es eine $S_{\mathrm{Ri}[v_0]}$-Formel $ri(x)$ gibt, so dass für jede $S_{\mathrm{Ri}[v_0]}$-Formel $\varphi(x)$ gilt

$$\mathcal{A}_{\mathrm{Ri}} \models \neg\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow ri(\ulcorner\varphi(x)\urcorner).$$

Nach Einsetzen von $ri(x)$ ergibt sich

$$\mathcal{A}_{\mathrm{Ri}} \models \neg ri(\ulcorner ri(x)\urcorner) \leftrightarrow ri(\ulcorner ri(x)\urcorner).$$

Und damit gilt auch

$$\mathcal{A}_{\mathrm{Ri}} \models ri(\ulcorner ri(x)\urcorner) \wedge \neg ri(\ulcorner ri(x)\urcorner).$$

Eine metasprachliche Variante ergibt sich analog zur Antinomie von Grelling.

Häufig wird die Antinomie von Richard auch ausgehend von einstelligen Funktionen anstatt einstelliger Begriffe formuliert. Für alle S_{Ri}-Terme r gelte $\mathcal{A}_{\mathrm{Ri}} \models \mathit{nf}(r) \not\equiv r$. Es wird vorausgesetzt, dass es einen $S_{\mathrm{Ri}[v_0]}$-Term $r(x)$ gibt, so dass für jeden $S_{\mathrm{Ri}[v_0]}$-Term $t(x)$ gilt

$$\mathcal{A}_{\mathrm{Ri}} \models \mathit{nf}(t(\ulcorner t(x)\urcorner)) \equiv r(\ulcorner t(x)\urcorner).$$

[19] Dies ist eine abgewandelte Form der von Richard geschilderten Antinomie. Die im Anschluss dargestellte, auf Funktionen statt einstelligen Begriffen basierende Fassung entspricht eher der ursprünglichen. Vgl. Richard (1905).

Nach Einsetzen von $r(x)$ ergibt sich

$$\mathcal{A}_{\mathrm{Ri}} \models nf(r(\ulcorner r(x) \urcorner)) \equiv r(\ulcorner r(x) \urcorner).$$

1.2.4 Die Antinomien von Berry und König

Die Antinomie von Berry

> Es werden Definitionen natürlicher Zahlen betrachtet. Da es nur endlich viele Definitionen mit weniger als 20 Worten gibt, gibt es genau eine natürliche Zahl m, auf die die Formel $\psi(x) :=$ "x ist kleinste Zahl unter denjenigen natürlichen Zahlen, die nicht mit weniger als 20 Worten definierbar sind" zutrifft. $\psi(x)$ stellt also eine Definition der natürlichen Zahl m dar. Da sie aus 16 Worten besteht, ergibt sich: Die natürliche Zahl m ist mit weniger als 20 Worten definierbar. Andererseits ist m nicht mit weniger als 20 Worten definierbar, da $\psi(x)$ auf m zutrifft und dies besagt.[20]

Rekonstruktion der Antinomie von Berry:
Die Symbolmenge $S_{\mathrm{N}}^{<}$ sei definiert als $S_{\mathrm{N}}^{<} := \{0, nf, <\}$, wobei 0 eine Objektkonstante, nf eine einstellige Funktionskonstante und $<$ eine zweistellige Relationskonstante sei. $\mathcal{A}_{\mathrm{N}}^{<}$ sei die $S_{\mathrm{N}}^{<}$-Struktur $(\mathbb{N}, 0^{\mathbb{N}}, (+1)^{\mathbb{N}}, <^{\mathbb{N}})$. Für $n \in \mathbb{N}$ sei $\boldsymbol{n}$ der Standardterm für n aus den $S_{\mathrm{N}}^{<}$-Termen.

S_{Be} sei die Symbolmenge der Sprache der Antinomie von Berry. S_{Be} enthalte die Symbolmenge $S_{\mathrm{N}}^{<}$. Die Sprache enthalte Standardterme für alle $S_{\mathrm{Be}[v_0]}$-Formeln. Die Sprache sei interpretiert über einer S_{Be}-Struktur $\mathcal{A}_{\mathrm{Be}} = (U_{\mathrm{Be}}, j_{\mathrm{Be}})$, deren Universum Obermenge von $\mathbb{N}$ sei und alle $S_{\mathrm{Be}[v_0]}$-Formeln enthalte und deren Interpretation den Standardterm einer Formel auf sich selbst abbilde. $\mathcal{A}_{\mathrm{Be}}$ enthalte $\mathcal{A}_{\mathrm{N}}^{<}$, nach Einschränkung auf die Symbolmenge $S_{\mathrm{N}}^{<}$, als Substruktur.

Da über die natürlichen Zahlen und die $S_{\mathrm{Be}[v_0]}$-Formeln quantifiziert werden muss, um die relevante Definition $\psi(x)$ zu bilden, seien $S_{\mathrm{Be}[v_0]}$-Formeln $nat(x)$,

[20] Die Antinomie geht auf G. B. Berry zurück und wird in Russell (1908), S. 153 geschildert. Vgl. für das Folgende auch die Formalisierungen in Boolos (1989) und Serény (2004).

fm(x) vorausgesetzt, so dass für jede Variablenbelegung β gilt:

$$\begin{aligned}\mathcal{A}_{\mathrm{Be}}^{\beta} &\models nat(x) \quad \text{gdw.} \quad \beta(x) \in \mathbb{N},\\ \mathcal{A}_{\mathrm{Be}}^{\beta} &\models fm(x) \quad \text{gdw.} \quad \beta(x) \text{ ist eine } S_{\mathrm{Be}[v_0]}\text{-Formel.}\end{aligned}$$

Definitionen natürlicher Zahlen n werden aufgefasst als $S_{\mathrm{Be}[v_0]}$-Formeln $\varphi(x)$, mit $\mathcal{A}_{\mathrm{Be}} \models \bigwedge x(\varphi(x) \leftrightarrow x \equiv \boldsymbol{n})$. Die entscheidende Voraussetzung zur Bildung der Definition $\psi(x)$ ist eine $S_{\mathrm{Be}[v_0,v_1]}$-Formel $def(x,y)$, so dass für alle $S_{\mathrm{Be}[v_0]}$-Formeln $\varphi(x)$ und alle $n \in \mathbb{N}$:

$$\mathcal{A}_{\mathrm{Be}} \models \bigwedge x(\varphi(x) \leftrightarrow x \equiv \boldsymbol{n}) \quad \text{gdw.} \quad \mathcal{A}_{\mathrm{Be}} \models def(\ulcorner\varphi(x)\urcorner, \boldsymbol{n}).$$

Ferner muss über die Länge von Definitionen, also die Anzahl der in ihr enthaltenen Worte, gesprochen werden können. Es wird also ein $S_{\mathrm{Be}[v_0]}$-Term $l(x)$ vorausgesetzt, so dass für alle $S_{\mathrm{Be}[v_0]}$-Formeln $\varphi(x)$ der Länge n gilt: $\mathcal{A}_{\mathrm{Be}} \models l(\ulcorner\varphi(x)\urcorner) \equiv \boldsymbol{n}$. Jetzt kann die Definition $\psi(x)$ gebildet werden:

$$\begin{aligned}\psi(x) :=\,& nat(x) \wedge \neg\bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y,x))\\ & \wedge \bigwedge z(nat(z) \wedge z < x \rightarrow \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y,z))).\end{aligned}$$

Es gibt genau eine natürliche Zahl m, auf die $\psi(x)$ zutrifft. Es gilt also $\mathcal{A}_{\mathrm{Be}} \models \bigwedge x(\psi(x) \leftrightarrow x \equiv \boldsymbol{m})$. Das bedeutet aber, dass m mit weniger als 20 Worten definierbar ist, gleichzeitig jedoch, da $\psi(x)$ auf m zutrifft, nicht mit weniger als 20 Worten definierbar ist.[21] Damit ergibt sich der Widerspruch als:

$$\begin{aligned}\mathcal{A}_{\mathrm{Be}} &\models \neg\bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y,\boldsymbol{m})) \text{ und}\\ \mathcal{A}_{\mathrm{Be}} &\models \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y,\boldsymbol{m})).\end{aligned}$$

Die Antinomie von König

> Es werden Definitionen von Ordinalzahlen betrachtet. Da es nur abzählbar viele Definitionen gibt, gibt es genau eine Ordinalzahl α, auf die die Formel $\psi(x) :=$ "x ist kleinste Ordinalzahl, die nicht

[21] Es wird hier offen gelassen, was die Länge einer Definition der formalen Sprache ist. Damit die Länge von $\psi(x)$ kleiner 20 ist, müssen eventuell zusätzliche Zeichen in die Symbolmenge S_{Be} aufgenommen werden.

> definierbar ist" zutrifft. $\psi(x)$ stellt also eine Definition dar. Es ergibt sich also: Die Ordinalzahl α ist definierbar. Andererseits ist α nicht defnierbar, da $\psi(x)$ auf α zutrifft und dies besagt.[22]

Die Antinomie von König hat offensichtlich eine ähnliche Struktur wie die Antinomie von Berry. Sie wird darum nicht weiter gesondert betrachtet.

1.3 Der Satz von Tarski, Unvollständigkeit und Unentscheidbarkeit

Die Darstellung des Satzes von Tarski, des ersten Gödelschen Unvollständigkeitssatzes und des Unentscheidbarkeitssatzes, bzw. ihrer Beweise ist so gestaltet, dass nur die für die weitere Untersuchung wichtigen Teile der Beweise in aller Ausführlichkeit durchgeführt werden. Dies sind die Teile, die in ihrer Struktur den semantischen Antinomien, speziell der Lügner-Antinomie und den Antinomien von Grelling und Richard, entsprechen.

Parallel zur Darstellung der Antinomien wird auch hier ein Schwerpunkt auf die Unterscheidung von Objekt- und Metasprache gelegt. Die Verwendung eines Wahrheitsprädikats (bzw. Beweisbarkeitsprädikats) wird der Verwendung eines Falschheitsprädikats (bzw. Unbeweisbarkeitsprädikats) gegenübergestellt. Es wird besonders darauf geachtet, was jeweils in Bezug auf die objektsprachliche Negation vorausgesetzt wird. In Zusammenhang hiermit sind alle drei Sätze in mehreren Fassungen formuliert, die mathematisch gesehen in einigen Fällen äquivalent sind. Sie sind dennoch als eigene Sätze formuliert und bewiesen, da sie zu unterschiedlichen Sätzen führen, wenn man sie verallgemeinert. Dies wird im anschließenden Kapitel 2 über abstrakte Fassungen dieser Sätze deutlich.

Ziel des Abschnitts über diese drei Sätze ist die Herausarbeitung einer gemeinsamen Struktur ihrer Beweise, die anschließend in einer abstrakten Fassung mündet, unter die auch die Lügner-Antinomie und die Antinomien von Grelling und Richard fallen. Hierbei werden verschiedene Beweise dieser Sätze berücksichtigt. So werden Tarskis Beweis für den Satz von Tarski und sein Beweis des Unvollständigkeitssatzes, Gödels eigener semantischer und syntaktischer Beweis, der Beweis von Rosser für den Unvollständigkeitssatz und die Beweise von

[22] Vgl. König (1905).

Church und Rosser für den Unentscheidbarkeitssatz alle in ihrem Kern auf eine bestimmte gemeinsame Struktur zurückgeführt.[23]

1.3.1 Der Fixpunktsatz

Der Fixpunktsatz wird häufig in Darstellungen von Beweisen des Satzes von Tarski und des Unvollständigkeitssatzes verwendet.

S sei im Folgenden eine Symbolmenge, die die Symbolmenge S_{N} enthält, und Φ sei eine Menge von S-Sätzen.[24]

Definition 1.6. Eine (partielle) Funktion $f : \mathbb{N}^r \to \mathbb{N}$ heißt *repräsentierbar in* Φ gdw. es eine $S_{[v_0,...,v_r]}$-Formel $\varphi(x_0, ..., x_r)$ gibt, so dass für alle $n_0, ..., n_r \in \mathbb{N}$:

- $[(n_0, ..., n_{r-1}) \in f^{-1}[\mathbb{N}]$ und $f(n_0, ..., n_{r-1}) = n_r]$ gdw. $\Phi \vdash \varphi(\boldsymbol{n_0}, ..., \boldsymbol{n_r})$
- Wenn $(n_0, ..., n_{r-1}) \in f^{-1}[\mathbb{N}]$, dann $\Phi \vdash \dot{\bigvee} x_r \varphi(\boldsymbol{n_0}, ..., \boldsymbol{n_{r-1}}, x_r)$.

$\varphi(x_0, ..., x_r)$ *repräsentiert* f *in* Φ.
Φ *erlaubt Repräsentierungen bezüglich Funktionen* gdw. alle totalen berechenbaren Funktionen über $\mathbb{N}$ in Φ repräsentierbar sind.

Um die Schreibweise für repräsentierbare Funktionen zu vereinfachen, wird metasprachlich der ι-Operator eingeführt. Ist eine Funktion $f : \mathbb{N}^r \to \mathbb{N}$ in Φ repräsentierbar, dann gibt es in diesem Sinn einen $S_{[v_0,...,v_{r-1}]}$-*Term* $t(x_0, ..., x_{r-1})$, der f in Φ repräsentiert. Es gilt also:

$$[(n_0, ..., n_{r-1}) \in f^{-1}[\mathbb{N}] \text{ und } f(n_0, ..., n_{r-1}) = n_r] \quad \text{gdw.} \quad \Phi \vdash t(\boldsymbol{n_0}, ..., \boldsymbol{n_{r-1}}) \equiv \boldsymbol{n_r}.$$

Diese Termschreibweise wird als eine metasprachliche Abkürzung verstanden. $\varphi(x_0, ..., x_r)$ sei eine Formel, die eine Funktion repräsentiert. $n_0, ..., n_{r-1}$ liege im Definitionsbereich der Funktion. Statt "$\bigwedge x_r \varphi(\boldsymbol{n_0}, ..., \boldsymbol{n_{r-1}}, x_r) \to \psi(x_r)$" wird dann einfach "$\psi(t(\boldsymbol{n_0}, ..., \boldsymbol{n_{r-1}}))$" geschrieben.

[23]Einen Überblick über die dargestellten Beweise des Unvollständigkeits- und Unentscheidbarkeitssatzes bietet Rosser (1939).

[24]Die Symbolmenge wird möglichst allgemein gehalten. Da S_{N} in S enthalten ist, können Standardterme für die natürlichen Zahlen in der Sprache formuliert werden.

Satz 1.7 (Fixpunktsatz). [25] *Die Symbolmenge S sei höchstens abzählbar. $\ulcorner\cdot\urcorner : \{\varphi|\varphi$ ist S-Formel$\} \to \mathbb{N}$ sei injektiv. $f : \mathbb{N} \to \mathbb{N}$ sei gegeben durch*

$$f(n) = \begin{cases} \ulcorner\varphi(\boldsymbol{n})\urcorner, & \textit{falls } n = \ulcorner\varphi(x)\urcorner \textit{ für eine } S_{[v_0]}\textit{-Formel } \varphi(x) \\ 0, & \textit{sonst.} \end{cases}$$

f sei in Φ repräsentierbar. Dann gibt es zu jeder $S_{[v_0]}$-Formel $\psi(x)$ einen Fixpunkt, d. h. einen S-Satz χ, so dass $\Phi \vdash \chi \leftrightarrow \psi(\ulcorner\chi\urcorner)$.

Beweis. $\alpha(x)$ repräsentiere f in Φ.[26] Dann gilt für alle $S_{[v_0]}$-Formeln $\varphi(x)$: $\Phi \vdash \alpha(\ulcorner\varphi(x)\urcorner) \equiv \ulcorner\varphi(\ulcorner\varphi(x)\urcorner)\urcorner$. χ sei der S-Satz $\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))$. χ geht also aus $\psi(\alpha(x))$ durch Einsetzen des Standardterms für $\psi(\alpha(x))$ in $\psi(\alpha(x))$ hervor. Dann ist χ Fixpunkt; denn es gilt $\Phi \vdash \alpha(\ulcorner\psi(\alpha(x))\urcorner) \equiv \ulcorner\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))\urcorner$ und damit – nach Einsetzen von $\alpha(\ulcorner\psi(\alpha(x))\urcorner)$ und $\ulcorner\chi\urcorner = \ulcorner\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))\urcorner$ in $\psi(x)$ – $\Phi \vdash \psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner)) \leftrightarrow \psi(\ulcorner\chi\urcorner)$, woraus nach der Definition von χ die Fixpunkteigenschaft folgt: $\Phi \vdash \chi \leftrightarrow \psi(\ulcorner\chi\urcorner)$. □

S sei im Folgenden höchstens abzählbar und $\ulcorner\cdot\urcorner : \{\varphi|\varphi$ ist S-Formel$\} \to \mathbb{N}$ sei eine effektive Gödelisierung.[27]

Satz 1.8. Φ *erlaube Repräsentierungen bezüglich Funktionen. Dann gibt es zu jeder $S_{[v_0]}$-Formel $\psi(x)$ einen Fixpunkt, d. h. einen S-Satz χ, so dass $\Phi \vdash \chi \leftrightarrow \psi(\ulcorner\chi\urcorner)$.*

Beweis. $f : \mathbb{N} \to \mathbb{N}$ sei gegeben durch

$$f(n) = \begin{cases} \ulcorner\varphi(\boldsymbol{n})\urcorner, & \text{falls } n = \ulcorner\varphi(x)\urcorner \text{ für eine } S_{[v_0]}\text{-Formel } \varphi(x) \\ 0, & \text{sonst.} \end{cases}$$

f ist berechenbar und damit in Φ repräsentierbar. Aus dem Fixpunktsatz folgt die Behauptung. □

[25] Der Fixpunktsatz findet sich in dieser Form z. B. in Ebbinghaus u. a. (1992), S. 212. Explizit wurde der Fixpunktsatz erstmals von Carnap (Carnap (1934), S. 90-92) formuliert und mit Hilfe von Gödels unentscheidbarem Satz bewiesen. Vgl. Kahle (2007), S. 6 und Kleene (1986), S. 134 Fußnote h.

[26] Es wird nur die eine Richtung der Äquivalenz dieser Voraussetzung verwendet: Wenn $f(n_0, ..., n_{r-1}) = n_r$, dann $\Phi \vdash t(\boldsymbol{n_0}, ..., \boldsymbol{n_{r-1}}) \equiv \boldsymbol{n_r}$.

[27] Die Abbildung von der Menge der Formeln in die natürlichen Zahlen ist injektiv, berechenbar und die Bildmenge entscheidbar.

1.3.2 Der Satz von Tarski

Definition 1.9. a) Eine Relation $R \subseteq \mathbb{N}^{r+1}$ heißt *repräsentierbar in* Φ gdw. es eine $S_{[v_0,...,v_r]}$-Formel $\varphi(x_0,...,x_r)$ gibt, so dass für alle $n_0,...,n_r \in \mathbb{N}$:

$$(n_0,...,n_r) \in R \quad \text{gdw.} \quad \Phi \vdash \varphi(\boldsymbol{n_0},...,\boldsymbol{n_r}).$$

b) Eine Relation $R \subseteq \mathbb{N}^{r+1}$ heißt *stark repräsentierbar in*[28] Φ gdw. es eine $S_{[v_0,...,v_r]}$-Formel $\varphi(x_0,...,x_r)$ gibt, s. d. für alle $n_0,...,n_r \in \mathbb{N}$:[29]

$$(n_0,...,n_r) \in R \quad \text{gdw.} \quad \Phi \vdash \varphi(\boldsymbol{n_0},...,\boldsymbol{n_r}),$$
$$(n_0,...,n_r) \notin R \quad \text{gdw.} \quad \Phi \vdash \neg\varphi(\boldsymbol{n_0},...,\boldsymbol{n_r}).$$

$\varphi(x_0,...,x_r)$ *repräsentiert* R *in* Φ (*stark*).
Φ *erlaubt Repräsentierungen bezüglich Relationen* gdw. alle entscheidbaren Relationen über $\mathbb{N}$ in Φ stark repräsentierbar sind.

Die Symbolmenge S_{Ar} sei definiert als $S_{\mathrm{Ar}} := \{0, 1, +, \cdot\}$, wobei 0,1 Objektkonstanten und $+, \cdot$ zweistellige Funktionskonstanten seien. $\mathcal{N}$ sei die S_{Ar}-Struktur $(\mathbb{N}, 0^{\mathbb{N}}, 1^{\mathbb{N}}, +^{\mathbb{N}}, \cdot^{\mathbb{N}})$. Für $n \in \mathbb{N}$ sei $\boldsymbol{n}$ der Standardterm für n aus den S_{Ar}-Termen.

Satz 1.10. Th($\mathcal{N}$) *erlaubt Repräsentierungen (bezüglich Funktionen und Relationen).*

Satz 1.11 (Satz von Tarski). *Es gibt keine $S_{\mathrm{Ar}[v_0]}$-Formel $w(x)$, so dass für alle S_{Ar}-Sätze φ gilt*

$$\mathcal{N} \models \varphi \leftrightarrow w(\ulcorner\varphi\urcorner).$$

Beweis. $w(x)$ sei eine $S_{\mathrm{Ar}[v_0]}$-Formel, so dass für alle S_{Ar}-Sätze φ gilt $\mathcal{N} \models \varphi \leftrightarrow w(\ulcorner\varphi\urcorner)$. Dann gibt es einen Fixpunkt χ der Negation $\neg w(x)$ von $w(x)$; es

[28] Eine Übertragung des Begriffs der starken Repräsentierbarkeit auf Funktionen könnte erreicht werden, indem zu den Bedingungen der Repräsentierbarkeit von Funktionen hinzugefügt wird, dass die Menge der natürlichen Zahlen, auf denen die partielle Funktion nicht definiert ist, repräsentierbar ist. Für totale Funktionen fällt damit der Begriff der starken Repräsentierbarkeit mit dem der einfachen Repräsentierbarkeit zusammen.

[29] Die Richtung "Wenn $\Phi \vdash \neg\varphi(\boldsymbol{n_0},...,\boldsymbol{n_r})$, dann $(n_0,...,n_r) \notin R$." folgt, falls Φ widerspruchsfrei ist, schon aus der ersten Bedingung.

gilt also $\mathcal{N} \models \chi \leftrightarrow \neg w(\ulcorner \chi \urcorner)$. Nach Einsetzen von χ für φ ergibt sich

$$\mathcal{N} \models \neg w(\ulcorner \chi \urcorner) \leftrightarrow w(\ulcorner \chi \urcorner) \text{ und damit}$$
$$\mathcal{N} \models w(\ulcorner \chi \urcorner) \wedge \neg w(\ulcorner \chi \urcorner)$$

Widerspruch. □

Die hier angegebene Formulierung des Satzes und des Beweises entspricht in ihren wesentlichen Schritten der ursprünglichen Fassung Tarskis.[30] In dieser wird gezeigt, dass ein adäquater Wahrheitsbegriff für eine Objektsprache nicht in einer Metasprache definiert werden kann, die die gleiche Stufe wie die Objektsprache hat. Tarski bezieht sich an dieser Stelle auf eine Sprache höherer Stufe ohne Konstanten, die über einem unendlichen Universum interpretiert ist.[31]

Später zeigt Tarski, dass sich in einer erweiterten Sprache ein Wahrheitsbegriff für die (nicht erweiterte) Sprache gemäß der Wahrheitskonvention definieren lässt.[32] Auf dieser Basis kann die folgende Form des Satzes formuliert werden. Sie ist eine metasprachliche Variante.

Satz 1.12. *Die Menge der Gödelnummern von Sätzen aus* $\mathrm{Th}(\mathcal{N})$ *ist nicht repräsentierbar in* $\mathrm{Th}(\mathcal{N})$. *Genauer gibt es sogar kein Wahrheitsprädikat für* $\mathcal{N}$, *d. h. es gibt keine* $S_{\mathrm{Ar}[v_0]}$*-Formel* $w(x)$, *so dass für alle* S_{Ar}*-Sätze* φ *gilt:*

$$\mathcal{N} \models \varphi \quad \textit{gdw.} \quad \mathcal{N} \models w(\ulcorner \varphi \urcorner).$$

[30] Siehe Tarski (1935), S. 522 ff. Tarski verweist an dieser Stelle in Fußnote 88 und im Nachwort, S. 543 ff. darauf, dass er die Methode des Beweises, insbesondere der Konstruktion des Fixpunktes, dem Beweis Gödels für dessen ersten Unvollständigkeitssatz entnommen hat. Ein Unterschied zwischen Tarskis Originalfassung und der hier gegebenen ist, dass statt der direkten Selbsteinsetzung des Standardterms einer $S_{\mathrm{Ar}[v_0]}$-Formel in sich, also der Funktion f im Fixpunktsatz, eine etwas andere Funktion gewählt wird, die für den hier interessanten ersten Fall so wiedergegeben werden kann (andere Fälle werden hier einfach auf 0 abgebildet):

$$f'(n) := \begin{cases} \ulcorner \bigvee v_0(v_0 \equiv \ulcorner \varphi(x) \urcorner \wedge \varphi(v_0)) \urcorner, & \text{falls } n = \ulcorner \varphi(x) \urcorner \text{ für eine } S_{[v_0]}\text{-Formel } \varphi(x) \\ 0, & \text{sonst.} \end{cases}$$

Der Fixpunktsatz kann jedoch fast analog bewiesen werden. Der $S_{\mathrm{Ar}[v_0]}$-Term $\alpha'(x)$ repräsentiere f' in $\mathcal{N}$. Dann ergibt sich als Fixpunkt der S_{Ar}-Satz $\bigvee v_0(v_0 \equiv \ulcorner \psi(\alpha'(x)) \urcorner \wedge \psi(\alpha'(v_0)))$, denn es gilt
$\mathcal{N} \models \bigvee v_0(v_0 \equiv \ulcorner \psi(\alpha'(x)) \urcorner \wedge \psi(\alpha'(v_0))) \leftrightarrow \psi(\alpha'(\ulcorner \psi(\alpha'(x)) \urcorner))$ und
$\mathcal{N} \models \psi(\alpha'(\ulcorner \psi(\alpha'(x)) \urcorner)) \leftrightarrow \psi(\ulcorner \bigvee v_0(v_0 \equiv \ulcorner \psi(\alpha'(x)) \urcorner \wedge \psi(\alpha'(v_0))) \urcorner)$.

[31] Zur Analyse der Form des ursprünglichen Beweises vgl. Abschnitt 2.1 der vorliegenden Arbeit.

[32] Tarski (1935), S. 542.

Der Satz folgt sofort aus Satz 1.11, da von einem klassischen Logiksystem ausgegangen wird. Da später in der abstrakten Form in Satz 2.10 ein klassisches Logiksystem nicht mehr vorausgesetzt wird, wird hier der Übersicht halber der Beweis noch einmal geführt:

Beweis. $w(x)$ sei eine $S_{\mathrm{Ar}[v_0]}$-Formel, so dass für alle S_{Ar}-Sätze φ gilt: $\mathcal{N} \models \varphi$ gdw. $\mathcal{N} \models w(\ulcorner\varphi\urcorner)$. Nach dem Fixpunktsatz gibt es einen Fixpunkt χ der Negation $\neg w(x)$ von $w(x)$; es gilt also $\mathcal{N} \models \chi \leftrightarrow \neg w(\ulcorner\chi\urcorner)$. Damit gilt auch: $\mathcal{N} \models \chi$ gdw. $\mathcal{N} \models \neg w(\ulcorner\chi\urcorner)$. Nach Einsetzen von χ für φ ergibt sich:

$$\begin{array}{rcl} \mathcal{N} \models \neg w(\ulcorner\chi\urcorner) & \text{gdw.} & \mathcal{N} \models w(\ulcorner\chi\urcorner), \text{ und damit:} \\ \mathcal{N} \models w(\ulcorner\chi\urcorner) & \text{und} & \mathcal{N} \models \neg w(\ulcorner\chi\urcorner). \end{array}$$

Widerspruch. □

Als Parallele zu den Antinomien von Grelling und Richard und im Hinblick auf die abstrakte Formulierung in Satz 2.10 werden hier zwei weitere Varianten des Satzes von Tarski mit Beweis dargestellt. Ausgangspunkt ist hier nicht mehr ein objektsprachlicher Wahrheitsbegriff bzw. ein Wahrheitsprädikat, sondern ein objektsprachlicher Falschheitsbegriff bzw. ein Falschheitsprädikat. In diesen Varianten werden Äquivalenz und Negation einmal objektsprachlich und einmal metasprachlich wiedergegeben. Die eine Variante bezieht sich also auf die objektsprachliche Negation von Sätzen und die andere auf die Nicht-Wahrheit von Sätzen. Beide Teile folgen sofort aus dem Satz von Tarski, da hier von einem klassischen Logiksystem ausgegangen wird, Äquivalenz und Negation also denen der Metasprache entsprechen.

Satz 1.13. *a) Es gibt keine $S_{\mathrm{Ar}[v_0]}$-Formel $f(x)$, so dass für alle S_{Ar}-Sätze φ gilt*

$$\mathcal{N} \models \neg\varphi \leftrightarrow f(\ulcorner\varphi\urcorner).$$

b) Die Menge der Gödelnummern von Sätzen aus $\{\varphi | \mathcal{N} \not\models \varphi\}$ ist nicht repräsentierbar in $\mathrm{Th}(\mathcal{N})$. Genauer gibt es sogar keine $S_{\mathrm{Ar}[v_0]}$-Formel $f(x)$, so dass für alle S_{Ar}-Sätze φ gilt:

$$\mathcal{N} \not\models \varphi \quad \textit{gdw.} \quad \mathcal{N} \models f(\ulcorner\varphi\urcorner).$$

Beweis. Zu a) $f(x)$ sei eine $S_{\mathrm{Ar}[v_0]}$-Formel, so dass für alle S_{Ar}-Sätze φ gilt $\mathcal{N} \models \neg\varphi \leftrightarrow f(\ulcorner\varphi\urcorner)$. Dann gibt es einen Fixpunkt χ von $f(x)$; es gilt also $\mathcal{N} \models \chi \leftrightarrow f(\ulcorner\chi\urcorner)$ und $\mathcal{N} \models \neg\chi \leftrightarrow \neg f(\ulcorner\chi\urcorner)$. Nach Einsetzen von χ für φ ergibt sich

$$\begin{aligned}&\mathcal{N} \models f(\ulcorner\chi\urcorner) \leftrightarrow \neg f(\ulcorner\chi\urcorner) \text{ und damit}\\ &\mathcal{N} \models f(\ulcorner\chi\urcorner) \wedge \neg f(\ulcorner\chi\urcorner)\end{aligned}$$

Widerspruch.
Zu b) $f(x)$ sei eine $S_{\mathrm{Ar}[v_0]}$-Formel, so dass für alle S_{Ar}-Sätze φ gilt: $\mathcal{N} \not\models \varphi$ gdw. $\mathcal{N} \models f(\ulcorner\varphi\urcorner)$. Dann gibt es einen Fixpunkt χ von $f(x)$, es gilt also $\mathcal{N} \models \chi \leftrightarrow f(\ulcorner\chi\urcorner)$. Damit gilt auch: $\mathcal{N} \not\models \chi$ gdw. $\mathcal{N} \not\models f(\ulcorner\chi\urcorner)$. Nach Einsetzen von χ für φ ergibt sich:

$$\begin{aligned}&\mathcal{N} \models f(\ulcorner\chi\urcorner) \quad \text{gdw.} \quad \mathcal{N} \not\models f(\ulcorner\chi\urcorner), \text{ und damit:}\\ &\mathcal{N} \models f(\ulcorner\chi\urcorner) \quad \text{und} \quad \mathcal{N} \not\models f(\ulcorner\chi\urcorner).\end{aligned}$$

Widerspruch. □

Allgemeiner lässt sich Satz 1.12 wie im folgenden Satz formulieren. Eine analoge Verallgemeinerung ist natürlich auch für die drei genannten Varianten möglich.

Satz 1.14. *S sei eine höchstens abzählbare Symbolmenge, die die Symbolmenge S_{N} enthält, und $\mathcal{A}$ sei eine S-Struktur.* $\mathrm{Th}(\mathcal{A})$ *erlaube Repräsentierungen bezüglich Funktionen. Dann ist die Menge der Gödelnummern von Sätzen aus* $\mathrm{Th}(\mathcal{A})$ *nicht repräsentierbar in* $\mathrm{Th}(\mathcal{A})$. *Genauer gibt es sogar kein Wahrheitsprädikat für $\mathcal{A}$, d. h. es gibt keine $S_{[v_0]}$-Formel $w(x)$, so dass für alle S-Sätze φ gilt: $\mathcal{A} \models \varphi$ gdw. $\mathcal{A} \models w(\ulcorner\varphi\urcorner)$.*

Beweis. Siehe den Beweis zu Satz 1.11. □

Der Satz von Tarski beantwortet grob gesagt die Frage nach der Ausdrückbarkeit der semantischen Zuordnung des Wahrheitswertes eines Satzes zu diesem Satz als Prädikat der Sprache. Und zwar beantwortet er sie negativ. Ein entsprechender Satz, der sich auf die Ausdrückbarkeit der semantischen Zuordnung eines von einem Term bezeichneten Objekts des Universums zu diesem Term als einstelliger Term der Sprache bezieht und der wiederum diese Ausdrückbarkeit als nicht möglich darstellt, ist der nun folgende Satz. Er wird dem Satz von Tarski entsprechend in vier Versionen dargestellt. Auch für die Antinomien können analog "Term"-Versionen formuliert werden. Dies geschieht in Abschnitt 2.4. Zunächst folgt ein dem Fixpunktsatz entsprechendes Lemma. Alle Sätze werden der Kürze halber nur für die Symbolmenge S_{Ar} und die S_{Ar}-Struktur $\mathcal{N}$ angegeben.[33]

Lemma 1.15. [34] *Zu jedem $S_{\mathrm{Ar}[v_0]}$-Term $t(x)$ gibt es einen Fixpunkt, d. h. einen S_{Ar}-Term u mit* $\mathrm{Th}(\mathcal{N}) \vdash u \equiv t(\ulcorner u \urcorner)$.

Beweis. Da $\ulcorner \cdot \urcorner$ eine effektive Gödelisierung ist, ist die Abbildung $f : \mathbb{N} \to \mathbb{N}$, die gegeben ist durch

$$f(n) = \begin{cases} \ulcorner t(\boldsymbol{n}) \urcorner, & \text{falls } n = \ulcorner t(x) \urcorner \text{ für einen } S_{\mathrm{Ar}[v_0]}\text{-Term } t(x) \\ 0, & \text{sonst,} \end{cases}$$

berechenbar. Da $\mathrm{Th}(\mathcal{N})$ Repräsentierungen erlaubt, gibt es einen $S_{\mathrm{Ar}[v_0]}$-Term $\alpha(x)$, der f repräsentiert. Dann gilt für alle $S_{\mathrm{Ar}[v_0]}$-Terme $t(x)$: $\mathrm{Th}(\mathcal{N}) \vdash \alpha(\ulcorner t(x) \urcorner) \equiv \ulcorner t(\ulcorner t \urcorner) \urcorner$. Ein Fixpunkt der gesuchten Art ist der S_{Ar}-Term $t(\alpha(\ulcorner t(\alpha(x)) \urcorner))$. Denn es gilt $\mathrm{Th}(\mathcal{N}) \vdash \alpha(\ulcorner t(\alpha(x)) \urcorner) \equiv \ulcorner t(\alpha(\ulcorner t(\alpha(x)) \urcorner)) \urcorner$, und damit $\mathrm{Th}(\mathcal{N}) \vdash t(\alpha(\ulcorner t(\alpha(x)) \urcorner)) \equiv t(\ulcorner t(\alpha(\ulcorner t(\alpha(x)) \urcorner)) \urcorner)$. ☐

Satz 1.16. *a) Es gibt keinen $S_{\mathrm{Ar}[v_0]}$-Term $s(x)$, so dass für alle S_{Ar}-Terme t gilt*

$$\mathcal{N} \models t \equiv s(\ulcorner t \urcorner).$$

b) Die partielle Funktion $f : \mathbb{N} \to \mathbb{N}$, die gegeben ist durch

$$f(n) = j_{\mathcal{N}}(t), \text{ falls } n = \ulcorner t \urcorner \text{ für einen } S_{\mathrm{Ar}}\text{-Term } t$$

ist nicht repräsentierbar in $\mathrm{Th}(\mathcal{N})$. *Genauer gibt es sogar keinen Interpretationsterm für $\mathcal{N}$, d. h. es gibt keinen $S_{\mathrm{Ar}[v_0]}$-Term $s(x)$, so dass für alle*

[33] Zum Satz von Tarski für Terme vgl. Hilbert u. Bernays (1939), S. 263-278.

[34] Das Lemma lässt sich folgendermaßen ohne Termschreibweise der repräsentierenden Formel formulieren und beweisen: Zu jeder $S_{\mathrm{Ar}[v_0,v_1]}$-Formel $\varphi(x_0, x_1)$, so dass $\mathrm{Th}(\mathcal{N}) \vdash \ddot{\bigvee} x_1 \varphi(\boldsymbol{n}, x_1)$ für alle $n \in \mathbb{N}$, gibt es einen Fixpunkt, d. h. eine $S_{\mathrm{Ar}[v_1]}$-Formel $\chi(x)$ mit $\mathrm{Th}(\mathcal{N}) \vdash \ddot{\bigvee} x \chi(x)$ und $\mathrm{Th}(\mathcal{N}) \vdash \bigwedge x : \chi(x) \to \varphi(\ulcorner \chi(x) \urcorner, x)$. Beweis: $\alpha(x_0, x_1)$ repräsentiere die Abbildung f, übertragen auf die Selbsteinsetzung von $S_{\mathrm{Ar}[v_0,v_1]}$-Formeln. Dann ist $\bigwedge z : \alpha(\ulcorner \bigwedge z \alpha(x_0, z) \to \varphi(z, x_1) \urcorner, z) \to \varphi(z, x_1)$ Fixpunkt; denn es gilt $\mathrm{Th}(\mathcal{N}) \vdash \bigwedge z : \alpha(\ulcorner \bigwedge z \alpha(x_0, z) \to \varphi(z, x_1) \urcorner, z) \leftrightarrow z \equiv \ulcorner \bigwedge z : \alpha(\ulcorner \bigwedge z \alpha(x_0, z) \to \varphi(z, x_1) \urcorner, z) \to \varphi(z, x_1) \urcorner$.

S_{Ar}-Terme t gilt:

$$\textit{Für alle } q \in \mathbb{N}: \quad \mathcal{N} \models t \equiv \boldsymbol{q} \textit{ gdw. } \mathcal{N} \models s(\ulcorner t \urcorner) \equiv \boldsymbol{q}.$$

c) Es gibt keinen $S_{\mathrm{Ar}[v_0]}$-Term $r(x)$, so dass für alle S_{Ar}-Terme t gilt

$$\mathcal{N} \models t + 1 \equiv r(\ulcorner t \urcorner).$$

d) Die Funktion $f : \mathbb{N} \to \mathbb{N}$, die gegeben ist durch

$$f(n) = j_{\mathcal{N}}(t) + 1, \textit{ falls } n = \ulcorner t \urcorner \textit{ für einen } S_{\mathrm{Ar}}\textit{-Term } t,$$

ist nicht repräsentierbar in Th($\mathcal{N}$). *Genauer gibt es sogar keinen $S_{\mathrm{Ar}[v_0]}$-Term $r(x)$, so dass für alle S_{Ar}-Terme t gilt:*[35]

$$\textit{Für alle } q \in \mathbb{N}: \quad \mathcal{N} \models t \equiv \boldsymbol{q} \textit{ gdw. } \mathcal{N} \models r(\ulcorner t \urcorner) \equiv \boldsymbol{q+1}.$$

Beweis. Zu a) $s(x)$ sei ein solcher $S_{\mathrm{Ar}[v_0]}$-Term, so dass für alle S_{Ar}-Terme t gilt $\mathcal{N} \models t \equiv s(\ulcorner t \urcorner)$. $s^+(x)$ sei der $S_{\mathrm{Ar}[v_0]}$-Term $s(x) + 1$. Dann gibt es einen Fixpunkt u von $s^+(x)$; es gilt also $\mathcal{N} \models u \equiv s^+(\ulcorner u \urcorner)$. Nach Einsetzen von u für t ergibt sich

$$\mathcal{N} \models s^+(\ulcorner u \urcorner) \equiv s(\ulcorner u \urcorner).$$

Widerspruch.

Zu b) $s(x)$ sei ein solcher $S_{\mathrm{Ar}[v_0]}$-Term, so dass für alle S_{Ar}-Terme t gilt: Für alle $q \in \mathbb{N}$: $\mathcal{N} \models t \equiv \boldsymbol{q}$ gdw. $\mathcal{N} \models s(\ulcorner t \urcorner) \equiv \boldsymbol{q}$. $s^+(x)$ sei der $S_{\mathrm{Ar}[v_0]}$-Term $s(x)+1$. Dann gibt einen Fixpunkt u von $s^+(x)$; es gilt also $\mathcal{N} \models u \equiv s^+(\ulcorner u \urcorner)$. Damit gilt auch: Für alle $q \in \mathbb{N}$: $\mathcal{N} \models u \equiv \boldsymbol{q}$ gdw. $\mathcal{N} \models s^+(\ulcorner u \urcorner) \equiv \boldsymbol{q}$. Nach Einsetzen von u für t ergibt sich:

$$\text{Für alle } q \in \mathbb{N}: \quad \mathcal{N} \models s^+(\ulcorner u \urcorner) \equiv \boldsymbol{q} \text{ gdw. } \mathcal{N} \models s(\ulcorner u \urcorner) \equiv \boldsymbol{q}.$$

Widerspruch.

Zu c) $r(x)$ sei ein solcher $S_{\mathrm{Ar}[v_0]}$-Term, so dass für alle S_{Ar}-Terme t gilt $\mathcal{N} \models t + 1 \equiv r(\ulcorner t \urcorner)$. Dann gibt es einen Fixpunkt u von $r(x)$; es gilt also $\mathcal{N} \models u \equiv$

[35] $\boldsymbol{q+1}$ ist der Standardterm für $q+1$ aus den S_{N}-Termen.

$r(\ulcorner u \urcorner)$ und $\mathcal{N} \models u + 1 \equiv r(\ulcorner u \urcorner) + 1$. Nach Einsetzen von u für t ergibt sich

$$\mathcal{N} \models r(\ulcorner u \urcorner) \equiv r(\ulcorner u \urcorner) + 1.$$

Widerspruch.

Zu d) r sei ein solcher $S_{\mathrm{Ar}[v_0]}$-Term, so dass für alle S_{Ar}-Terme t gilt: Für alle $q \in \mathbb{N}$: $\mathcal{N} \models t \equiv \boldsymbol{q+1}$ gdw. $\mathcal{N} \models r(\ulcorner t \urcorner) \equiv \boldsymbol{q}$. Dann gibt es einen Fixpunkt u von $r(x)$; es gilt also $\mathcal{N} \models u \equiv r(\ulcorner u \urcorner)$. Damit gilt auch: Für alle $q \in \mathbb{N}$: $\mathcal{N} \models u \equiv \boldsymbol{q}$ gdw. $\mathcal{N} \models r(\ulcorner u \urcorner) \equiv \boldsymbol{q}$. Nach Einsetzen von u für t ergibt sich:

$$\text{Für alle } q \in \mathbb{N}: \quad \mathcal{N} \models r(\ulcorner u \urcorner) \equiv \boldsymbol{q+1} \text{ gdw. } \mathcal{N} \models r(\ulcorner u \urcorner) \equiv \boldsymbol{q}.$$

Widerspruch. □

1.3.3 Die Gödelschen Unvollständigkeitssätze

Zunächst wird die Form des ersten Unvollständigkeitssatzes hergeleitet, die Tarski[36] entwickelt. Diese Version des Unvollständigkeitssatzes kann als semantisch bezeichnet werden, da sie auf dem semantischen Begriff der Wahrheit in einer Struktur aufbaut. In der vorliegenden Darstellung liegt der Schwerpunkt darauf, zu zeigen, in welcher Art diese Form des Unvollständigkeitssatzes und Tarskis Beweis hierfür auf dem Satz von Tarski als Grundgerüst beruhen.

Im Anschluss wird die Form des ersten Unvollständigkeitssatzes hergeleitet, die Gödel[37] selbst entwickelt. Genauer geht es in diesem Abschnitt um die rein syntaktische Form des Satzes, die in der zitierten Arbeit von Gödel ausführlich dargestellt wird.[38] In derselben Arbeit gibt Gödel auch eine semantische Form des Unvollständigkeitssatzes an, die er informell zu Beginn beweist.[39] Auf diese Form und Gödels Beweis hierfür wird erst in Abschnitt 2.1 eingegangen. Als Beweis für die syntaktische Form des Unvollständigkeitssatzes wird ein Beweis angegeben, der den Fixpunktsatz verwendet. Auf Gödels Originalbeweis, der sich nicht explizit auf eine Fixpunktkonstellation bezieht wie sie im Fixpunktsatz der vorliegenden Arbeit auftritt, wird ebenfalls erst in Abschnitt 2.1 näher

[36] Vgl. Tarski (1935).
[37] Vgl. Gödel (1931).
[38] Gödel (1931), S. 187 ff.
[39] Gödel (1931), S. 174 f.

eingegangen. Es folgt der auf Gödels Satz aufbauende Unvollständigkeitssatz von Rosser. Ziel ist es, zu zeigen, in welcher Art diese Sätze bzw. ihre Beweise auf einem Grundgerüst beruhen, das eine syntaktische Parallele zum Satz von Tarski bzw. Satz 1.12 ist.

Die Beweise werden jeweils nur so weit dargestellt, wie dies für dieses Ziel erforderlich ist. Von den Eigenschaften des Grundgerüsts wird dann im späteren Abschnitt 2.1 so abstrahiert, dass auch die Lügner-Antinomie, die Antinomien von Grelling und Richard und die Beweise des Unentscheidbarkeitssatzes von der Abstraktion umfasst werden.

Lemma 1.17. *Zu jeder aufzählbaren Menge $M \subseteq \mathbb{N}$ gibt es eine $S_{\mathrm{Ar}[v_0,v_1]}$-Formel $\varphi(x,y)$, s. d. für alle $n \in \mathbb{N}$: $n \in M$ gdw. $\mathcal{N} \models \bigvee x\varphi(x,\boldsymbol{n})$.*

Beweis. Zu jeder aufzählbaren Menge $M \subseteq \mathbb{N}$ und jedem Aufzählungsverfahren ist die zweistellige Relation R_M über $\mathbb{N}$, die gegeben ist durch

$$R_M := \{(n_1, n_2) | n_2 \text{ steht an der Stelle } n_1 \text{ der Aufzählung von } M\},$$

entscheidbar. Nach Satz 1.10 gibt es also eine $S_{\mathrm{Ar}[v_0,v_1]}$-Formel $\varphi(x,y)$, die R_M in $\mathrm{Th}(\mathcal{N})$ repräsentiert. Damit gilt für alle $n \in \mathbb{N}$: $n \in M$ gdw. $\mathcal{N} \models \bigvee x\varphi(x,\boldsymbol{n})$. □

S sei im Folgenden eine höchstens abzählbare Symbolmenge, die die Symbolmenge S_{N} enthält, und $\mathcal{A}$ sei eine S-Struktur.

Definition 1.18. Eine Menge von S- Sätzen Φ heißt *ω-widerspruchsfrei* gdw. für jede $S_{[v_0]}$-Formel $\varphi(x)$ gilt: Wenn $\Phi \vdash \neg\varphi(\boldsymbol{n})$ für alle $n \in \mathbb{N}$, dann $\Phi \nvdash \bigvee x\varphi(x)$.

Lemma 1.19. $\mathrm{Th}(\mathcal{A})$ *sei ω-widerspruchsfrei und erlaube Repräsentierungen. Dann gibt es zu jeder aufzählbaren Menge M natürlicher Zahlen eine $S_{[v_0,v_1]}$-Formel $\varphi(x,y)$, s. d. für alle $n \in \mathbb{N}$: $n \in M$ gdw. $\mathcal{A} \models \bigvee x\varphi(x,\boldsymbol{n})$.*

Beweis. Siehe den Beweis zu Lemma 1.17. □

Um nach dem Satz von Tarski (in der metasprachlichen Variante in Satz 1.12) bzw. Satz 1.14 zu einem Widerspruch zu kommen, genügt also schon die Aufzählbarkeit von $\mathrm{Th}(\mathcal{N})$ bzw. $\mathrm{Th}(\mathcal{A})$ (die zu ihrer Entscheidbarkeit äquivalent ist[40]). Damit sind $\mathrm{Th}(\mathcal{N})$ und $\mathrm{Th}(\mathcal{A})$ als Theorien erster Stufe nicht axiomatisierbar (aus der semantischen Vollständigkeit von $\vdash$ ergäbe sich die Aufzählbarkeit von $\mathrm{Th}(\mathcal{N})$ bzw. $\mathrm{Th}(\mathcal{A})$). Als Theorie höherer Stufe ist $\mathrm{Th}(\mathcal{N})$ axiomatisierbar, woraus folgt, dass die Logik höherer Stufe semantisch unvollständig ist. Unabhängig von der Stufe lässt sich dies auch folgendermaßen ausdrücken: Für jede aufzählbare Menge von S_{Ar}- bzw. S-Sätzen Φ gilt: Wenn $\Phi^{\vdash} \subseteq \mathrm{Th}(\mathcal{N})$ bzw. $\Phi^{\vdash} \subseteq \mathrm{Th}(\mathcal{A})$, dann $\Phi^{\vdash} \neq \mathrm{Th}(\mathcal{N})$ bzw. $\Phi^{\vdash} \neq \mathrm{Th}(\mathcal{A})$. Eine aufzählbare Menge von in $\mathcal{N}$ bzw. $\mathcal{A}$ wahren S_{Ar}- bzw. S-Sätzen ist also nie syntaktisch vollständig. Dies ist die Folgerung, die in Tarski (1935), S. 527 f. gezogen wird.

Nach dem Satz von Tarski (Satz 1.12) bzw. Satz 1.14 gibt es zwar kein Wahrheitsprädikat für $\mathcal{N}$ bzw. $\mathcal{A}$, aus Lemma 1.17 ergibt sich jedoch, dass es ein Ableitbarkeitsprädikat für jede aufzählbare Menge von S'-Sätzen einer abzählbaren Symbolmenge S' gibt. Dies wird in Lemma 1.20 dargestellt. Wendet man nun den Beweis des Satzes von Tarski (Satz 1.12) bzw. den Beweis von Satz 1.14 auf ein solches Ableitbarkeitsprädikat, für eine aufzählbare Menge Φ von in $\mathcal{N}$ bzw. $\mathcal{A}$ wahren S_{Ar}- bzw. S-Sätzen, als potenzielles Wahrheitsprädikat an, erhält man speziell den Satz $\neg\bigvee x \mathrm{bew}_\Phi(x, \alpha(\ulcorner\neg\bigvee x \mathrm{bew}_\Phi(x, \alpha(y))\urcorner))$ als nicht aus Φ ableitbaren, jedoch in $\mathcal{N}$ bzw. $\mathcal{A}$ wahren Satz. Hierbei sei $\mathrm{bew}_\Phi(x, y)$ eine S_{Ar}- bzw. S-Formel, die die Relation "n_2 steht an der Stelle n_1 einer Aufzählung von $\Phi^{\vdash}$" in $\mathrm{Th}(\mathcal{N})$ bzw. $\mathrm{Th}(\mathcal{A})$ repräsentiert. Dies wird in Satz 1.21 bzw. Satz 1.22 dargestellt. Satz 1.21 bzw. Satz 1.22 entsprechen der Version des ersten Gödelschen Unvollständigkeitssatzes, die Tarski im Nachwort zu Tarski (1935), S. 543 f. beweist.

Lemma 1.20. Φ *sei eine aufzählbare Menge von S'-Sätzen. Dann gibt es ein Ableitbarkeitsprädikat* $\bigvee x \mathrm{bew}_\Phi(x, y)$ *mit einer* $S_{\mathrm{Ar}[v_0,v_1]}$*-Formel* $\mathrm{bew}_\Phi(x, y)$*, d. h. es gilt für jeden S'-Satz φ: φ ist aus Φ ableitbar gdw.* $\mathcal{N} \models \bigvee x \mathrm{bew}_\Phi(x, \ulcorner\varphi\urcorner)$*.*

Beweis. Da Φ aufzählbar ist, ist die Menge der aus Φ ableitbaren S'-Sätze aufzählbar. Die Aussage folgt also aus Lemma 1.17.[41] □

[40] Die Überlegungen dieses Absatzes sind daher von Lemma 1.17 bzw. Lemma 1.19 unabhängig.

[41] Nach dem Beweis zu Lemma 1.17 ergibt sich das Ableitbarkeitsprädikat $\bigvee x \mathrm{bew}_\Phi(x, y)$ aus der $S_{\mathrm{Ar}[v_0,v_1]}$-Formel $\mathrm{bew}_\Phi(x, y)$, die eine entscheidbare Relation in $\mathrm{Th}(\mathcal{N})$ repräsentiert: Eine feste Aufzählung der Ableitungen von S'-Sätzen aus Sätzen aus Φ sei gegeben. Dann

Satz 1.21. *Zu jeder aufzählbaren Menge* Φ *von in* $\mathcal{N}$ *wahren* S_{Ar}*-Sätzen ist* $\chi := \neg\bigvee x \mathrm{bew}_\Phi(x, \alpha(\ulcorner\neg\bigvee x \mathrm{bew}_\Phi(x, \alpha(y))\urcorner))$ *ein unentscheidbarer* S_{Ar}*-Satz:* $\Phi \nvdash \chi$ *und* $\Phi \nvdash \neg\chi$*. Ferner gilt* $\mathcal{N} \models \chi$*.*

Beweis. Es gibt kein Wahrheitsprädikat, da sich ansonsten ein Widerspruch durch die Existenz eines Fixpunkts zur Verneinung des Wahrheitsprädikates ergäbe. Die Herleitung dieses Widerspruchs wird nun übertragen auf $\bigvee x \mathrm{bew}_\Phi(x, y)$ in der Rolle eines möglichen Wahrheitsprädikats. χ sei der nach dem Beweis des Fixpunktsatzes konstruierte Fixpunkt zu $\neg\bigvee x \mathrm{bew}_\Phi(x, y)$. Wäre $\bigvee x \mathrm{bew}_\Phi(x, y)$ ein Wahrheitsprädikat, würde also insbesondere gelten: $\mathcal{N} \models \chi$ gdw. $\mathcal{N} \models \bigvee x \mathrm{bew}_\Phi(x, \ulcorner\chi\urcorner)$, so ergäbe sich aus der Fixpunkteigenschaft[42] der Widerspruch: $\mathcal{N} \models \bigvee x \mathrm{bew}_\Phi(x, \ulcorner\chi\urcorner)$ gdw. $\mathcal{N} \models \neg\bigvee x \mathrm{bew}_\Phi(x, \ulcorner\chi\urcorner)$. Aus der Definition von $\bigvee x \mathrm{bew}_\Phi(x, y)$ folgt jedoch nur:

$$\text{Wenn } \mathcal{N} \models \bigvee x \mathrm{bew}_\Phi(x, \ulcorner\chi\urcorner), \text{ dann } \mathcal{N} \models \neg\bigvee x \mathrm{bew}_\Phi(x, \ulcorner\chi\urcorner).$$

Hieraus ergibt sich $\mathcal{N} \not\models \bigvee x \mathrm{bew}_\Phi(x, \ulcorner\chi\urcorner)$. Und da dies $\mathcal{N} \models \neg\bigvee x \mathrm{bew}_\Phi(x, \ulcorner\chi\urcorner)$ impliziert, folgt über die Fixpunkteigenschaft $\mathcal{N} \models \chi$. Damit gilt $\Phi \nvdash \chi$ und $\Phi \nvdash \neg\chi$.

Der nach dem Beweis des Fixpunktsatzes konstruierte Fixpunkt χ ist genau $\neg\bigvee x \mathrm{bew}_\Phi(x, \alpha(\ulcorner\neg\bigvee x \mathrm{bew}_\Phi(x, \alpha(y))\urcorner))$. □

Satz 1.22. $\mathrm{Th}(\mathcal{A})$ *sei* ω*-widerspruchsfrei.* $\mathrm{Th}(\mathcal{A})$ *erlaube Repräsentierungen. Zu jeder aufzählbaren Menge* Φ *von in* $\mathcal{A}$ *wahren* S*-Sätzen ist* $\chi := \neg\bigvee x \mathrm{bew}_\Phi(x, \ulcorner\neg\bigvee x \mathrm{bew}_\Phi(x, y)\urcorner)$ *ein unentscheidbarer* S*-Satz:* $\Phi \nvdash \chi$ *und* $\Phi \nvdash \neg\chi$*. Ferner gilt* $\mathcal{A} \models \chi$*.*

Beweis. Siehe den Beweis zu Satz 1.21. □

Im Folgenden geht es um den Unentscheidbarkeitssatz nach Rosser, Gödels syntaktische Form des Unvollständigkeitssatzes und den darauf aufbauenden Satz von Rosser. Hierzu wird zunächst der Satz von Tarski auf die Ableitbarkeitsrelation an Stelle der Wahrheitsrelation übertragen.

gibt es eine $S_{\mathrm{Ar}[v_0,v_1]}$-Formel $\mathrm{bew}_\Phi(x, y)$, s. d. für alle $n_1, n_2 \in \mathbb{N}$ gilt: Die n_1-te Ableitung aus Sätzen aus Φ ist eine Ableitung des S'-Satzes mit der Gödelnummer n_2 gdw. $\mathcal{N} \models \mathrm{bew}_\Phi(\boldsymbol{n_1}, \boldsymbol{n_2})$.

[42] $\mathcal{N} \models \chi \leftrightarrow \neg\bigvee x \mathrm{bew}_\Phi(x, \ulcorner\chi\urcorner)$

Die Menge PA bestehe aus folgenden S_{Ar}-Sätzen:

- $\bigwedge x : x + 1 \not\equiv 0$
- $\bigwedge xy : x + 1 \equiv y + 1 \rightarrow x \equiv y$
- für alle $\varphi(x) \in S_{\mathrm{Ar}_{[v_0]}\supseteq}$ mit genau den weiteren freien Variablen $y_0, ..., y_n$ der Satz $\bigwedge y_0...y_n : \varphi(0) \wedge \bigwedge x(\varphi(x) \rightarrow \varphi(x+1)) \rightarrow \bigwedge x\varphi(x)$
- $\bigwedge x : x + 0 \equiv x$
- $\bigwedge xy : x + (y + 1) \equiv (x + y) + 1$
- $\bigwedge x : x \cdot 0 \equiv 0$
- $\bigwedge xy : x \cdot (y + 1) \equiv x \cdot y + x$

PA ist damit ein Axiomensystem für die Peano-Arithmetik erster Stufe.[43]

S sei im Folgenden eine höchstens abzählbare Symbolmenge, die die Symbolmenge S_{Ar} enthält. *PA* sei ab jetzt eine widerspruchsfreie Menge von S-Sätzen, die die Menge PA enthält.[44] Dem Satz von Tarski in seinen vier Varianten entspricht Satz 1.24.

Satz 1.23. *PA erlaubt Repräsentierungen (bezüglich Funktionen und Relationen).*

Beweis. Siehe Tarski u. a. (1953), S. 56, Theorem 6. □

Satz 1.24. *a) Wenn die Menge der Gödelnummern von Sätzen aus $PA^{\vdash}$ in PA repräsentierbar ist, genauer genügt schon: wenn es eine $S_{[v_0]}$-Formel bew(x) gibt, so dass für alle S-Sätze φ gilt:*

$$PA \vdash \varphi \quad \text{gdw.} \quad PA \vdash bew(\ulcorner\varphi\urcorner),$$

so gilt für jeden Fixpunkt χ der Negation $\neg bew(x)$ von $bew(x)$:

$$PA \not\vdash bew(\ulcorner\chi\urcorner) \quad \text{und} \quad PA \not\vdash \neg bew(\ulcorner\chi\urcorner).$$

b) Es gibt keine $S_{[v_0]}$-Formel bew(x), so dass für alle S-Sätze φ gilt

$$PA \vdash \varphi \leftrightarrow bew(\ulcorner\varphi\urcorner).$$

[43] Die folgenden Überlegungen lassen sich auch für schwächere Systeme wie die Robinson Arithmetik durchführen. Vgl. hierzu Tarski u. a. (1953), S. 56, Theorem 6. Gödel selbst führt die Überlegungen für ein stärkeres System, die Peano-Arithmetik höherer Stufe, durch.

[44] Die Bezeichnung für diese Menge ist kursiv im Gegensatz zu der Bezeichnung der Menge der Peano-Axiome, da es sich nicht um eine feste Menge handelt. "*PA*" ist also eine Variable.

c) Die Menge der Gödelnummern von S-Sätzen aus $\{\varphi | PA \nvdash \varphi\}$ *ist nicht repräsentierbar in PA. Genauer gibt es sogar keine* $S_{[v_0]}$*-Formel* $nbew(x)$[45]*, so dass für alle S-Sätze* φ *gilt:*

$$PA \nvdash \varphi \quad gdw. \quad PA \vdash nbew(\ulcorner\varphi\urcorner).$$

d) Es gibt keine $S_{[v_0]}$*-Formel* $nbew(x)$*, so dass für alle S-Sätze* φ *gilt*

$$PA \vdash \neg\varphi \leftrightarrow nbew(\ulcorner\varphi\urcorner).$$

Beweis. Zu a) $bew(x)$ sei eine $S_{[v_0]}$-Formel, so dass für alle S-Sätze φ gilt: $PA \vdash \varphi$ gdw. $PA \vdash bew(\ulcorner\varphi\urcorner)$. χ sei ein Fixpunkt der Negation $\neg bew(x)$ von $bew(x)$; es gilt also $PA \vdash \chi \leftrightarrow \neg bew(\ulcorner\chi\urcorner)$. Damit gilt auch: $PA \vdash \chi$ gdw. $PA \vdash \neg bew(\ulcorner\chi\urcorner)$. Nach Einsetzen von χ für φ ergibt sich:

$$PA \vdash \neg bew(\ulcorner\chi\urcorner) \quad \text{gdw.} \quad PA \vdash bew(\ulcorner\chi\urcorner), \text{ und damit:}$$
$$PA \nvdash bew(\ulcorner\chi\urcorner) \quad \text{und} \quad PA \nvdash \neg bew(\ulcorner\chi\urcorner).$$

Zu b) $bew(x)$ sei eine $S_{[v_0]}$-Formel, so dass für alle S-Sätze φ gilt: $PA \vdash \varphi \leftrightarrow bew(\ulcorner\varphi\urcorner)$. χ sei ein Fixpunkt der Negation $\neg bew(x)$ von $bew(x)$; es gilt also $PA \vdash \chi \leftrightarrow \neg bew(\ulcorner\chi\urcorner)$. Nach Einsetzen von χ für φ ergibt sich

$$PA \vdash \neg bew(\ulcorner\chi\urcorner) \leftrightarrow bew(\ulcorner\chi\urcorner) \text{ und damit}$$
$$PA \vdash bew(\ulcorner\chi\urcorner) \wedge \neg bew(\ulcorner\chi\urcorner).$$

Widerspruch.
Zu c) $nbew(x)$ sei eine solche $S_{[v_0]}$-Formel, so dass für alle S-Sätze φ gilt: $PA \nvdash \varphi$ gdw. $PA \vdash nbew(\ulcorner\varphi\urcorner)$. Dann gibt es einen Fixpunkt χ von $nbew(x)$, es gilt also $PA \vdash \chi \leftrightarrow nbew(\ulcorner\chi\urcorner)$. Damit gilt auch: $PA \nvdash \chi$ gdw. $PA \nvdash nbew(\ulcorner\chi\urcorner)$. Nach Einsetzen von χ für φ ergibt sich:

$$PA \vdash nbew(\ulcorner\chi\urcorner) \quad \text{gdw.} \quad PA \nvdash nbew(\ulcorner\chi\urcorner), \text{ und damit:}$$
$$PA \vdash nbew(\ulcorner\chi\urcorner) \quad \text{und} \quad PA \nvdash nbew(\ulcorner\chi\urcorner).$$

Widerspruch.

[45] Über die $S_{[v_0]}$-Formel $nbew(x)$ ist sonst nichts weiter vorausgesetzt, sie muss also kein negierter Satz sein. "*nbew*" soll "nicht-ableitbar" andeuten.

Zu d) $nbew(x)$ sei eine $S_{[v_0]}$-Formel, so dass für alle S-Sätze φ gilt $PA \vdash \neg\varphi \leftrightarrow nbew(\ulcorner\varphi\urcorner)$. Dann gibt es einen Fixpunkt χ von $nbew(x)$, es gilt also $PA \vdash \chi \leftrightarrow nbew(\ulcorner\chi\urcorner)$ und $PA \vdash \neg\chi \leftrightarrow \neg nbew(\ulcorner\chi\urcorner)$. Nach Einsetzen von χ für φ ergibt sich

$$PA \vdash nbew(\ulcorner\chi\urcorner) \leftrightarrow \neg nbew(\ulcorner\chi\urcorner) \text{ und damit}$$
$$PA \vdash nbew(\ulcorner\chi\urcorner) \wedge \neg nbew(\ulcorner\chi\urcorner).$$

Widerspruch. □

Satz 1.25 (Unentscheidbarkeitssatz nach Rosser). *$PA^{\vdash}$ ist nicht entscheidbar.*

Beweis. Wäre $PA^{\vdash}$ entscheidbar, so wäre die Menge der Gödelnummern von Sätzen aus $\{\varphi | PA \nvdash \varphi\}$ in PA repräsentierbar, was oben ausgeschlossen wurde (Satz 1.24 c)).[46] □

Falls PA entscheidbar ist, ist es damit (syntaktisch) unvollständig. Denn wäre PA vollständig, so wäre $PA^{\vdash}$ entscheidbar.

Im Anschluss an den Satz von Tarski wird mit einer Art Annäherung an die nicht in $\mathcal{N}$ repräsentierbare Relation der Wahrheit in $\mathcal{N}$ fortgefahren. Dazu wird eine aufzählbare Menge Φ von in $\mathcal{N}$ wahren S_{Ar}-Sätzen und die Menge der aus ihr ableitbaren Sätze betrachtet. Dann wird die Tatsache, dass die Menge der aus Φ ableitbaren Sätze in $\mathcal{N}$ repräsentierbar ist, mit dem Satz von Tarski in Verbindung gebracht. Das Ergebnis ist in Satz 1.21 formuliert. Analog wird im Folgenden dargestellt, inwiefern die Relation der Ableitbarkeit aus PA in PA repräsentiert werden kann und wie dies in Verbindung mit Satz 1.24 gebracht werden kann. Das Ergebnis ist der erste Gödelsche Unvollständigkeitssatz und die Ergänzung von Rosser. PA sei ab jetzt aufzählbar.

Die $S_{\mathrm{Ar}[v_0]}$-Formel $\bigvee x \mathrm{bew}_{PA}(x, y)$, das Ableitbarkeitsprädikat, erfüllt die folgende Teilbedingung einer Repräsentierung.

Lemma 1.26. *Für alle S_{Ar}-Sätze φ gilt:*

$$\text{Wenn } PA \vdash \varphi, \text{ dann } \mathrm{PA} \vdash \bigvee x \mathrm{bew}_{PA}(x, \ulcorner\varphi\urcorner).$$

[46] Vgl. Rosser (1936), Theorem III. Der Beweis von Rosser ist allerdings etwas anders aufgebaut und benutzt die Repräsentierbarkeit der Menge der Gödelnummern der ableitbaren Sätze.

Beweis. Der Beweis läuft parallel zum Beweis von Lemma 1.20. Da PA Repräsentierungen erlaubt, gibt es eine $S_{\mathrm{Ar}[v_0,v_1]}$-Formel $\mathrm{bew}_{PA}(x,y)$, so dass, bei gegebener Aufzählung der Ableitungen von S_{Ar}-Sätzen aus Sätzen aus PA, für alle $n_1, n_2 \in \mathbb{N}$ gilt: Die n_1-te Ableitung aus Sätzen aus PA ist eine Ableitung des S_{Ar}-Satzes mit der Gödelnummer n_2 gdw. PA $\vdash \mathrm{bew}_{PA}(\boldsymbol{n_1}, \boldsymbol{n_2})$. Damit gilt für jeden S_{Ar}-Satz φ: Wenn $PA \vdash \varphi$, dann PA $\vdash \bigvee x \mathrm{bew}_{PA}(x, \ulcorner\varphi\urcorner)$. □

Lemma 1.27. PA *ist ω-widerspruchsfrei.*

Lemma 1.28. *PA sei ω-widerspruchsfrei. Dann gilt für alle S_{Ar}- Sätze φ:*

$$PA \vdash \varphi \text{ gdw. } PA \vdash \bigvee x \mathrm{bew}_{PA}(x, \ulcorner\varphi\urcorner).$$

Beweis. Es gelte $PA \vdash \bigvee x \mathrm{bew}_{PA}(x, \ulcorner\varphi\urcorner)$. Da PA ω-widerspruchsfrei ist, gibt es ein $n \in \mathbb{N}$ mit $PA \nvdash \neg\mathrm{bew}_{PA}(\boldsymbol{n}, \ulcorner\varphi\urcorner)$. Da $\mathrm{bew}_{PA}(x,y)$ die Relation in PA stark repräsentiert, die angibt, ob die n_1-te Ableitung aus Sätzen aus PA eine Ableitung des S_{Ar}-Satzes mit der Gödelnummer n_2 ist, ist die n-te Ableitung aus Sätzen aus PA eine Ableitung des Satzes φ. Es gilt also $PA \vdash \varphi$. Mit Lemma 1.26 folgt die Behauptung.[47] □

Das Ableitbarkeitsprädikat $\bigvee x \mathrm{bew}^{\mathrm{R}}_{\Phi}(x,y)$ von Rosser für eine aufzählbare Menge Φ von S-Sätzen ist definiert als Entsprechung zum umgangssprachlichen "Es gibt ein x, so dass die x-te Ableitung aus Sätzen aus Φ eine Ableitung des S-Satzes mit der Gödelnummer y ist und alle Ableitungen vor der x-ten keine Ableitungen des dazu negierten Satzes sind":[48]

Definition 1.29.

$$\mathrm{bew}^{\mathrm{R}}_{\Phi}(x,y) := \mathrm{bew}_{\Phi}(x,y) \wedge \bigwedge z[\bigvee w(z+w=x) \rightarrow \neg\mathrm{bew}_{\Phi}(z, \mathrm{neg}(y))],$$

wobei $\mathrm{neg}(y)$ ein $S_{\mathrm{Ar}[v_0]}$-Term sei, der die berechenbare Funktion, die die Gödelnummer eines Satzes auf die Gödelnummer seiner Negation abbildet, repräsentiert.

[47] Dies ist zusammen mit dem Beweis von Lemma 1.30 und dem Beweis von Satz 1.31 a) (zweiter Teil) die einzige Stelle, an der auf den Begriff der starken Repräsentierung Bezug genommen wird. Mit $\mathrm{bew}_{PA}(x,y)$ wird auf eine $S_{\mathrm{Ar}[v_0]}$-Formel Bezug genommen wird, die die Relation repräsentiert, die angibt, ob die n_1-te Ableitung aus Sätzen aus PA eine Ableitung des S_{Ar}-Satzes mit der Gödelnummer n_2 ist. Gleichzeitig wird benötigt, dass die Negation dieser Formel das Komplement der Relation repräsentiert.

[48] Siehe Rosser (1936), S. 89 f.

Es gilt: $\Phi \vdash \varphi$ gdw. $\mathcal{N} \models \bigvee x \mathrm{bew}^{\mathrm{R}}_{\Phi}(x, \ulcorner\varphi\urcorner)$. Dies wird jedoch im Folgenden nicht benutzt; es wird direkt zur Entsprechung im Beweis des Unvollständigkeitssatzes zu Lemma 1.26 und Lemma 1.28 übergegangen.

Lemma 1.30. *Für alle S-Sätze φ gilt:*

$$\textit{Wenn } PA \vdash \varphi, \textit{ dann } \mathrm{PA} \vdash \bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner\varphi\urcorner);$$
$$\textit{wenn } PA \vdash \neg\varphi, \textit{ dann } \mathrm{PA} \vdash \neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner\varphi\urcorner).$$

Beweis. Teil 1 gilt analog zu Lemma 1.26 bzw. 1.28. Teil 2: $\neg\varphi$ sei ableitbar. Dann gibt es $n \in \mathbb{N}$, so dass $\mathrm{PA} \vdash \mathrm{bew}_{PA}(\boldsymbol{n}, \ulcorner\neg\varphi\urcorner)$. Für alle $m < n, m \in \mathbb{N}$ gilt $\mathrm{PA} \vdash \neg\mathrm{bew}_{PA}(\boldsymbol{m}, \ulcorner\varphi\urcorner)$. Und damit $\mathrm{PA} \vdash \mathrm{bew}_{PA}(x, \ulcorner\varphi\urcorner) \to \bigvee w(\boldsymbol{n} + w = x)$. Also auch $\mathrm{PA} \vdash \neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner\varphi\urcorner)$. □

Nun wird Satz 1.24 auf $\bigvee x \mathrm{bew}_{PA}(x, y)$ bzw. $\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, y)$ angewandt.

Der Gödelsatz $\mathrm{gö}_{\Phi}$ zu einer aufzählbaren Menge Φ von S-Sätzen sei definiert als Fixpunkt (als der im Beweis des Fixpunktsatzes konstruierte Fixpunkt) der Verneinung des Ableitbarkeitsprädikats $\bigvee x \mathrm{bew}_{\Phi}(x, y)$. $\mathrm{gö}_{\Phi}$ ist also der Satz $\neg\bigvee x \mathrm{bew}_{\Phi}(x, \alpha(\ulcorner\neg\bigvee x \mathrm{bew}_{\Phi}(x, \alpha(y))\urcorner))$. Und es gilt
$\mathrm{PA} \vdash \mathrm{gö}_{PA} \leftrightarrow \neg\bigvee x \mathrm{bew}_{PA}(x, \ulcorner\mathrm{gö}_{PA}\urcorner)$. Entsprechend sei auch der Gödelsatz $\mathrm{gö}^{\mathrm{R}}_{\Phi}$ als Fixpunkt der Verneinung des Ableitbarkeitsprädikats von Rosser definiert.

Satz 1.31 (Der erste Gödelsche/Rossersche Unvollständigkeitssatz).
a) $PA \nvdash \neg\bigvee x \mathrm{bew}_{PA}(x, \ulcorner\mathrm{gö}_{PA}\urcorner)$, jedoch gilt für alle $n \in \mathcal{N}$:
$\mathrm{PA} \vdash \neg\mathrm{bew}_{PA}(\boldsymbol{n}, \ulcorner\mathrm{gö}_{PA}\urcorner)$.
b)[49] *PA sei ω-widerspruchsfrei. Dann gilt $PA \nvdash \bigvee x \mathrm{bew}_{PA}(x, \ulcorner\mathrm{gö}_{PA}\urcorner)$ und $PA \nvdash \neg\bigvee x \mathrm{bew}_{PA}(x, \ulcorner\mathrm{gö}_{PA}\urcorner)$.*
c) $PA \nvdash \neg\neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner\mathrm{gö}^{\mathrm{R}}_{PA}\urcorner)$ und $PA \nvdash \neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner\mathrm{gö}^{\mathrm{R}}_{PA}\urcorner)$.

Beweis. Zu a) Aus der Definition von $\bigvee x \mathrm{bew}_{PA}(x, y)$ folgt nur die eine Richtung der Voraussetzung in Satz 1.24 a), angewandt auf den Fixpunkt also:

$$\text{Wenn } PA \vdash \mathrm{gö}_{PA}, \text{ dann } PA \vdash \bigvee x \mathrm{bew}_{PA}(x, \ulcorner\mathrm{gö}_{PA}\urcorner).$$

[49] In Ehrenfeucht u. Feferman (1960) wird gezeigt, dass jede aufzählbare Menge auch unter der Bedingung der einfachen Widerspruchsfreiheit in *PA* repräsentierbar ist. Vorausgesetzt *PA* ist aufzählbar. Damit muss die ω-Widerspruchsfreiheit hier nicht gefordert werden.

Es ergibt sich damit auch nur:

$$\text{Wenn } PA \vdash \neg\bigvee x \mathrm{bew}_{PA}(x, \ulcorner \mathrm{gö}_{PA} \urcorner), \text{ dann } PA \vdash \bigvee x \mathrm{bew}_{PA}(x, \ulcorner \mathrm{gö}_{PA} \urcorner);$$
$$\text{und damit } PA \nvdash \neg\bigvee x \mathrm{bew}_{PA}(x, \ulcorner \mathrm{gö}_{PA} \urcorner).$$

Da $\mathrm{gö}_{PA}$ Fixpunkt zu $\neg\bigvee x \mathrm{bew}_{PA}(x, y)$ ist, gilt damit auch $PA \nvdash \mathrm{gö}_{PA}$. Und da $\mathrm{bew}_{PA}(x, y)$ die Relation in PA stark repräsentiert, die angibt, ob die n_1-te Ableitung aus Sätzen aus PA eine Ableitung des S_{Ar}-Satzes mit der Gödelnummer n_2 ist, gilt für alle $n \in \mathcal{N}$

$$\mathrm{PA} \vdash \neg \mathrm{bew}_{PA}(\boldsymbol{n}, \ulcorner \mathrm{gö}_{PA} \urcorner).$$

Zu b) Da PA ω-widerspruchsfrei ist, gilt nach Lemma 1.28 für alle S_{Ar}-Sätze φ: $PA \vdash \varphi$ gdw. $PA \vdash \bigvee x \mathrm{bew}_{PA}(x, \ulcorner \varphi \urcorner)$. Aus Satz 1.24 a) folgt damit $PA \nvdash \neg\bigvee x \mathrm{bew}_{PA}(x, \ulcorner \mathrm{gö}_{PA} \urcorner)$ und $PA \nvdash \bigvee x \mathrm{bew}_{PA}(x, \ulcorner \mathrm{gö}_{PA} \urcorner)$.
Zu c) Parallel zum Beweis zu a) ergibt sich, $PA \nvdash \neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner \mathrm{gö}^{\mathrm{R}}_{PA} \urcorner)$. Da[50] $\mathrm{gö}^{\mathrm{R}}_{PA}$ Fixpunkt zu $\neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, y)$ ist, gilt auch

$$PA \vdash \neg \mathrm{gö}^{\mathrm{R}}_{PA} \quad \text{gdw.} \quad PA \vdash \neg\neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner \mathrm{gö}^{\mathrm{R}}_{PA} \urcorner).$$

Nach Lemma 1.30 gilt:

$$\text{Wenn } PA \vdash \neg \mathrm{gö}^{\mathrm{R}}_{PA}, \text{ dann PA} \vdash \neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner \mathrm{gö}^{\mathrm{R}}_{PA} \urcorner).$$

Es ergibt sich also:

$$\text{Wenn } PA \vdash \neg\neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner \mathrm{gö}^{\mathrm{R}}_{PA} \urcorner), \text{ dann PA} \vdash \neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner \mathrm{gö}^{\mathrm{R}}_{PA} \urcorner);$$
$$\text{und damit } PA \nvdash \neg\neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner \mathrm{gö}^{\mathrm{R}}_{PA} \urcorner).$$

□

Die Originalformulierung der Sätze bei Gödel und Rosser ist eine von der obigen leicht verschiedene. Die Teile a), b) und c) können in einer der Originalformu-

[50] Für diesen Teil des Beweises besteht zunächst keine direkte Verbindung zu Satz 1.24. Diese wird erst in Abschnitt 2.1 hergestellt, indem auch dieser Teil von Satz 1.31 als Spezialfall eines abstrakten Satzes dargestellt wird. Der Begriff der Repräsentierbarkeit wird dabei erweitert, so dass von einer Repräsentierbarkeit in der Menge der S_{Ar}-Sätze ψ mit $PA \vdash \neg\psi$ gesprochen werden kann.

lierung entsprechenden Fassung etwa wie folgt wiedergegeben werden:
a) $PA \nvdash \neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(y))\urcorner))$[51], jedoch gilt für alle $n \in \mathcal{N}$:
$\mathrm{PA} \vdash \neg\mathrm{bew}_{PA}(\boldsymbol{n}, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(y))\urcorner))$.
b) PA sei ω-widerspruchsfrei. Dann gilt
$PA \nvdash \bigvee x\mathrm{bew}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(y))\urcorner))$ und
$PA \nvdash \neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(y))\urcorner))$.
c) $PA \nvdash \neg\neg\bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(y))\urcorner))$ und
$PA \nvdash \neg\bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(y))\urcorner))$.[52]

Die Fassungen sind äquivalent, da
$PA \vdash \alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(y))\urcorner) \equiv \ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(y))\urcorner))\urcorner$.
Entsprechendes gilt auch für den Fall des Ableitbarkeitsprädikats von Rosser. Gödels ursprünglicher Beweis – und damit auch Rossers, der darauf aufbaut – unterscheidet sich in seiner Struktur leicht von dem obigen. Auf den ursprünglichen Beweis und die Unterschiede zu dem oben gegebenen wird in Abschnitt 2.1 eingegangen.

Satz 1.32 (Der zweite Gödelsche Unvollständigkeitssatz).
Der S_{Ar}-Satz kon_{PA} sei definiert als $\mathrm{kon}_{PA} := \neg\bigvee x\mathrm{bew}_{PA}(x, \ulcorner\bigvee yy \neq y\urcorner)$. Dann gilt

$$\mathcal{N} \models \mathrm{kon}_{PA} \quad \textit{und} \quad PA \nvdash \mathrm{kon}_{PA}.$$

Beweis. [53] Nach Voraussetzung gilt für alle S_{Ar}-Sätze φ: Wenn $PA \vdash \varphi$, dann $PA \vdash \bigvee x\mathrm{bew}_{PA}(x, \ulcorner\varphi\urcorner)$. Der Beweis des Satzes lässt sich darauf aufbauen, dass $\mathrm{bew}_{PA}(x, y)$ so gewählt werden kann, dass diese Bedingung und die Regel des Modus ponens auch aus Sicht von PA gelten. Genauer, dass für alle S_{Ar}-Sätze φ, ψ gilt:
A1: $PA \vdash \bigvee x\mathrm{bew}_{PA}(x, \ulcorner\varphi\urcorner) \rightarrow \bigvee x\mathrm{bew}_{PA}(x, \ulcorner\bigvee x\mathrm{bew}_{PA}(x, \ulcorner\varphi\urcorner)\urcorner)$,
A2: $PA \vdash \bigvee x\mathrm{bew}_{PA}(x, \ulcorner\varphi\urcorner) \wedge \bigvee x\mathrm{bew}_{PA}(x, \ulcorner\varphi \rightarrow \psi\urcorner) \rightarrow \bigvee x\mathrm{bew}_{PA}(x, \ulcorner\psi\urcorner)$.

Auf der Basis von PA lässt sich nun der erste Unvollständigkeitssatz nachvollziehen.

Nach A1 gilt

$$PA \vdash \bigvee x\mathrm{bew}_{PA}(x, \ulcorner\mathrm{gö}_{PA}\urcorner) \rightarrow \bigvee x\mathrm{bew}_{PA}(x, \ulcorner\bigvee x\mathrm{bew}_{PA}(x, \ulcorner\mathrm{gö}_{PA}\urcorner)\urcorner).$$

[51] D. h. $PA \nvdash \mathrm{gö}_{PA}$.
[52] D. h. $PA \nvdash \mathrm{gö}^{\mathrm{R}}_{PA}$ und $PA \nvdash \neg\mathrm{gö}^{\mathrm{R}}_{PA}$.
[53] Zum Beweis vgl. z. B. Smorynski (1978), S. 825 ff.

Da gö_{PA} Fixpunkt zu $\neg\bigvee x\text{bew}_{PA}(x,y)$ ist, gilt

$$PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\neg\bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner) \rightarrow \text{gö}_{PA}\urcorner).$$

Und damit nach A2

$$PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\neg\bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)\urcorner) \rightarrow \bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)\urcorner).$$

Insgesamt erhält man

$$PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\neg\bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)\urcorner) \rightarrow \\ \bigvee x\text{bew}_{PA}(x, \ulcorner\bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)\urcorner).$$

Dies entspricht dem Ergebnis des ersten Unvollständigkeitssatzes auf der Basis von *PA*.

Zur besseren Übersicht sei "$\bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)$" abgekürzt durch "$\varphi$". Da $PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\neg\varphi \rightarrow (\varphi \rightarrow \varphi \wedge \neg\varphi)\urcorner)$, gilt nach A2 und dem schon Bewiesenen: (*) $PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\neg\varphi\urcorner) \rightarrow \bigvee x\text{bew}_{PA}(x, \ulcorner\varphi \wedge \neg\varphi\urcorner)$. Also

$$PA \vdash \neg\bigvee x\text{bew}_{PA}(x, \ulcorner\varphi \wedge \neg\varphi\urcorner) \rightarrow \neg\bigvee x\text{bew}_{PA}(x, \ulcorner\neg\varphi\urcorner).$$

Auf Grund von A1 gilt $PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\varphi \wedge \neg\varphi\urcorner) \rightarrow \bigvee x\text{bew}_{PA}(x, \ulcorner\bigvee yy \neq y\urcorner)$ und damit

$$PA \vdash \neg\bigvee x\text{bew}_{PA}(x, \ulcorner\bigvee yy \neq y\urcorner) \rightarrow \neg\bigvee x\text{bew}_{PA}(x, \neg\bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)).$$

Da $PA \vdash \neg\bigvee x\text{bew}_{PA}(x, \ulcorner\neg\bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)\urcorner) \leftrightarrow \neg\bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)$ und $PA \nvdash \neg\bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)$, gilt auch $PA \nvdash \text{kon}_{PA}$.[54] □

Bemerkung. Entsprechend sei $\text{kon}^{\text{R}}_{PA}$ definiert als

[54] Führt man den ersten Unvollständigkeitssatz an dieser Stelle nicht auf der Grundlage der $S_{\text{Ar}[v_0]}$-Formel $\neg\bigvee x\text{bew}_{PA}(x,y)$ durch, sondern parallel zur Antinomie von Curry auf der Grundlage der $S_{\text{Ar}[v_0]}$-Formel $\bigvee x\text{bew}_{PA}(x,y) \rightarrow \psi$, so erhält man den Satz von Löb. Dabei sei ψ ein beliebiger S_{Ar}-Satz. Statt der Zeile (*) erhält man nämlich $PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\varphi \rightarrow \psi\urcorner) \rightarrow \bigvee x\text{bew}_{PA}(x, \ulcorner\psi\urcorner)$. Und auf Grund der Fixpunkteigenschaft $PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner) \rightarrow \bigvee x\text{bew}_{PA}(x, \ulcorner\psi\urcorner)$. Es gelte die Voraussetzung des Satzes von Löb: $PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\psi\urcorner) \rightarrow \psi$. Dann gilt $PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner) \rightarrow \psi$ und damit $PA \vdash \bigvee x\text{bew}_{PA}(x, \ulcorner\text{gö}_{PA}\urcorner)$ und $PA \vdash \psi$, was die Aussage des Satzes von Löb ist. Vgl. Boolos u. Jeffrey (2002), Kap. 16.

$\mathrm{kon}^{\mathrm{R}}_{PA} := \neg\bigvee x \mathrm{bew}^{\mathrm{R}}_{PA}(x, \ulcorner\bigvee yy \neq y\urcorner)$. Dann gilt PA $\vdash \mathrm{kon}^{\mathrm{R}}{}_{\mathrm{PA}}$.[55] Dies macht die Intensionalität des 2. Unvollständigkeitssatzes deutlich.

1.3.4 Das Halteproblem

Der Begriff der Registermaschine sei in einer üblichen Weise definiert.[56] Ferner sei eine effektive Gödelisierung[57] ($\ulcorner\cdot\urcorner$) von Registermaschinen gegeben. Es genügt hier die Betrachtung von Registermaschinen über einem einstelligen Alphabet $\{\mathrm{l}\}$, da berechenbare Funktionen von $\mathbb{N}$ nach $\mathbb{N}$ dargestellt werden sollen. Zur Darstellung des Halteproblems genügt es ferner, nur Registermaschinen zu betrachten, die nur ein Eingaberegister und die einzige Ausgabemöglichkeit l haben. Für eine solche Registermaschine Q gibt es dann zwei Möglichkeiten: Angesetzt auf die Eingabe $\boldsymbol{n}$[58] hält sie entweder mit Ergebnis l an – Bezeichnungsweise: $Q(\boldsymbol{n}) \hookrightarrow \mathrm{l}$ – oder sie hält nicht an (liefert kein Ergebnis) – Bezeichnungsweise: $Q(\boldsymbol{n}) \not\hookrightarrow \mathrm{l}$.[59]

Definition 1.33. Eine Relation $R \subseteq \mathbb{N}$ heißt *repräsentierbar* gdw. es eine Registermaschine Q gibt, so dass für alle $n \in \mathbb{N}$:

$$n \in R \quad \text{gdw.} \quad Q(\boldsymbol{n}) \hookrightarrow \mathrm{l};$$

Q repräsentiert R.

Satz 1.34 (Unentscheidbarkeit des Halteproblems). [60] *Die Menge der Gödelnummern von Registermaschinen Q mit $Q(\ulcorner Q\urcorner) \not\hookrightarrow \mathrm{l}$ ist nicht repräsentierbar. Genauer gibt es sogar keine Registermaschine $nHalt$, so dass für alle Registermaschinen Q gilt:*

$$Q(\ulcorner Q\urcorner) \not\hookrightarrow \mathrm{l} \quad \textit{gdw.} \quad nHalt(\ulcorner Q\urcorner) \hookrightarrow \mathrm{l}.$$

[55] Vgl. z. B. Smorynski (1978), S. 841.

[56] Eine solche Definition findet sich z. B. in Ebbinghaus u. a. (1992), S. 180 f.

[57] Vgl. S. 31.

[58] Für $n \in \mathbb{N}$ sei $\boldsymbol{n}$ die n-gliedrige Zeichenkette, bei der jeder Eintrag l ist.

[59] Über den Begriff einer wie oben eingeschränkten Registermaschine lässt sich der Begriff der Aufzählbarkeit einführen. Es gilt nämlich: Für jede Registermaschine Q ist $\{n \in \mathbb{N} | Q(\boldsymbol{n}) \hookrightarrow \mathrm{l}\}$ aufzählbar. Umgekehrt gibt es zu jeder aufzählbaren Menge M natürlicher Zahlen eine solche Registermaschine Q mit $M = \{n \in \mathbb{N} | Q(\boldsymbol{n}) \hookrightarrow \mathrm{l}\}$.

[60] In den hier wesentlichen Schritten entspricht die Fassung dieses Satzes und des folgenden Beweises der Version in Church (1936), die statt auf dem Begriff der Registermaschine auf dem λ-Kalkül von Church als Berechenbarkeitsbegriff basiert.

Eine Formulierung dieses Satzes in Bezug auf die Repräsentierbarkeit der Menge der Gödelnummern von Registermaschinen Q mit $Q(\ulcorner Q \urcorner) \hookrightarrow 1$ ergibt:

Satz 1.35. *a) Die Menge der Gödelnummern von Registermaschinen Q mit $Q(\ulcorner Q \urcorner) \hookrightarrow 1$ ist repräsentierbar.*
b) Halt *repräsentiere die Menge aus a). Dann gilt für jede Registermaschine $nHalt$, für die für keinen Eingabewert $\boldsymbol{n}$ sowohl* $\mathrm{Halt}(\boldsymbol{n}) \hookrightarrow 1$ *als auch $nHalt(\boldsymbol{n}) \hookrightarrow 1$ gilt:*

$$\mathrm{Halt}(\ulcorner nHalt \urcorner) \not\hookrightarrow 1 \quad \textit{und} \quad nHalt(\ulcorner nHalt \urcorner) \not\hookrightarrow 1.$$

Zu Satz 1.24 besteht folgender Zusammenhang: Für eine ω-widerspruchsfreie Menge PA von S_{Ar}-Sätzen, die die Menge PA enthält, gilt: Für jede Teilmenge $M \subseteq \mathbb{N}$ gilt: Wenn es eine Registermaschine Q gibt, so dass für alle $n \in \mathbb{N}$: $n \in M$ gdw. $Q(\boldsymbol{n}) \hookrightarrow 1$, dann gibt es eine $S_{\mathrm{Ar}[v_0]}$-Formel $\varphi(x)$, so dass für alle $n \in \mathbb{N}$: $n \in M$ gdw. $PA \vdash \varphi(\boldsymbol{n})$. Für eine widerspruchsfreie und aufzählbare Menge Φ von S-Sätzen (S enthalte S_N) gilt umgekehrt: Für jede Teilmenge $M \subseteq \mathbb{N}$ gilt: Wenn es eine $S_{[v_0]}$-Formel $\varphi(x)$ gibt, so dass für alle $n \in \mathbb{N}$: $n \in M$ gdw. $\Phi \vdash \varphi(\boldsymbol{n})$, dann gibt es eine Registermaschine Q, so dass für alle $n \in \mathbb{N}$: $n \in M$ gdw. $Q(\boldsymbol{n}) \hookrightarrow 1$.

Im Zusammenhang mit dem Satz von der Unentscheidbarkeit des Halteproblems ergeben sich verschiedene Unentscheidbarkeitsaussagen für Satzmengen:

- Die Symbolmenge S_∞ enthalte abzählbar unendlich viele Objekt-, Relations- und Funktionskonstanten jedes Typs der Stufe 0 oder 1. Die Menge der logisch wahren S^1_∞-Sätze (die Logik erster Stufe) ist unentscheidbar: Jedem Paar aus einer Registermaschine und einer Eingabefolge kann ein S^1_∞-Satz zugeordnet werden, der genau dann aus der leeren Menge ableitbar ist, wenn die Registermaschine irgendwann anhält.[61]
- Die Menge der in $\mathcal{N}$ wahren S_{Ar}-Sätze ist unentscheidbar: Jeder Registermaschine kann eine $S_{\mathrm{Ar}[v_0]}$-Formel zugeordnet werden, die für alle $n \in \mathbb{N}$ angewandt auf $\boldsymbol{n}$ genau dann in $\mathcal{N}$ wahr ist, wenn die Registermaschine angewandt auf $\boldsymbol{n}$ irgendwann anhält.[62]
- Der obige Punkt ist ein Spezialfall von: Für eine ω-widerspruchsfreie Menge PA von S_{Ar}-Sätzen, die die Menge PA enthält, ist $PA^{\vdash}$ unentscheidbar:

[61] Vgl. z. B. Ebbinghaus u. a. (1992), S. 193 ff.
[62] Vgl. z. B. Ebbinghaus u. a. (1992), S. 203 ff.

> Jeder Registermaschine kann eine $S_{\mathrm{Ar}[v_0]}$-Formel zugeordnet werden, die für alle $n \in \mathbb{N}$ angewandt auf $\boldsymbol{n}$ genau dann aus PA ableitbar ist, wenn die Registermaschine angewandt auf $\boldsymbol{n}$ irgendwann anhält.[63]

Hierbei ist relevant, dass die gewählte Zuordnung eines Satzes bzw. einer Formel zu einer Registermaschine berechenbar ist, da von der Entscheidbarkeit bestimmter Satzmengen auf die Entscheidbarkeit bestimmter Mengen von Registermaschinen geschlossen wird.

Durch die obigen Überlegungen wird deutlich, dass und wie auch der Beweis der Unentscheidbarkeit der Logik erster Stufe, der Beweis der Unentscheidbarkeit der Arithmetik und der Ansatz zum Beweis des Unvollständigkeitssatzes, der die Unentscheidbarkeit des Halteproblems benutzt, auf einen Satz (eben den Satz von der Unentscheidbarkeit des Halteproblems) als wesentliche Grundlage zurückgreifen, der seiner Struktur nach dem Satz von Tarski (in der Form von Satz 1.13) und Satz 1.24 gleicht.

[63] Vgl. Church (1936).

Kapitel 2

Abstrakte Darstellungen

2.1 Vergleich semantischer Antinomien und damit verwandter Widerspruchsbeweise – eine abstrakte Beschreibung

In diesem Abschnitt werden die Beziehungen der Lügner-Antinomie, der Antinomien von Grelling und Richard und verschiedener Beweise des Satzes von Tarski, des ersten Unvollständigkeitssatzes und des Unentscheidbarkeitssatzes zueinander untersucht. Genauer geht es um die Herausarbeitung gemeinsamer wesentlicher Strukturen in den genannten Antinomien und Beweisen. Zur Darstellung dieser Strukturen wird eine abstrakte Form der Antinomien und Widerspruchsbeweise entwickelt. Beginnend bei einer Antinomie oder einem Widerspruchsbeweis gibt es mehrere Aspekte, in deren Hinsicht abstrahiert werden kann, bzw. von denen abstrahiert werden kann. So kann zunächst vom syntaktischen Aufbau der Sprache, vom Wahrheits- oder Ableitbarkeitsbegriff und in diesem Zusammenhang von der semantischen Zuordnung und den vorausgesetzten Axiomen und Regeln abstrahiert werden. In dem folgenden Ansatz wird in allen diesen Bereichen abstrahiert. Insbesondere wird keine Negation mit Regeln oder mit semantischer Interpretation in der Sprache vorausgesetzt, so dass auch Darstellungen und Auflösungen von Antinomien, die mit einer nicht-klassischen Negation arbeiten, als Spezialfälle der abstrakten Sätze aufgefasst werden können. Da keine Negation verwendet wird, können die abstrakten Sätze in ihren Beweisen auch nicht den konkreten Widerspruch, sofern es sich um einen objektsprachlichen Widerspruch handelt, oder die konkrete Unvollständigkeit ausdrücken. Die Formulierungen enden in diesem Fall einen Schritt vor dem eigentlichen Widerspruch mit einer Äquivalenz oder einem Äquivalenzschema.

Ziel ist dabei nicht, möglichst schwache Voraussetzungen der konkreten Beweise auch in den abstrakten Formen abbilden zu können, sondern sozusagen nur den Rumpf der Beweise. Daher wird z. B. in Bezug auf den ersten Unvollständigkeitssatz in der abstrakten Form nicht wiedergegeben, dass die ω-Widerspruchsfreiheit von PA zusammen mit der Eigenschaft, Repräsentationen zu erlauben, als Voraussetzung genügt. Im abstrakten Satz taucht nur die gröbere Voraussetzung auf, dass die Menge der aus PA ableitbaren Sätze in PA repräsentierbar sein muss.

Zunächst wird kurz auf andere Ansätze zum Vergleich der oben genannten Antinomien und Widerspruchsbeweise eingegangen. Eine ausführliche Auseinandersetzung mit dem Ansatz von Serény[1] erfolgt in dem gesonderten Abschnitt 2.2.

Auf einem nicht abstrakten Niveau entwickelt Stegmüller[2] Schritt für Schritt aus der Antinomie von Richard den ersten Unvollständigkeitssatz. Seine Ausführungen in diesem Abschnitt haben informellen Charakter. Übertragen auf das formale System des letzten Kapitels gelangt er von der Antinomie von Richard zu folgender Form des Unvollständigkeitssatzes:

> Φ sei eine aufzählbare, widerspruchsfreie Menge von in $\mathbb{N}$ wahren S_{Ar}-Sätzen. Dann gibt es eine $S_{\mathrm{Ar}[v_0]}$-Formel $ri(x)$, so dass für alle $S_{\mathrm{Ar}[v_0]}$-Formeln $\varphi(x)$ gilt:
>
> $$\Phi \nvdash \varphi(\ulcorner\varphi(x)\urcorner) \quad \text{gdw.} \quad \mathbb{N} \models ri(\ulcorner\varphi(x)\urcorner).$$
>
> Daher gilt:
>
> $$\Phi \nvdash ri(\ulcorner ri(x)\urcorner) \quad \text{gdw.} \quad \mathbb{N} \models ri(\ulcorner ri(x)\urcorner).$$
>
> Und es ergibt sich: $\Phi \nvdash ri(\ulcorner ri(x)\urcorner)$ und $\mathbb{N} \models ri(\ulcorner ri(x)\urcorner)$, also auch $\Phi \nvdash \neg ri(\ulcorner ri(x)\urcorner)$.

Diese Form des Unvollständigkeitsbeweises entspricht eher dem semantischen Beweis von Gödel.[3] Er fällt unter den in der vorliegenden Arbeit formulierten abstrakten Satz 2.4 b). Betrachtet man diesen Beweis, so muss man Stegmüller

[1] Vgl. Serény (2003).
[2] Vgl. Stegmüller (1973), S. 3-11.
[3] Zum semantischen Beweis siehe S. 62 der vorliegenden Arbeit.

Recht geben, dass in dieser Hinsicht der Unvollständigkeitssatz eine zur Antinomie von Richard bzw. Grelling parallele Struktur aufweist. Betrachtet man andere Unvollständigkeitsbeweise, ergeben sich Parallelen zu anderen Antinomien.[4]

Mostowski leitet aus dem semantischen Beweis Tarskis, Gödels syntaktischem Beweis und Rossers Beweis des ersten Unvollständigkeitssatzes jeweils Verallgemeinerungen dieses Satzes[5] ab, die sich auf die in den jeweiligen Beweisen benutzten Voraussetzungen an die Satzmenge beziehen. Anschließend zeigt er ihre Unvergleichbarkeit in Bezug auf die Stärke der jeweiligen Voraussetzungen. Für diese drei Sätze, bzw. in ihren Beweisen bei Mostowski, spielt der Begriff der Definierbarkeit in einer Satzmenge eine wichtige Rolle, der für widerspruchsfreie Satzmengen mit dem Begriff der starken Repräsentierbarkeit, wie er in der vorliegenden Arbeit eingeführt wird, zusammenfällt. Alle drei Sätze werden bezogen auf einen Satz[6], dem in der Terminologie der vorliegenden Arbeit grob gesagt die Sätze 1.24 a) und c) entsprechen. Mostowski deutet also diesen Satz als gemeinsames Kernstück der verschiedenen Beweise an. Seine Verallgemeinerungen gehen aber in die entgegengesetzte Richtung, indem gerade die unterschiedlichen Voraussetzungen, die die Bedingungen in den Kernstücken der Beweise ermöglichen, deutlich gemacht werden.

Smullyan[7] gibt mehrere Fassungen des Satzes von Tarski und des ersten Unvollständigkeitssatzes in unterschiedlichen Abstraktionsgraden der Objektsprache an. Zentral ist der Begriff der Repräsentierbarkeit, der mit dem der vorliegenden Arbeit übereinstimmt. In Smullyans Darstellung werden die verschiedenen abstrakten Sätze jedoch nicht systematisch zueinander in Beziehung gesetzt, wie es das Ziel in dieser Arbeit ist. Im folgenden Kapitel wird an vielen Stellen direkt auf Gemeinsamkeiten mit der Darstellung Smullyans verwiesen.

Auch Serény[8] entwickelt auf der Basis einer abstrakten Sprache, die ähnlich der von Smullyan verwendeten ist, und aufbauend auf den Begriff der Repräsentierbarkeit, abstrakte Versionen verschiedener Fassungen des Satzes von Tarski,

[4]Dies wird auf S. 73 näher ausgeführt.

[5]Siehe Mostowski (1964), S. 97, Theorem 1, 2, 3. Diese Verallgemeinerungen sind zum Teil schon in Tarski (1935) bzw. Rosser (1936) enthalten und stellen in Bezug auf den syntaktischen Aufbau der Sprache keine Abstraktionen dar.

[6]Siehe Mostowski (1964), S. 88, Theorem 1 und 5.

[7]Vgl. Smullyan (1992).

[8]Vgl. Serény (2003).

des ersten Unvollständigkeitssatzes und des Unentscheidbarkeitssatzes. Alle abstrakten Sätze werden als Anwendungsfälle *eines* Satzes dargestellt, der wiederum eine abstrakte Version der Antinomie von Richard darstellt. Dies geschieht im Unterschied zu der vorliegenden Arbeit, in der diese Rolle von zwei zwar miteinander verwandten, jedoch unterschiedlichen abstrakten Sätzen eingenommen wird. Die Konstruktion von Serény erfordert damit auch eine andere Interpretation des syntaktischen Beweises von Gödel innerhalb des abstrakten Satzes. Dazu setzt er eine Negation in der Objektsprache voraus, mit der Eigenschaft, dass für die Menge A der ableitbaren abstrakten Sätze und für alle abstrakten Sätze φ gilt: $\varphi \in A$ gdw. $\neg\neg\varphi \in A$.

Auch in Bezug auf die Arbeit von Serény wird an den entsprechenden Stellen auf Gemeinsamkeiten eingegagen. In Abschnitt 2.2 erfolgt ein ausführlicher Vergleich mit dem Ansatz von Serény.

2.1.1 Darstellung ohne Fixpunktsatz

Die für das Zustandekommen der Antinomien und Widersprüche wesentlichen Aspekte der vorkommenden Sprachen werden unter dem Begriff einer abstrakten Sprache zusammengefasst.

Definition 2.1. Eine *abstrakte Sprache*[9] S ist ein Tupel $(F_S, F'_S, T_S, T'_S, e_S, n_S)$ aus

- einer Menge F_S (Menge der S-Formeln)[10],
- einer Menge $F'_S \subseteq F_S$ (Menge der S-Sätze),
- einer Menge T_S (Menge der S-Terme),
- einer Menge $T'_S \subseteq T_S$ (Menge der geschlossenen S-Terme),
- einer Abbildung $e_S : (F_S \setminus F'_S) \times T_S \to F_S$ (Einsetzung eines Terms t in eine Formel $\varphi(x)$ mit genau einer freien Variable, $\varphi(t)$)
  - mit $e_S[(F_S \setminus F'_S) \times T'_S] \subseteq F'_S$ (die Einsetzung eines geschlossenen Terms in eine Formel mit genau einer freien Variable ist ein Satz)

[9]Die hier definierte abstrakte Sprache hat Ähnlichkeit mit der in Smullyan (1992), S. 5 und Serény (2003), S. 8 eingeführten.

[10]Die Beschreibungen in Klammern geben die intendierte Bedeutung der Mengen und Abbildungen an und gehören nicht zur Definition. Im Folgenden werden meist diese Bezeichnungen, die von speziellen Anwendungsfällen her stammen, verwendet, da sie leichter lesbar sind als die für den abstrakten Fall eingeführten Schreibweisen.

  - und $e_S[(F_S\setminus F'_S)\times(T_S\setminus T'_S)] \subseteq F_S\setminus F'_S$ (die Einsetzung eines Terms mit genau einer freien Variable in eine Formel mit genau einer freien Variable ist eine Formel mit genau einer freien Variable),
- einer injektiven Abbildung $n_S : F_S \to T'_S$ (Standardterm oder Standardname einer Formel φ, $\ulcorner\varphi\urcorner$).

Die Menge der $S_{[v_0]}$-Formeln bzw. -Terme sei die Menge $F_S\setminus F'_S$ bzw. $T_S\setminus T'_S$.

Die in diesem Abschnitt untersuchten Antinomien und Widerspruchsbeweise sind alle in einer Sprache formuliert, die sich als abstrakte Sprache in obigem Sinn auffassen lässt: In der Darstellung der Lügner-Antinomie und der Antinomie von Grelling wird vorausgesetzt, dass es in der Sprache Standardterme für die Sätze bzw. Formeln der Sprache gibt. In der Darstellung der Antinomie von Richard, im Satz von Tarski und in Satz 1.24 werden eine injektive Abbildung $\ulcorner\cdot\urcorner$ von der Menge der Formeln der Sprache nach $\mathbb{N}$ und Standardterme für die natürlichen Zahlen, und damit indirekt auch Standardterme für die Formeln der Sprache, vorausgesetzt. Die Menge der Registermaschinen mit einem Eingabefeld kann als Menge der $S_{[v_0]}$-Formeln aufgefasst werden, die Menge der endlichen Folgen über l als Menge T'_S. Die Menge der geordneten Paare aus Registermaschinen und endlichen Folgen über l kann dann als Menge F'_S aufgefasst werden. Über die vorausgesetzte Gödelisierung ergeben sich die Standardterme für Formeln. Die Gödelisierung wird in Abschnitt 1.3.4 nur für Registermaschinen, also $S_{[v_0]}$-Formeln, eingeführt, womit der hier dargestellte Beweis für die Unentscheidbarkeit des Halteproblems auskommt. Die Gödelisierung kann aber natürlich auf geordnete Paare aus Registermaschinen und endlichen Folgen über l, also auf F_S, ausdehnt werden.

Wesentlich für die Formulierung der Antinomien und Widerspruchsbeweise ist weiter die Auszeichnung einer Teilmenge (W) aus der Menge der Sätze, die je nach Anwendungsfall der Menge der wahren bzw. ableitbaren Sätze in den konkreten Fällen entspricht. In den drei Antinomien und dem Satz von Tarski entspricht diese der Menge der in der ausgezeichneten Struktur wahren Sätze. In Satz 1.24 entspricht sie der Menge der aus einer Satzmenge, die die Peano-Axiome erster Stufe enthält, ableitbaren Sätze. Im Unentscheidbarkeitssatz nach Church entspricht sie der Menge der geordneten Paare aus Registermaschinen und Eingaben, bei denen die Registermaschine mit der jeweiligen Eingabe mit dem Ergebnis l stoppt.

Wenn es darum geht, die Struktur der Antinomien und Widerspruchsbeweise deutlich zu machen, ist es jedoch irrelevant, um welche Menge von Sätzen es sich handelt. Darum wird hier einfach von einer festen aber beliebigen Menge W von wahren bzw. ableitbaren Sätzen ausgegangen. In Bezug auf diese Menge kann der Begriff der Repräsentierbarkeit einer Relation definiert werden:

Definition 2.2. S sei eine abstrakte Sprache. W sei eine Menge von S-Sätzen. Eine Relation $R \subseteq F_S$ heißt *repräsentierbar in* W gdw. es eine $S_{[v_0]}$-Formel $\varphi(x)$ gibt, so dass für alle S-Formeln ψ:

$$\psi \in R \quad \text{gdw.} \quad \varphi(\ulcorner\psi\urcorner) \in W.$$

$\varphi(x)$ *repräsentiert* R *in* W.

Der Begriff der Repräsentierbarkeit einer Relation über der Menge der S-Formeln in einer Menge von S-Sätzen W entspricht, über die Gödelisierung, dem der Repräsentierbarkeit einer Relation über $\mathbb{N}$ zum einen in einer Menge von S-Sätzen (S enthalte die Symbolmenge S_{N}, Definition 1.9), zum anderen durch eine Registermaschine (Definition 1.33). Als Menge W wird dabei die Menge der aus der Satzmenge ableitbaren Sätze bzw. der Paare aus Registermaschinen und Eingaben, bei denen die Registermaschine mit der jeweiligen Eingabe mit dem Ergebnis 1 stoppt, gewählt.

Die Antinomien von Grelling und Richard können nun in ihrer metasprachlichen Fassung auf abstrakte Sprachen übertragen werden. Sie werden dabei nicht als Antinomien, sondern als Sätze formuliert. In Teil a) wird angegeben, auf welche Äquivalenzen in der Metasprache geschlossen werden kann, in Teil b) wird aus dem metasprachlich auftretenden Widerspruch auf die Nichtrepräsentierbarkeit einer Formelmenge geschlossen.

Satz 2.3. *S sei eine abstrakte Sprache und W eine Menge von S-Sätzen.*
a) Wenn die Menge der $S_{[v_0]}$-Formeln $\varphi(x)$ mit $\varphi(\ulcorner\varphi(x)\urcorner) \in W$ in W durch eine $S_{[v_0]}$-Formel $\vartheta_1(x)$ repräsentiert wird, so gilt für jede $S_{[v_0]}$-Formel $\vartheta_2(x)$:[11]

$$\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in W \quad \textit{gdw.} \quad \vartheta_1(\ulcorner\vartheta_2(x)\urcorner) \in W.$$

[11] Vgl. den Beweis 2 zum Korollar in Smullyan (1992), S. 60.

b) Die Menge der $S_{[v_0]}$-Formeln $\varphi(x)$ mit $\varphi(\ulcorner\varphi(x)\urcorner) \notin W$ ist nicht in W repräsentierbar.[12]

Beweis. Zu a) Hier ist nichts zu beweisen.
Zu b) Angenommen $\vartheta_2(x)$ ist eine $S_{[v_0]}$-Formel, so dass für alle $S_{[v_0]}$-Formeln $\varphi(x)$ gilt: $\varphi(\ulcorner\varphi(x)\urcorner) \notin W$ gdw. $\vartheta_2(\ulcorner\varphi(x)\urcorner) \in W$. Nach Einsetzen von $\vartheta_2(x)$ für $\varphi(x)$ ergibt sich: $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin W$ gdw. $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in W$. Widerspruch. □

Satz 2.3 a) kann als Entsprechung der folgenden Variante der Antinomie von Grelling oder einer entsprechenden Fassung der Antinomie von Richard aufgefasst werden:[13]

> Es wird ausgegangen von dem Begriff "autologisch". Ein einstelliger Begriff heißt *autologisch* gdw. er auf sich selbst zutrifft.
> Es ergibt sich nun: Wenn der Begriff "nicht autologisch" autologisch ist, trifft er auf sich selbst zu, d. h. er ist nicht autologisch. Wenn der Begriff "nicht autologisch" nicht autologisch ist, trifft er auf sich selbst zu, d. h. er ist autologisch.

Satzes 2.3 b) entspricht den metasprachlichen Fassungen der Antinomien von Grelling und Richard, wie sie in Abschnitt 1.2.2 dargestellt werden.

Gödels syntaktischer Beweis des ersten Unvollständigkeitssatzes in Gödel (1931), S. 187 ff. fällt unter Satz 2.3 a).[14] Hierzu wird W interpretiert als die Menge der aus einer Menge PA von S_{Ar}-Sätzen ableitbaren S_{Ar}-Sätze. Dabei sei vorausgesetzt, dass PA aufzählbar und ω-widerspruchsfrei ist und die Menge PA enthält. Da die Relation der Ableitbarkeit aus PA in PA durch[15] $\bigvee x \mathrm{bew}_{PA}(x,y)$ repräsentiert wird und die Menge der $S_{\mathrm{Ar}[v_0]}$-Formeln $\varphi(x)$ mit $PA \vdash \varphi(\ulcorner\varphi(x)\urcorner)$ damit durch $\bigvee x \mathrm{bew}_{PA}(x,\alpha(y))$, ergibt die Anwendung von Satz 2.3 a) für $\vartheta_2(y) = \neg\bigvee x \mathrm{bew}_{PA}(x,\alpha(y))$:
$PA \vdash \neg\bigvee x \mathrm{bew}_{PA}(x,\alpha(\ulcorner\neg\bigvee x \mathrm{bew}_{PA}(x,\alpha(y))\urcorner))$ gdw.
$PA \vdash \bigvee x \mathrm{bew}_{PA}(x,\alpha(\ulcorner\neg\bigvee x \mathrm{bew}_{PA}(x,\alpha(y))\urcorner))$.

[12] Teil b) entspricht Theorem (T)1. in Smullyan (1992), S. 9 und Proposition (Liar Theorem) und Theorem 2(iii) in Serény (2003), S. 10 bzw. S. 16.
[13] In Abschnitt 1.2.2 werden die Antinomien von Grelling und Richard nur ausgehend von dem Begriff "heterologisch", "trifft nicht auf sich selbst zu" dargestellt.
[14] Vgl. die Bemerkungen zur Originalformulierung von Gödel auf Seite 48.
[15] Die folgenden Bezeichnungen sind Abschnitt 1.3.3 entnommen.

Daraus ergibt sich $PA \nvdash \bigvee x \mathrm{bew}_{PA}(x, \alpha(\ulcorner \neg \bigvee x \mathrm{bew}_{PA}(x, \alpha(y)) \urcorner))$ und $PA \nvdash \neg \bigvee x \mathrm{bew}_{PA}(x, \alpha(\ulcorner \neg \bigvee x \mathrm{bew}_{PA}(x, \alpha(y)) \urcorner))$.

Ausgangspunkt für die Bestimmung des Ableitbarkeitsprädikats $\bigvee x \mathrm{bew}_{PA}(x, y)$ in Abschnitt 1.3.3 war die Repräsentierung der Relation "die n_1-te Ableitung aus Sätzen aus PA ist eine Ableitung des S_{Ar}-Satzes mit der Gödelnummer n_2" durch $\mathrm{bew}_{PA}(x, y)$ in PA.[16] Gödels Ausgangspunkt in seiner Originalformulierung ist die Repräsentierung der Relation "die n_1-te Ableitung aus Sätzen aus PA ist keine Ableitung des S_{Ar}-Satzes mit der Gödelnummer n_2" in PA durch eine $S_{\mathrm{Ar}[v_0]}$-Formel. Demnach ergibt sich sein Beweisprädikat, auf die Bezeichnungen dieser Arbeit übertragen, als $\neg \bigwedge x \neg \mathrm{bew}_{PA}(x, y)$. Gödels Unentscheidbarkeitssatz führt als nicht aus PA ableitbare S_{Ar}-Sätze die Sätze $\bigwedge x \neg \mathrm{bew}_{PA}(x, \alpha(\ulcorner \bigwedge x \neg \mathrm{bew}_{PA}(x, \alpha(y)) \urcorner))$ und $\neg \bigwedge x \neg \mathrm{bew}_{PA}(x, \alpha(\ulcorner \bigwedge x \neg \mathrm{bew}_{PA}(x, \alpha(y)) \urcorner))$ auf. Für Gödels Beweis wesentlich ist jedoch, wie in der vorliegenden Darstellung, die Repräsentierbarkeit der Ableitbarkeit aus PA – hier durch $\bigvee x \mathrm{bew}_{PA}(x, y)$, in Gödels Originalbeweis durch $\neg \bigwedge x \neg \mathrm{bew}_{PA}(x, y)$.

Interpretiert man in 2.3 b) W als Menge von Paaren aus Registermaschinen und Eingaben, bei denen die Registermaschine mit der jeweiligen Eingabe mit dem Ergebnis l stoppt, ergibt sich aus Satz 2.3 b) der Satz von der Unentscheidbarkeit des Halteproblems.

Satz 2.3 reicht zur Interpretation der Beweise des letzten Kapitels noch nicht aus. Denn zum Teil werden nicht Satzmengen in sich selbst oder in ihrem Komplement bzw. Formelmengen aus Formeln, deren Selbsteinsetzung in einer Satzmenge liegt, in ebendieser oder in ihrem Komplement repräsentiert, sondern Teilmengen oder Obermengen von ihnen. Der folgende Satz, der sich auf Teilmengen von W bezieht, ist eine Verallgemeinerung des letzten Satzes.

Satz 2.4. *S sei eine abstrakte Sprache und W eine Menge von S-Sätzen.*
a) Für $A \subseteq W$ repräsentiere die $S_{[v_0]}$-Formel $\vartheta_1(x)$ die Menge der $S_{[v_0]}$-Formeln $\varphi(x)$ mit $\varphi(\ulcorner \varphi(x) \urcorner) \in A$ in W. Dann gilt für jede $S_{[v_0]}$-Formel $\vartheta_2(x)$:

$$(i) \quad \textit{Wenn } \vartheta_1(\ulcorner \vartheta_2(x) \urcorner) \in W, \textit{ dann } \vartheta_2(\ulcorner \vartheta_2(x) \urcorner) \in W.$$

$$(ii) \quad \textit{Wenn } \vartheta_1(\ulcorner \vartheta_2(x) \urcorner) \in A, \textit{ dann } \vartheta_2(\ulcorner \vartheta_2(x) \urcorner) \in A.$$

[16] Siehe Fußnote 41 auf S. 40.

b)[17] *Für $A \subseteq W$ repräsentiere die $S_{[v_0]}$-Formel $\vartheta_2(x)$ die Menge der $S_{[v_0]}$-Formeln $\varphi(x)$ mit $\varphi(\ulcorner\varphi(x)\urcorner) \notin A$ in W. Dann gilt:*

$$\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin A \textit{ und } \vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in W.$$

Beweis. Zu a) Nach Voraussetzung gilt für alle $S_{[v_0]}$-Formeln $\varphi(x)$: $\varphi(\ulcorner\varphi(x)\urcorner) \in A$ gdw. $\vartheta_1(\ulcorner\varphi(x)\urcorner) \in W$. Nach Einsetzen von $\vartheta_2(x)$ für $\varphi(x)$ ergibt sich:

$$\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in A \quad \text{gdw.} \quad \vartheta_1(\ulcorner\vartheta_2(x)\urcorner) \in W.$$

Nach Voraussetzung gilt $A \subseteq W$. Damit folgen (i) und (ii).

Zu b) Nach Voraussetzung gilt für alle $S_{[v_0]}$-Formeln $\varphi(x)$: $\varphi(\ulcorner\varphi(x)\urcorner) \notin A$ gdw. $\vartheta_2(\ulcorner\varphi(x)\urcorner) \in W$. Nach Einsetzen von $\vartheta_2(x)$ für $\varphi(x)$ ergibt sich:

$$\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin A \quad \text{gdw.} \quad \vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in W.$$

Nach Voraussetzung gilt $A \subseteq W$ und damit:

$$\text{Wenn } \vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in A, \quad \text{dann} \quad \vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin A.$$

Also gilt: $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin A$ und $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in W$. □

Der Satz lässt sich allgemein für Mengen A formulieren.[18] Da jedoch in den hier relevanten Fällen nur Teilmengen von W und Obermengen von W vorkommen, folgt nur die Variante für eine Obermenge A von W.

Satz 2.5. *S sei eine abstrakte Sprache und W eine Menge von S-Sätzen.*
a) Für $A \supseteq W$ repräsentiere die $S_{[v_0]}$-Formel $\vartheta_1(x)$ die Menge der $S_{[v_0]}$-Formeln $\varphi(x)$ mit $\varphi(\ulcorner\varphi(x)\urcorner) \in A$ in W. Dann gilt für jede $S_{[v_0]}$-Formel $\vartheta_2(x)$:

$$\text{(i)} \quad \textit{Wenn } \vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in W, \textit{ dann } \vartheta_1(\ulcorner\vartheta_2(x)\urcorner) \in W.$$
$$\text{(ii)} \quad \textit{Wenn } \vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in A, \textit{ dann } \vartheta_1(\ulcorner\vartheta_2(x)\urcorner) \in A.$$

[17] Teil b) entspricht Theorem (GT) in Smullyan (1992), S. 7.

[18] Auch die Teile a) und b) des Satzes lassen sich dann zusammenfassen zu dem folgenden Satz: Für jede Menge A von S-Sätzen, für die die Menge der $S_{[v_0]}$-Formeln $\varphi(x)$ mit $\varphi(\ulcorner\varphi(x)\urcorner) \in A$ in W durch eine $S_{[v_0]}$-Formel $\vartheta(x)$ repräsentiert wird, und für jede $S_{[v_0]}$-Formel $\vartheta'(x)$ gilt: $\vartheta'(\ulcorner\vartheta'(x)\urcorner) \in A$ gdw. $\vartheta(\ulcorner\vartheta'(x)\urcorner) \in W$. Diese Abstraktion geht nicht mehr über den Begriff der Repräsentierbarkeit hinaus.

b) Für $A \supseteq W$ repräsentiere die $S_{[v_0]}$-Formel $\vartheta_2(x)$ die Menge der $S_{[v_0]}$-Formeln $\varphi(x)$ mit $\varphi(\ulcorner\varphi(x)\urcorner) \notin A$ in W. Dann gilt:

$$\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin W \text{ und } \vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in A.$$

Beweis. Analog zum Beweis von Satz 2.4. □

Die semantische Version des Beweises des ersten Unvollständigkeitssatzes, die Gödel informell in Gödel (1931), S. 174 f. beschreibt, fällt unter Satz 2.4 b). Dabei wird W interpretiert als die Menge der in $\mathcal{N}$ wahren S_{Ar}-Sätze und A als die Menge der aus einer aufzählbaren Menge PA von in $\mathcal{N}$ wahren S_{Ar}-Sätzen ableitbaren S_{Ar}-Sätze.

Da die Relation der Nicht-Ableitbarkeit aus PA in $\mathrm{Th}(\mathbb{N})$ durch $\neg\bigvee x \mathrm{bew}_{PA}(x,y)$ repräsentiert wird, ergibt die Anwendung von Satz 2.4 b):
$PA \nvdash \neg\bigvee x\mathrm{bew}_{PA}(x,\alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x,\alpha(y))\urcorner))$. So wird an dieser Stelle von Gödel dargestellt, dass dieser Satz nicht ableitbar ist.

Um zu zeigen, dass auch seine Negation nicht ableitbar ist, argumentiert Gödel wie folgt. Dabei sei der Satz $\neg\bigvee x\mathrm{bew}_{PA}(x,\alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x,\alpha(y))\urcorner))$ abgekürzt durch gö_{PA}. Wenn $PA \vdash \neg\text{gö}_{PA}$, dann auch $\mathcal{N} \models \neg\text{gö}_{PA}$. Und damit $\mathcal{N} \not\models \text{gö}_{PA}$, was im Widerspruch zur Annahme $PA \vdash \text{gö}_{PA}$ impliziert. In dieser Argumentation wird die Voraussetzung $PA^{\vdash} \subseteq \mathrm{Th}(\mathbb{N})$ verwendet und die Implikation: Wenn $\mathcal{N} \not\models \text{gö}_{PA}$, dann $PA \vdash \text{gö}_{PA}$. Dies kann wiederum als Anwendung von Satz 2.4 b) gedeutet werden, in dessen Beweis die Äquivalenz: $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin A$ gdw. $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in W$ zentral ist.

Gödels syntaktischer Beweis des ersten Unvollständigkeitssatzes in Gödel (1931), S. 190, der die ω-Widerspruchsfreiheit nicht voraussetzt, fällt unter Satz 2.5 a)(i).[19] Hierzu wird W interpretiert als die Menge der aus einer Menge PA von S_{Ar}-Sätzen ableitbaren S_{Ar}-Sätze, wobei vorausgesetzt wird, dass PA aufzählbar ist und die Menge PA enthält. A wird interpretiert als die Menge $PA^{\vdash} \cup \{\varphi(\ulcorner\varphi(x)\urcorner) | \varphi(x) \text{ ist } S_{\mathrm{Ar}[v_0]}\text{-Formel mit } PA \vdash \bigvee x\mathrm{bew}_{PA}(x,\alpha(\ulcorner\varphi(x)\urcorner))\}$.

Die Anwendung von Satz 2.5 a)(i) für $\vartheta_2(y) = \neg\bigvee x\mathrm{bew}_{PA}(x,\alpha(y))$ ergibt:
Wenn $PA \vdash \neg\bigvee x\mathrm{bew}_{PA}(x,\alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x,\alpha(y))\urcorner))$,

[19] Siehe die Bemerkungen zur Originalformulierung von Gödel auf Seite 48.

dann $PA \vdash \bigvee x\mathrm{bew}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(y))\urcorner))$.
Dies impliziert $PA \nvdash \neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}_{PA}(x, \alpha(y))\urcorner))$.[20]

Rossers Beweis seiner Form des ersten Unvollständigkeitssatzes in Rosser (1936), S. 89 f. fällt ebenfalls unter Satz 2.5 a)(i).[21] Genauer gesagt wird Satz 2.5 a)(i) zweimal angewendet. Hierzu wird W zum einen interpretiert als die Menge der aus einer Menge PA von S_{Ar}-Sätzen ableitbaren S_{Ar}-Sätze, wobei vorausgesetzt wird, dass PA aufzählbar ist und die Menge PA enthält. A wird interpretiert als die Menge
$PA^{\vdash} \cup \{\varphi(\ulcorner\varphi(x)\urcorner)|\varphi(x)$ ist $S_{\mathrm{Ar}[v_0]}$-Formel mit $PA \vdash \bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(\ulcorner\varphi(x)\urcorner))\}$.
Die Anwendung von Satz 2.5 a)(i) ergibt dann wieder
$PA \nvdash \neg\bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(y))\urcorner))$.[22]

Zum anderen wird W interpretiert als die Menge der S_{Ar}-Sätze, deren Negation aus PA ableitbar ist. A wird dazu interpretiert als die Menge
$W \cup \{\varphi(\ulcorner\varphi(x)\urcorner)|\varphi(x)$ ist $S_{\mathrm{Ar}[v_0]}$-Formel mit $PA \vdash \neg\bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(\ulcorner\varphi(x)\urcorner))\}$.
Die Anwendung von Satz 2.5 a)(i) ergibt dann
$PA \nvdash \neg\neg\bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(\ulcorner\neg\bigvee x\mathrm{bew}^{\mathrm{R}}_{PA}(x, \alpha(y))\urcorner))$.

Die Antinomien in Kapitel 1.2 sind rein objektsprachlich formuliert. Um eine solche objektsprachliche Variante von Satz 2.3 formulieren zu können, wird eine Äquivalenz und eine Negation in abstrakten Sprachen benötigt. Die Äquivalenz soll dabei gewährleisten, dass alle Argumentationen der Metasprache, die sich auf die Äquivalenz beziehen, auch in der Objektsprache durchgeführt werden können. An die Negation werden zunächst keine Anforderungen gestellt, da gerade an diesem Punkt mögliche Ansätze zur Vermeidung von Antinomien bestehen. Die Negation wird also zunächst nur als eine Abbildung innerhalb der objektsprachlichen Formeln angesehen, die eine Form von Assoziativität bezüglich der Einsetzung von Termen erfüllt.

Definition 2.6. $S = (F_S, F'_S, T_S, T'_S, e_S, n_S)$ sei eine abstrakte Sprache. W sei eine Menge von S-Sätzen.

---

[20] Gödels Ausgangspunkt ist die Repräsentierung der Relation "die n_1-te Ableitung aus Sätzen aus PA ist keine Ableitung des S_{Ar}-Satzes mit der Gödelnummer n_2" in PA durch eine $S_{\mathrm{Ar}[v_0]}$-Formel (vgl. S. 60).

Die weitere Folgerung, die Gödel an dieser Stelle zieht, lautet in der Formulierung dieser Arbeit: Für alle $n \in \mathbb{N}$ gilt $PA \vdash \neg\mathrm{bew}_{PA}(\boldsymbol{n}, \alpha(\ulcorner\bigwedge x\neg\mathrm{bew}_{PA}(x, \alpha(y))\urcorner))$. Siehe auch Satz 1.31. Dieser Teil bezieht sich auf die spezielle Situation der Aufzählbarkeit der repräsentierbaren Relation und taucht daher in der abstrakten Fassung nicht auf.

[21] Siehe die Bemerkung zur Originalformulierung von Rosser auf Seite 48.

[22] Vgl. die abstrakte Darstellung in Smullyan (1992), S. 76 f.

a) $\leftrightarrow\ : F'_S \times F'_S \to F'_S$ (Äquivalenz zweier Sätze) sei eine Abbildung. $(S, W, \leftrightarrow)$ ist eine *abstrakte Sprache mit Äquivalenz* gdw.:

- für alle S-Sätze φ_1, φ_2 gilt: wenn $\varphi_1 \leftrightarrow \varphi_2 \in W$, dann $\varphi_2 \leftrightarrow \varphi_1 \in W$,
- für alle S-Sätze $\varphi_1, \varphi_2, \varphi_3$ gilt: wenn $\varphi_1 \leftrightarrow \varphi_2 \in W$ und $\varphi_2 \leftrightarrow \varphi_3 \in W$, dann $\varphi_1 \leftrightarrow \varphi_3 \in W$,
- wenn $\varphi_1 \leftrightarrow \varphi_2 \in W$, dann: $\varphi_1 \in W$ gdw. $\varphi_2 \in W$.[23]

b) $(S, W, \leftrightarrow)$ sei eine abstrakte Sprache mit Äquivalenz. $\neg : F_S \to F_S$ (Negation einer Formel) sei eine Abbildung. $(S, W, \leftrightarrow, \neg)$ ist eine *abstrakte Sprache mit Negation* gdw.:

- für alle $S_{[v_0]}$-Formeln $\varphi(x)$ und alle S-Terme t gilt: $\neg(\varphi(x)) \in F_S \setminus F'_S$ und $\neg(\varphi(t)) = \neg(\varphi(x))(t)$.

Die rein objektsprachliche Variante von Satz 2.3 lautet nun:

Satz 2.7.
a) $(S, W, \leftrightarrow)$ sei eine abstrakte Sprache mit Äquivalenz. Wenn es eine $S_{[v_0]}$-Formel $\vartheta_1(x)$ gibt, so dass für jede $S_{[v_0]}$-Formel $\varphi(x)$ gilt

$$\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow \vartheta_1(\ulcorner\varphi(x)\urcorner) \in W,$$

so gilt für jede $S_{[v_0]}$-Formel $\vartheta_2(x)$

$$\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \leftrightarrow \vartheta_1(\ulcorner\vartheta_2(x)\urcorner) \in W.$$

b) $(S, W, \leftrightarrow, \neg)$ sei eine abstrakte Sprache mit Negation. Wenn es eine $S_{[v_0]}$-Formel $\vartheta_2(x)$ gibt, so dass für jede $S_{[v_0]}$-Formel $\varphi(x)$ gilt

$$\neg\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow \vartheta_2(\ulcorner\varphi(x)\urcorner) \in W,$$

so gilt

$$\neg\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \leftrightarrow \vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in W.$$

[23] Die letzte Bedingung wird nur zur Übereinstimmung gewisser Begriffe für Sprachen mit und ohne Äquivalenz benötigt. Vgl. Definition 2.8.

Beweis. Hier ist nichts zu beweisen. □

Angewendet auf die Sprache der Antinomie von Grelling als abstrakte Sprache mit Negation (bezüglich der Menge der in $\mathcal{A}_{\mathrm{Gr}}$ wahren Sätze dieser Sprache), besagt Satz 2.7 b), dass unter der Annahme einer $S_{\mathrm{Gr}[v_0]}$-Formel $het(x)$, so dass für jede $S_{\mathrm{Gr}[v_0]}$-Formel $\varphi(x)$ gilt $\mathcal{A}_{\mathrm{Gr}} \models \neg\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow het(\ulcorner\varphi(x)\urcorner)$, die "antinomische" Aussage

$$\mathcal{A}_{\mathrm{Gr}} \models \neg het(\ulcorner het(x)\urcorner) \leftrightarrow het(\ulcorner het(x)\urcorner)$$

gilt. Entsprechendes gilt für die Antinomie von Richard.

2.1.2 Darstellung mit Fixpunktsatz

Die Lügner-Antinomie, der Satz von Tarski und Satz 1.24 lassen sich als Verfeinerungen der Sätze 2.3 bzw. 2.7 deuten: Die Voraussetzungen der Lügner-Antinomie, des Satzes von Tarski und des Satzes 1.24 besagen, dass die Einsetzungsfunktion (der Name einer Formel wird in die Formel eingesetzt) repräsentierbar ist und die Menge der wahren oder ableitbaren bzw. der nicht wahren oder nicht ableitbaren Sätze repräsentierbar ist oder ein objektsprachlicher Wahrheits- bzw. Falschheitsbegriff adäquat definierbar ist. Dies ist hinreichend dafür, dass auch die Bedingungen der Sätze 2.3 bzw. 2.7 für die jeweiligen Anwendungsfälle erfüllt sind. Die Lügner-Antinomie, den Satz von Tarski und Satz 1.24 betreffend wird jedoch häufig – wie auch in der vorliegenden Arbeit – mit einem Fixpunktsatz argumentiert, der in seiner abstrakten Fassung nach der Einführung des Repräsentierbarkeitsbegriffs für Funktionen formuliert werden kann.

Definition 2.8. a) S sei eine abstrakte Sprache und W eine Menge von S-Sätzen. Eine totale Funktion $f : F_S \to F_S$ heißt *repräsentierbar in* W gdw. es einen $S_{[v_0]}$-Term $t(x)$ gibt, so dass für alle S-Formeln ψ_1, ψ_2:

> $f(\psi_1) = \psi_2$ gdw.
> für alle $S_{[v_0]}$-Formeln $\varphi(x)$ gilt: $\varphi(t(\ulcorner\psi_1\urcorner)) \in W$ gdw. $\varphi(\ulcorner\psi_2\urcorner) \in W$.

$t(x)$ *repräsentiert* f *in* W.

b) $(S, W, \leftrightarrow)$ sei eine abstrakte Sprache mit Äquivalenz. Eine totale Funktion $f : F_S \to F_S$ heißt *repräsentierbar in* $(S, W, \leftrightarrow)$ gdw. es einen $S_{[v_0]}$-Term $t(x)$ gibt, so dass für alle S-Formeln ψ_1, ψ_2:

$$\begin{aligned}&f(\psi_1) = \psi_2 \text{ gdw.}\\ &\text{für alle } S_{[v_0]}\text{-Formeln } \varphi(x) : \varphi(t(\ulcorner\psi_1\urcorner)) \leftrightarrow \varphi(\ulcorner\psi_2\urcorner) \in W.\end{aligned}$$

$t(x)$ *repräsentiert* f *in* S.

Satz 2.9 (Abstrakter Fixpunktsatz). *S sei eine abstrakte Sprache und W eine Menge von S-Sätzen. $f : F_S \to F_S$ sei gegeben durch*

$$f(\varphi(x)) = \begin{cases} \varphi(\ulcorner\varphi(x)\urcorner), & \textit{falls } \varphi(x) \textit{ eine } S_{[v_0]}\textit{-Formel ist,}\\ t^*, & \textit{sonst.}\end{cases}$$

Dabei sei t^* eine beliebige S-Formel.
a) f sei in W repräsentierbar. Dann gibt es zu jeder $S_{[v_0]}$-Formel $\psi(x)$ einen S-Satz χ, so dass: $\chi \in W$ gdw. $\psi(\ulcorner\chi\urcorner) \in W$.
b) $(S, W, \leftrightarrow)$ sei eine abstrakte Sprache mit Äquivalenz. f sei in $(S, W, \leftrightarrow)$ repräsentierbar. Dann gibt es zu jeder $S_{[v_0]}$-Formel $\psi(x)$ einen Fixpunkt, d. h. einen S-Satz χ, so dass $\chi \leftrightarrow \psi(\ulcorner\chi\urcorner) \in W$.

Beweis. Zu a) $\alpha(x)$ repräsentiere f in W. Für jede $S_{[v_0]}$-Formel $\varphi(x)$ gilt: $\psi(\alpha(\ulcorner\varphi(x)\urcorner)) \in W$ gdw. $\psi(\ulcorner\varphi(\ulcorner\varphi(x)\urcorner)\urcorner) \in W$. Dann erfüllt $\chi := \psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))$ die Bedingung, denn mit $\psi(\alpha(x))$ für $\varphi(x)$ erhält man: $\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner)) \in W$ gdw. $\psi(\ulcorner\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))\urcorner) \in W$.
Zu b) $\alpha(x)$ repräsentiere f in $(S, W, \leftrightarrow)$. Für jede $S_{[v_0]}$-Formel $\varphi(x)$ gilt $\psi(\alpha(\ulcorner\varphi(x)\urcorner)) \leftrightarrow \psi(\ulcorner\varphi(\ulcorner\varphi(x)\urcorner)\urcorner) \in W$. Dann erfüllt $\chi := \psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))$ die Bedingung, denn mit $\psi(\alpha(x))$ für $\varphi(x)$ erhält man
$\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner)) \leftrightarrow \psi(\ulcorner\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))\urcorner) \in W$. □

Parallel zur Lügner-Antinomie, zum Satz von Tarski und Satz 1.24 kann nun folgende abstrakte Version bewiesen werden:

Satz 2.10. *S sei eine abstrakte Sprache und W eine Menge von S-Sätzen.*
a) Die Abbildung f aus Satz 2.9 sei in W repräsentierbar. Wenn die Menge W in W repräsentierbar ist, genauer genügt schon: wenn es eine $S_{[v_0]}$-Formel

$\gamma_1(x)$ gibt, so dass für alle S-Sätze φ gilt:

$$\varphi \in W \quad \text{gdw.} \quad \gamma_1(\ulcorner\varphi\urcorner) \in W,$$

so gibt es zu jeder $S_{[v_0]}$-Formel $\gamma_2(x)$ einen S-Satz χ, so dass:

$$\gamma_2(\ulcorner\chi\urcorner) \in W \quad \text{gdw.} \quad \gamma_1(\ulcorner\chi\urcorner) \in W.$$

b) Die Abbildung f aus Satz 2.9 sei in W repräsentierbar. Die Menge $F'_S\backslash W$ der S-Sätze, die nicht in W liegen, ist nicht repräsentierbar in W. Genauer gibt es sogar keine $S_{[v_0]}$-Formel $\gamma_2(x)$, so dass für alle S-Sätze φ gilt:

$$\varphi \in F'_S\backslash W \quad \text{gdw.} \quad \gamma_2(\ulcorner\varphi\urcorner) \in W.$$

c)[24] *$(S, W, \leftrightarrow)$ sei eine abstrakte Sprache mit Äquivalenz. Die Abbildung f aus Satz 2.9 sei in $(S, W, \leftrightarrow)$ repräsentierbar. Wenn es eine $S_{[v_0]}$-Formel $\gamma_1(x)$ gibt, so dass für alle S-Sätze φ gilt*

$$\varphi \leftrightarrow \gamma_1(\ulcorner\varphi\urcorner) \in W,$$

so gibt es zu jeder $S_{[v_0]}$-Formel $\gamma_2(x)$ einen S-Satz χ, so dass:

$$\gamma_2(\ulcorner\chi\urcorner) \leftrightarrow \gamma_1(\ulcorner\chi\urcorner) \in W.$$

d) $(S, W, \leftrightarrow, \neg)$ sei eine abstrakte Sprache mit Negation. Die Abbildung f aus Satz 2.9 sei in $(S, W, \leftrightarrow)$ repräsentierbar. Wenn es eine $S_{[v_0]}$-Formel $\gamma_2(x)$ gibt, so dass für alle S-Sätze φ gilt

$$\neg\varphi \leftrightarrow \gamma_2(\ulcorner\varphi\urcorner) \in W,$$

so gilt

$$\gamma_2(\ulcorner\chi\urcorner) \leftrightarrow \neg\gamma_2(\ulcorner\chi\urcorner) \in W.$$

Beweis. Zu a) $\gamma_1(x)$ sei eine $S_{[v_0]}$-Formel, so dass für alle S-Sätze φ gilt: $\varphi \in W$ gdw. $\gamma_1(\ulcorner\varphi\urcorner) \in W$. χ sei ein S-Satz gemäß Satz 2.9, so dass: $\chi \in W$ gdw. $\gamma_2(\ulcorner\chi\urcorner) \in W$. Nach Einsetzen von χ für φ ergibt sich: $\gamma_2(\ulcorner\chi\urcorner) \in W$ gdw. $\gamma_1(\ulcorner\chi\urcorner) \in W$.

[24] Vgl. Smullyan (1992), Theorem 7, S. 104 f.

Zu b) Angenommen es gibt eine solche Formel $\gamma_2(x)$, so dass für alle S-Sätze φ gilt: $\varphi \in F'_S \setminus W$ gdw. $\gamma_2(\ulcorner\varphi\urcorner) \in W$. χ sei ein S-Satz gemäß Satz 2.9, so dass: $\chi \in F'_S \setminus W$ gdw. $\gamma_2(\ulcorner\chi\urcorner) \in F'_S \setminus W$. Nach Einsetzen von χ für φ ergibt sich: $\gamma_2(\ulcorner\chi\urcorner) \in W$ gdw. $\gamma_2(\ulcorner\chi\urcorner) \in F'_S \setminus W$. Widerspruch.
Zu c) $\gamma_1(x)$ sei eine $S_{[v_0]}$-Formel, so dass für alle S-Sätze φ gilt $\varphi \leftrightarrow \gamma_1(\ulcorner\varphi\urcorner) \in W$. χ sei ein Fixpunkt von $\gamma_2(x)$. Es gilt also $\chi \leftrightarrow \gamma_2(\ulcorner\chi\urcorner) \in W$. Nach Einsetzen von χ für φ ergibt sich $\gamma_2(\ulcorner\chi\urcorner) \leftrightarrow \gamma_1(\ulcorner\chi\urcorner) \in W$.
Zu d) $\gamma_2(x)$ sei eine $S_{[v_0]}$-Formel, so dass für alle S-Sätze φ gilt $\neg\varphi \leftrightarrow \gamma_2(\ulcorner\varphi\urcorner) \in W$. χ sei ein Fixpunkt von $\gamma_2(x)$, es gilt also $\chi \leftrightarrow \gamma_2(\ulcorner\chi\urcorner) \in W$. Mit der gleichen Argumentation wie im Fixpunktsatz erhält man auch $\neg\chi \leftrightarrow \neg\gamma_2(\ulcorner\chi\urcorner) \in W$, da $\neg\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \leftrightarrow \neg\gamma_2(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W$. Nach Einsetzen von χ für φ ergibt sich $\gamma_2(\ulcorner\chi\urcorner) \leftrightarrow \neg\gamma_2(\ulcorner\chi\urcorner) \in W$. $\square$

Wendet man Satz 2.10 c) auf die Sprache der Lügner-Antinomie als abstrakte Sprache mit Negation an und setzt voraus, dass die Selbsteinsetzungsfunktion (der Standardterm einer Formel wird in die Formel eingesetzt) durch einen $S_{\text{Lü}[v_0]}$-Term $\alpha(x)$ in $\text{Th}(\mathcal{A}_{\text{Lü}})$ repräsentiert wird und es eine $S_{\text{Lü}[v_0]}$-Formel $w(x)$ gibt, so dass für jeden $S_{\text{Lü}}$-Satz φ gilt $\mathcal{A}_{\text{Lü}} \models \varphi \leftrightarrow w(\ulcorner\varphi\urcorner)$, so erhält man die "antinomische" Aussage

$$\mathcal{A}_{\text{Lü}} \models \neg w(\ulcorner\chi\urcorner) \leftrightarrow w(\ulcorner\chi\urcorner).$$

Dabei ist χ ein Fixpunkt zu $\neg w(\alpha(x))$, also z. B. der im Fixpunktsatz konstruierte Satz $\neg w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner))$.

Angewandt auf die Sprache der Arithmetik zur Symbolmenge S_{Ar} als abstrakte Sprache mit Negation bezüglich der Menge der in $\mathcal{N}$ wahren Sätze als Menge W, stellt Satz 2.10 a)-d) den Satz von Tarski in den verschiedenen Varianten der Sätze 1.11, 1.12, 1.13 dar. Da $\text{Th}(\mathcal{N})$ Repräsentierungen erlaubt, ist die Funktion f aus Satz 2.9, die auf Grund der Standardnamenbildung durch Gödelisierung berechenbar ist, in $\text{Th}(\mathcal{N})$ repräsentierbar. Die Voraussetzung von Satz 2.10 ist also erfüllt.

Angewandt auf die gleiche Sprache als abstrakte Sprache mit Negation, jetzt jedoch bezüglich der Menge der aus *PA* ableitbaren Sätze als Menge W, stellt Satz 2.10 a)-d) den Satz 1.24 a)-d) dar. Da *PA* Repräsentierungen erlaubt, ist die Funktion f aus Satz 2.9 in *PA* repräsentierbar. Die Voraussetzung von Satz 2.10 ist also erfüllt.

Satz 2.10 und die Sätze 2.3 und 2.7 stehen in engem Zusammenhang. Satz 2.10 kann sozusagen als Verfeinerung der Sätze 2.3 und 2.7 aufgefasst werden, denn aus den Voraussetzungen von Satz 2.10 a)-d) ergeben sich die Voraussetzungen der Sätze 2.3a) und b) bzw. 2.7 a) und b). Aber auch die Konklusion des Satzes 2.10 geht nicht über die der Sätze 2.3 und 2.7 hinaus. Dies soll im Folgenden gezeigt werden. Der $S_{[v_0]}$-Term $\alpha(x)$ repräsentiere die Funktion f aus Satz 2.9 in W.

Zu Satz 2.10 a)
Unter der Voraussetzung von Satz 2.10 a) gilt für jede $S_{[v_0]}$-Formel $\varphi(x)$:

$$\gamma_1(\alpha(\ulcorner\varphi(x)\urcorner)) \in W \quad \text{gdw.} \quad \varphi(\ulcorner\varphi(x)\urcorner) \in W.$$

$\gamma_1(\alpha(x))$ erfüllt also die Bedingung aus Satz 2.3 a). Speziell für die $S_{[v_0]}$-Formel $\gamma_2(\alpha(x))$ für $\varphi(x)$ ergibt sich dies in folgenden zwei Schritten:

$$\begin{aligned}
\gamma_1(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in W &\text{ gdw. } \gamma_1(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W,\\
&\text{ da } \alpha(x) f \text{ in } W \text{ repräsentiert, und:}\\
\gamma_1(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W &\text{ gdw. } \gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in W\\
&\text{ nach Voraussetzung.}
\end{aligned}$$

Die Aussage aus Satz 2.10 a) ergibt sich nun hieraus, da $\alpha(x)$ f in W repräsentiert:

$$\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in W \text{ gdw. } \gamma_2(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W.$$

Zu Satz 2.10 b)
Unter der Voraussetzung von Satz 2.10 b) gilt für jede $S_{[v_0]}$-Formel $\varphi(x)$:

$$\varphi(\ulcorner\varphi(x)\urcorner) \notin W \quad \text{gdw.} \quad \gamma_2(\alpha(\ulcorner\varphi(x)\urcorner)) \in W.$$

$\gamma_2(\alpha(x))$ erfüllt also die Bedingung aus Satz 2.3 b). Speziell für die $S_{[v_0]}$-Formel

$\gamma_2(\alpha(x))$ für $\varphi(x)$ ergibt sich dies in folgenden zwei Schritten:

$$\begin{aligned}\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in W \text{ gdw. } & \gamma_2(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W, \\ & \text{da } \alpha(x) f \text{ in } W \text{ repräsentiert, und:} \\ \gamma_2(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W \text{ gdw. } & \gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \notin W \\ & \text{nach Voraussetzung.}\end{aligned}$$

Der genaue Widerspruch des Beweises von Satz 2.10 b) ergibt sich nun hieraus, da $\alpha(x)$ f in W repräsentiert:

$$\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \notin W \text{ gdw. } \gamma_2(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \notin W.$$

Zu Satz 2.10 c)
Unter der Voraussetzung von Satz 2.10 c) gilt für jede $S_{[v_0]}$-Formel $\varphi(x)$:

$$\gamma_1(\alpha(\ulcorner\varphi(x)\urcorner)) \leftrightarrow \varphi(\ulcorner\varphi(x)\urcorner) \in W.$$

$\gamma_1(\alpha(x))$ erfüllt also die Bedingung aus Satz 2.7 a). Speziell für die $S_{[v_0]}$-Formel $\gamma_2(\alpha(x))$ für $\varphi(x)$ ergibt sich dies in folgenden zwei Schritten:

$$\begin{aligned}& \gamma_1(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \leftrightarrow \gamma_1(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W, \\ & \qquad \text{da } \alpha(x) f \text{ in } W \text{ repräsentiert, und:} \\ & \gamma_1(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \leftrightarrow \gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in W \\ & \qquad \text{nach Voraussetzung.}\end{aligned}$$

Die Aussage aus Satz 2.10 c) ergibt sich nun hieraus, da $\alpha(x)$ f in W repräsentiert:

$$\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \leftrightarrow \gamma_2(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W.$$

Zu Satz 2.10 d)
Unter der Voraussetzung von Satz 2.10 d) gilt für jede $S_{[v_0]}$-Formel $\varphi(x)$:

$$\neg\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow \gamma_2(\alpha(\ulcorner\varphi(x)\urcorner)) \in W.$$

$\gamma_2(\alpha(x))$ erfüllt also die Bedingung aus Satz 2.7 b). Speziell für die $S_{[v_0]}$-Formel

$\gamma_2(\alpha(x))$ für $\varphi(x)$ ergibt sich dies in folgenden zwei Schritten:

$$\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \leftrightarrow \gamma_2(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W,$$

da $\alpha(x) f$ in W repräsentiert, und:

$$\gamma_2(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \leftrightarrow \neg\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in W$$

nach Voraussetzung.

Die Aussage aus Satz 2.10 d) ergibt sich nun hieraus, da $\alpha(x)$ f in W repräsentiert:

$$\neg\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \leftrightarrow \neg\gamma_2(\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner) \in W.$$

Der Unterschied in dem Ansatz der Sätze 2.3 und 2.7 einerseits und dem Ansatz des Satzes 2.10 andererseits kann also aufgefasst werden als Unterschied in der Bezeichnung des Gödelsatzes[25] $\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))$. In den zuerst genannten Sätzen wird der Gödelsatz nicht mit dem Standardnamen bezeichnet, sondern als Wert $\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)$ der Einsetzungsfunktion $\alpha(x)$ betrachtet. Im zuletzt genannten Satz wird der Gödelsatz mit seinem Standardnamen $\ulcorner\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))\urcorner$ bezeichnet. In diesem Sinn haben auch die Argumentationen und Beweise, die über den Fixpunktsatz laufen, die Struktur, die in den Sätzen 2.3 und 2.7 wiedergegeben wird.

Dieser Unterschied soll noch einmal konkret an der Lügner-Antinomie und der Antinomie von Grelling in der "autologisch"-Variante deutlich gemacht werden. Voraussetzung in der Lügner-Antinomie ist eine $S_{\text{Lü}[v_0]}$-Formel $w(x)$, so dass für alle $S_{\text{Lü}}$-Sätze φ gilt $\mathcal{A}_{\text{Lü}} \models \varphi \leftrightarrow w(\ulcorner\varphi\urcorner)$. Der Lügner-Satz kann konstruiert werden, wenn es einen $S_{\text{Lü}[v_0]}$-Term $\alpha(x)$ gibt, der semantisch der Selbsteinsetzung von $S_{\text{Lü}[v_0]}$-Formeln in sich entspricht. Setzt man den Term in die obige Formel ein, so erhält man eine $S_{\text{Lü}[v_0]}$-Formel, die die Voraussetzungen der Antinomie von Grelling in der "autologisch"-Variante erfüllt. Denn es gilt für alle $S_{\text{Lü}[v_0]}$-Formeln $\varphi(x)$: $\mathcal{A}_{\text{Lü}} \models w(\alpha(\ulcorner\varphi(x)\urcorner)) \leftrightarrow w(\ulcorner\varphi(\ulcorner\varphi(x)\urcorner)\urcorner)$, da $\alpha(x)$ die Selbsteinsetzung einer einstelligen Formel repräsentiert, und $\mathcal{A}_{\text{Lü}} \models w(\ulcorner\varphi(\ulcorner\varphi(x)\urcorner)\urcorner) \leftrightarrow \varphi(\ulcorner\varphi(x)\urcorner)$, zusammen: $\mathcal{A}_{\text{Lü}} \models w(\alpha(\ulcorner\varphi(x)\urcorner)) \leftrightarrow \varphi(\ulcorner\varphi(x)\urcorner)$. Der Widerspruch der Antinomie von Grelling ergibt sich also auch hier nach Einsetzen von $\neg w(\alpha(x))$ für $\varphi(x)$ in den einzelnen Schritten:

[25]Zur Definition des Gödelsatzes siehe Seite 46.

$$\mathcal{A}_{\text{Lü}} \models w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner)) \leftrightarrow w(\ulcorner \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))\urcorner)$$

und

$$\mathcal{A}_{\text{Lü}} \models w(\ulcorner \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))\urcorner) \leftrightarrow \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner)),$$

zusammen also

$$\mathcal{A}_{\text{Lü}} \models w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner)) \leftrightarrow \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner)).$$

Demgegenüber wird in der Lügner-Antinomie folgendermaßen argumentiert: Der Schritt

$$\mathcal{A}_{\text{Lü}} \models w(\ulcorner \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))\urcorner) \leftrightarrow \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))$$

wird übernommen. Nun wird auf die rechte Seite der Äquivalenz nochmals angewendet, dass $\alpha(x)$ die Selbsteinsetzung einer einstelligen Formel repräsentiert:

$$\mathcal{A}_{\text{Lü}} \models \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner)) \leftrightarrow \neg w(\ulcorner \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))\urcorner).$$

Zusammen erhält man

$$\mathcal{A}_{\text{Lü}} \models w(\ulcorner \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))\urcorner) \leftrightarrow \neg w(\ulcorner \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))\urcorner).$$

Verwendet wird also in beiden Fällen in gleicher Weise die Eigenschaft des Wahrheitsbegriffs $w(x)$. Im ersten Fall, also der Nachzeichnung der Antinomie von Grelling mit den Voraussetzungen der Lügner-Antinomie, wird der Satz $w(\ulcorner \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))\urcorner)$, der eine Aussage über den Lügner-Satz $\neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))$ darstellt, so durch Anwendung von $\alpha(x)$ umformuliert, dass er eine Aussage über den Lügner-Satz als Funktionswert von $\alpha(x)$ darstellt: $w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))$. Im zweiten Fall, also der eigentlichen Lügner-Antinomie, wird umgekehrt der Satz $\neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))$, der eine Aussage über den Lügner-Satz als Funktionswert von $\alpha(x)$ ist, so durch Anwendung von $\alpha(x)$ umformuliert, dass er eine Aussage über den Lügner-Satz $\neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))$ darstellt: $\neg w(\ulcorner \neg w(\alpha(\ulcorner \neg w(\alpha(x))\urcorner))\urcorner)$.

Der Unterschied zwischen der Antinomie von Grelling in der “autologisch”-Variante und der Lügner-Antinomie kann also zusammenfassend beschrieben werden als Unterschied erstens in der Feinheit der Voraussetzung – die Voraussetzung der Lügner-Antinomie ist mit der Annahme von $w(x)$ und $\alpha(x)$ feiner als die Voraussetzung der Antinomie von Grelling mit der Annahme von *aut* – und zweitens in der Art des Bezugs auf den Lügner-Satz – die Lügner-Antinomie bezieht sich direkt auf ihn, die Antinomie von Grelling bezieht sich auf ihn als Funktionswert.

Die Sätze 2.3 und 2.7 einerseits und Satz 2.10 andererseits stellen also in abstrakter Form die vier Sätze dar, die aus der Lügner-Antinomie, den Antinomien von Grelling und Richard, dem Satz von Tarski, dem ersten Gödelschen Unvollständigkeitssatz und dem Unentscheidbarkeitssatz abstrahiert werden können. Die Abstraktion bezieht sich hierbei auf den Aspekt der Repräsentierbarkeit einer Formelmenge in einer Satzmenge. Abstrahiert wird insbesondere von den Eigenschaften einer Negation in den speziellen Objektsprachen, insbesondere in Bezug auf einen Wahrheitsbegriff.

Die zwei Fassungen der vier abstrakten Sätze in den Sätzen 2.3 und 2.7 einerseits und Satz 2.10 andererseits unterscheiden sich dabei nicht wesentlich, wie die obigen Bemerkungen zeigen.

Der ursprüngliche syntaktische Beweis von Gödel entspricht der Abstraktion in Satz 2.3 a).[26] Dies ist die abstrakte Version der Antinomie von Grelling in Bezug auf den Begriff "autologisch". Gödel selbst zieht einen Vergleich seiner Beweise in erster Linie zur Antinomie von Richard aber auch zur Lügner-Antinomie. Er führt den Vergleich allerdings nicht näher aus.[27] Die abstrakte Version der Antinomie von Richard wird in Satz 2.3 b) wiedergegeben. Dies ist auch die abstrakte Version der Antinomie von Grelling in Bezug auf den Begriff "heterologisch". Wenn die Unterscheidung zwischen einem verneinten Wahrheitsprädikat und einem Falschheitsprädikat bzw. zwischen der Verneinung des Begriffs "autologisch" und dem Begriff "heterologisch" in den abstrakten Sätzen getroffen wird, dann ist der syntaktische Beweis Gödels der Antinomie von Grelling in Bezug auf den Begriff "autologisch" zuzuordnen. Führt man diesen Beweis auf der Basis des Fixpunktsatzes durch, wie in Abschnitt 1.3 geschehen, erhält man als abstrakte Version Satz 2.10 a), der auch die abstrakte Version der Lügner-Antinomie ist. Aus dieser Sichtweise, also mit diesem Beweis als Grundlage, hat der Unvollständigkeitssatz die Form der Lügner-Antinomie.

[26]Siehe Seite 59.
[27]Siehe Gödel (1931), S. 175.

Gödels semantischer Beweis entspricht der Abstraktion in Satz 2.4 b).[28] Dieser Satz wiederum ist abgeleitet aus Satz 2.3 b), also der abstrakten Version der Antinomie von Richard bzw. der Antinomie von Grelling bezogen auf den Begriff "heterologisch".

Der ursprüngliche Beweis von Tarski verläuft nach meiner Deutung auf Seite 32 gemäß den Abstraktionen in Satz 2.10. Tarski[29] selbst bezieht sich in seiner Konstruktion eines Widerspruchs ausdrücklich auf die Lügner-Antinomie, die auch der Abstraktion in Satz 2.10 entspricht. Er beschreibt die Idee seines Beweises als Analogie zur Lügner-Antinomie. Der Beweis selbst weist allerdings Züge des Ansatzes auf, der ohne den Fixpunktsatz vorgeht.

Als Ergebnis lässt sich festhalten, dass die Lügner-Antinomie, die Antinomien von Grelling und Richard und die verschiedenen Beweise des Satzes von Tarski, des ersten Gödelschen Unvollständigkeitssatzes und des Unentscheidbarkeitssatzes in wesentlichen Zügen durch die abstrakten Sätze 2.3 a) und b) wiedergegeben werden. Dies ist umso bemerkenswerter, als dass diese beiden Sätze direkte, auf die abstrakte Sprache übertragene Fassungen der Antinomie von Grelling sind, einmal bezogen auf den Begriff "heterologisch" und einmal bezogen auf den Begriff "autologisch".

Satz 2.10 a) und b) lassen sich parallel zu den Sätzen 2.4 und 2.5 allgemein für Teil- oder Obermengen von W ausdrücken.

Es werden jeweils nur die den Teilsätzen a)(i) 2.4 und a)(i) 2.5 entsprechenden Teile wiedergegeben. Nur für diese finden sich Anwendungsfälle unter den hier angegebenen Beweisen zu Sätzen der Logik.

Satz 2.11. *S sei eine abstrakte Sprache und W eine Menge von S-Sätzen. Die Abbildung f aus Satz 2.9 sei in W repräsentierbar.*
a) Die $S_{[v_0]}$-Formel $\gamma_1(x)$ repräsentiere die Menge $A \subseteq W$ in W. Dann gibt es zu jeder $S_{[v_0]}$-Formel $\gamma_2(x)$ einen S-Satz χ, so dass:

$$\textit{Wenn } \gamma_1(\ulcorner\chi\urcorner) \in W, \textit{ dann } \gamma_2(\ulcorner\chi\urcorner) \in W.$$

b) Die $S_{[v_0]}$-Formel $\gamma_1(x)$ repräsentiere die Menge $A \supseteq W$ in W. Dann gibt es

[28] Siehe Seite 62.
[29] Siehe Tarski (1935), S. 522 ff.

zu jeder $S_{[v_0]}$-Formel $\gamma_2(x)$ einen S-Satz χ, so dass:

$$\textit{Wenn } \gamma_2(\ulcorner\chi\urcorner) \in W, \textit{ dann } \gamma_1(\ulcorner\chi\urcorner) \in W.$$

Beweis. Zu a) Nach Voraussetzung gilt für alle S-Sätze φ: $\varphi \in A$ gdw. $\gamma_1(\ulcorner\varphi\urcorner) \in W$. χ sei ein S-Satz gemäß Satz 2.9, so dass:
$\chi \in W$ gdw. $\gamma_2(\ulcorner\chi\urcorner) \in W$. Nach Einsetzen von χ für $\varphi(x)$ ergibt sich:

$$\chi \in A \quad \text{gdw.} \quad \gamma_1(\ulcorner\chi\urcorner) \in W.$$

Nach Voraussetzung gilt $A \subseteq W$. Damit folgt a).
Zu b) Analog. □

Die Version des ersten Gödelschen Unvollständigkeitssatzes, die Tarski im Nachwort zu Tarski (1935), S. 543 f. beweist, bzw. Satz 1.21 fallen unter Satz 2.11 a). Dabei wird W interpretiert als die Menge der in $\mathcal{N}$ wahren S_{Ar}-Sätze und A als die Menge der aus einer aufzählbaren Menge PA von in $\mathcal{N}$ wahren S_{Ar}-Sätzen ableitbaren S_{Ar}-Sätze. Da die Relation der Ableitbarkeit aus PA in $\mathrm{Th}(\mathcal{N})$ durch $\bigvee x \mathrm{bew}_{PA}(x, y)$ repräsentiert wird, ergibt die Anwendung von Satz 2.11 a): Wenn $\mathcal{N} \models \bigvee x \mathrm{bew}_{\Phi}(x, \ulcorner\chi\urcorner)$, dann $\mathcal{N} \models \neg\bigvee x \mathrm{bew}_{\Phi}(x, \ulcorner\chi\urcorner)$. Es folgt also $\mathcal{N} \not\models \bigvee x \mathrm{bew}_{\Phi}(x, \ulcorner\chi\urcorner)$, was auch $\mathcal{N} \nvdash \chi$ bedeutet. Die weiteren Folgerungen im Unvollständigkeitssatz nach Tarski sind möglich, da in diesem speziellen Fall $\gamma_1(\ulcorner\chi\urcorner) \in W$ gdw. $\gamma_2(\ulcorner\chi\urcorner) \notin W$. Mit $\mathcal{N} \not\models \bigvee x \mathrm{bew}_{\Phi}(x, \ulcorner\chi\urcorner)$ gilt also $\mathcal{N} \models \neg\bigvee x \mathrm{bew}_{\Phi}(x, \ulcorner\chi\urcorner)$ und über die Fixpunkteigenschaft $\mathcal{N} \models \chi$ und $\mathcal{N} \nvdash \neg\chi$.

Die Sätze 1.31 a) und c) können als Einsetzungsfälle von Satz 2.11 b) betrachtet werden.[30]

2.2 Vergleich mit der abstrakten Beschreibung von Serény

Serény[31] behandelt in seiner Arbeit ein dem vorliegenden Kapitel ähnliches Thema. Er untersucht den Zusammenhang zwischen der Lügner-Antinomie, dem

[30] Vgl. S. 62. Dort werden die ursprünglichen Beweise dieser Sätze von Gödel bzw. Rosser als Spezialfälle von Satz 2.5 dargestellt. Parallel dazu verläuft die Darstellung der Sätze 1.31 a) und c) als Spezialfälle von Satz 2.11 b).

[31] Vgl. Serény (2003).

Satz von Tarski, dem Unvollständigkeits- und dem Unentscheidbarkeitssatz, indem er eine abstrakte Sprache verwendet und Antinomie sowie Sätze samt Beweisen in dieser Sprache darstellt. Die abstrakte Antinomie und die abstrakten Sätze fallen wiederum unter ein allgemeines Schema, das Serény "Theorem 1 (Generalized Liar Theorem)"[32] nennt.

Die Lügner-Antinomie hat bei Serény eine andere Form als in der vorliegenden Arbeit. Sie entspricht der hier als Antinomie von Grelling wiedergegebenen Antinomie. In der abstrakten Formulierung[33] hat die Antinomie bei Serény konsequenter Weise die Form von Satz 2.3 b), unter den die Antinomie von Grelling in der vorliegenden Arbeit fällt. Dieser Unterschied erklärt sich dadurch, dass Serény nicht zwischen der Verwendung eines Wahrheits- und der Verwendung eines Falschheitsprädikats unterscheidet. Er benutzt damit auch nicht zwei unterschiedliche abstrakte Sätze für die Verwendung eines Wahrheitsprädikats einerseits und die Verwendung eines Falschheitsprädikats andererseits. Weiter unterscheidet Serény nicht zwischen einem Ansatz mit Fixpunktsatz und einem Ansatz ohne Fixpunktsatz.

Eine Verallgemeinerung der abstrakten Lügner-Antinomie, die er in Theorem 1 (Generalized Liar Theorem)[34] formuliert, liefert Serény das Schema, unter das die abstrakten Formen der Antinomie und der Sätze fallen. Theorem 1 (Generalized Liar Theorem) entspricht im Wesentlichen Satz 2.4 b) und Satz 2.5 b). Übertragen auf die Bezeichnungsweisen der vorliegenden Arbeit lautet das Theorem von Serény:

> S sei eine abstrakte Sprache und V eine Menge von S-Sätzen. Die Abbildung f aus Satz 2.9 sei in V repräsentierbar. B sei eine Menge von S-Sätzen, so dass $F'_S \setminus B$ in V repräsentierbar ist. Dann gibt es eine $S_{[v_0]}$-Formel $\vartheta_2(x)$, so dass $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin B$ gdw. $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in V$.

Nun werden bei Serény der Satz von Tarski, der Unvollständigkeits- und der Unentscheidbarkeitssatz abstrakt formuliert und dann als Spezialfälle des obigen Satzes interpretiert. Dies geschieht in Theorem 2(i)-(iii)[35]. Übertragen auf die Bezeichnungsweisen der vorliegenden Arbeit lautet das Theorem:

[32] Siehe Serény (2003), S. 12.
[33] Siehe Proposition (Liar Theorem), Serény (2003), S. 10.
[34] Siehe Serény (2003), S. 12.
[35] Siehe Serény (2003), S. 15 f.

S sei eine abstrakte Sprache und W, A seien Mengen von S-Sätzen. $(S, W, \leftrightarrow, \neg)$ sei eine abstrakte Sprache mit Negation.

(i)(a) Für die Negation $\neg$ gelte folgende zusätzliche Voraussetzung: Für alle S-Sätze φ ist $\varphi \in A$ gdw. $\neg\neg\varphi \in A$.
Die Abbildung f aus Satz 2.9 sei in A repräsentierbar. A sei in A repräsentierbar. Dann gibt es eine $S_{[v_0]}$-Formel $\vartheta_1(x)$, so dass $\neg\vartheta_1(\ulcorner\neg\vartheta_1(x)\urcorner) \in A$ gdw. $\neg\neg\vartheta_1(\ulcorner\neg\vartheta_1(x)\urcorner) \in A$.[36]

(i)(b) Für die Negation $\neg$ gelte folgende zusätzliche Voraussetzung: Für alle S-Sätze φ ist $\varphi \in W$ gdw. $\neg\varphi \notin W$.
Die Abbildung f aus Satz 2.9 sei in W repräsentierbar. A sei in W repräsentierbar. Dann gibt es eine $S_{[v_0]}$-Formel $\vartheta_1(x)$, so dass $\neg\vartheta_1(\ulcorner\neg\vartheta_1(x)\urcorner) \notin A$ gdw. $\neg\vartheta_1(\ulcorner\neg\vartheta_1(x)\urcorner) \in W$.

(ii) Für die Negation $\neg$ gelte folgende zusätzliche Voraussetzung: Für alle S-Sätze φ ist $\varphi \in W$ gdw. $\neg\varphi \notin W$.
Die Abbildung f aus Satz 2.9 sei in W repräsentierbar. Dann ist W nicht in W repräsenierbar. Denn andernfalls gäbe es eine $S_{[v_0]}$-Formel $\vartheta_1(x)$, so dass $\neg\vartheta_1(\ulcorner\neg\vartheta_1(x)\urcorner) \notin W$ gdw. $\neg\vartheta_1(\ulcorner\neg\vartheta_1(x)\urcorner) \in W$.

(iii) Die Abbildung f aus Satz 2.9 sei in A repräsentierbar. Dann ist $F'_S \setminus A$ nicht in A repräsentierbar. Denn andernfalls gäbe es eine $S_{[v_0]}$-Formel $\vartheta_1(x)$, so dass $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin A$ gdw. $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in A$.

Dabei stellen (i)(a) und (i)(b) abstrakte Formen des Unvollständigkeitssatzes dar. (a) bezieht sich auf die syntaktische Form, (b) bezieht sich auf die semantische Form des Beweises von Gödel. (ii) stellt die abstrakte Form des Satzes von Tarski und (iii) den Unentscheidbarkeitssatz nach Rosser dar.

Ein wesentlicher Unterschied zu den abstrakten Fassungen der jeweiligen Sätze in der vorliegenden Arbeit besteht in Bezug auf die syntaktische Form des Unvollständigkeitssatzes. Serény muss hier für die Negation voraussetzen, dass für alle S-Sätze φ gilt: $\varphi \in A$ gdw. $\neg\neg\varphi \in A$. Dafür kann er alle abstrakten Sätze unter einem Satz (Generalized Liar Theorem) zusammenfassen, während in der vorliegenden Arbeit die folgenden zwei Sätze, formuliert analog zu Serénys Generalized Liar Theorem, verwendet werden:[37]

[36] Vgl. auch Smullyan (1992), Theorem 1, S. 59 und das Korollar hierzu, S. 60.
[37] Die Sätze sind Übertragungen der Sätze 2.4 und 2.5.

S sei eine abstrakte Sprache und V eine Menge von S-Sätzen. Die Abbildung f aus Satz 2.9 sei in V repräsentierbar.

(i) B sei eine Menge von S-Sätzen, so dass B in V repräsentierbar ist. Dann gibt es eine $S_{[v_0]}$-Formel $\vartheta_1(x)$, so dass für alle $S_{[v_0]}$-Formeln $\vartheta_2(x)$ gilt: $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in B$ gdw. $\vartheta_1(\ulcorner\vartheta_2(x)\urcorner) \in V$.

(ii) B sei eine Menge von S-Sätzen, so dass $F'_S \setminus B$ in V repräsentierbar ist. Dann gibt es eine $S_{[v_0]}$-Formel $\vartheta_2(x)$, so dass $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \notin B$ gdw. $\vartheta_2(\ulcorner\vartheta_2(x)\urcorner) \in V$.

Um den Beweis nach Serény zu deuten, sei V interpretiert als die Menge der S_{Ar}-Sätze, deren Negation aus einer Menge PA von S_{Ar}-Sätzen ableitbar ist. B sei interpretiert als $PA^{\vdash}$. Dabei wird vorausgesetzt, dass PA aufzählbar und ω-widerspruchsfrei ist und die Menge PA enthält. Da die Relation der Ableitbarkeit aus PA in PA durch $\bigvee x \mathrm{bew}_{PA}(x, y)$ repräsentiert wird und: $PA \vdash \neg\neg\bigvee x \mathrm{bew}_{PA}(x, y)$ gdw. $PA \vdash \bigvee x \mathrm{bew}_{PA}(x, y)$, wird diese Relation in V durch $\neg\bigvee x \mathrm{bew}_{PA}(x, y)$ repräsentiert. Die Anwendung des abstrakten Satzes von Serény ergibt also $PA \vdash \neg\bigvee x \mathrm{bew}_{PA}(x, \alpha(\ulcorner\neg\bigvee x \mathrm{bew}_{PA}(x, \alpha(y))\urcorner))$ gdw. $PA \vdash \neg\neg\bigvee x \mathrm{bew}_{PA}(x, \alpha(\ulcorner\neg\bigvee x \mathrm{bew}_{PA}(x, \alpha(y))\urcorner))$.

Gödels syntaktischer Beweis des Unvollständigkeitssatzes benutzt nicht die Voraussetzung, dass für alle S_{Ar}-Sätze φ gilt: $PA \vdash \varphi$ gdw. $PA \vdash \neg\neg\varphi$.

2.3 Eine abstrakte Version des zweiten Gödelschen Unvollständigkeitssatzes

Da der Beweis des zweiten Unvollständigkeitssatzes von Gödel so formuliert werden kann, dass er im Kern[38] nichts anderes als der Beweis des ersten Unvollständigkeitssatzes auf der Basis von PA ist, kann auch der zweite Unvollständigkeitssatz als Anwendung von Satz 2.10 a) verstanden werden. Als Metasprache wird in diesem Fall die Sprache der Peano-Arithmetik erster Stufe, bzw. eine widerspruchsfreie, aufzählbare Erweiterung hiervon, benutzt.

Es wird nun untersucht, inwieweit sich der zweite Unvollständigkeitssatz auf abstrakte Sprachen im Sinn des Abschnitts 2.1 übertragen lässt.

[38] Vgl. den Beweis zu Satz 1.32.

$(S, W, \leftrightarrow)$ sei eine abstrakte Sprache mit Äquivalenz. In der abstrakten Sprache wird nun zusätzlich zur Äquivalenz eine zweite, schwächere Verknüpfung zwischen Sätzen benötigt. Für alle S-Sätze φ_1, φ_2 sei $\varphi_1 \rightarrow \varphi_2$ ein S-Satz. Die Relation $\rightarrow$ sei schwächer als die Relation $\leftrightarrow$: Wenn $\varphi_1 \leftrightarrow \varphi_2 \in W$, dann auch $\varphi_1 \rightarrow \varphi_2 \in W$. Ferner gelte für alle S-Sätze $\varphi_1, \varphi_2, \varphi_3$: Wenn $\varphi_1 \rightarrow \varphi_2 \in W$ und $\varphi_2 \rightarrow \varphi_3 \in W$, dann $\varphi_1 \rightarrow \varphi_3 \in W$.

Satz 2.12. *Die Abbildung f aus Satz 2.9 sei in W repräsentierbar. $\gamma_1(x)$ sei eine $S_{[v_0]}$-Formel, so dass für alle S-Sätze φ: Wenn $\varphi \in W$, dann $\gamma_1(\ulcorner\varphi\urcorner) \in W$. Für $\gamma_1(x)$ wird nun angenommen, dass diese Bedingung und die Regel des Modus ponens auch aus Sicht von W gelten. Genauer, dass für alle S-Sätze φ, ψ:*
A1: $\gamma_1(\ulcorner\varphi\urcorner) \rightarrow \gamma_1(\ulcorner\gamma_1(\ulcorner\varphi\urcorner)\urcorner) \in W$
A2: Wenn $\gamma_1(\ulcorner\varphi \leftrightarrow \psi\urcorner) \in W$, dann $\gamma_1(\ulcorner\varphi\urcorner) \leftrightarrow \gamma_1(\ulcorner\psi\urcorner) \in W$.

Dann gilt für jede $S_{[v_0]}$-Formel $\gamma_2(x)$ und einen Fixpunkt χ von $\gamma_2(x)$

$$\gamma_1(\ulcorner\gamma_2(\ulcorner\chi\urcorner)\urcorner) \rightarrow \gamma_1(\ulcorner\gamma_1(\ulcorner\chi\urcorner)\urcorner) \in W.$$

Beweis. Wegen A1 gilt $\gamma_1(\ulcorner\chi\urcorner) \rightarrow \gamma_1(\ulcorner\gamma_1(\ulcorner\chi\urcorner)\urcorner) \in W$. Da χ Fixpunkt zu $\gamma_2(x)$ ist, gilt $\gamma_1(\ulcorner\gamma_2(\ulcorner\chi\urcorner) \leftrightarrow \chi\urcorner) \in W$. Und nach A2: $\gamma_1(\ulcorner\gamma_2(\ulcorner\chi\urcorner)\urcorner) \leftrightarrow \gamma_1(\ulcorner\chi\urcorner) \in W$. Insgesamt erhält man also $\gamma_1(\ulcorner\gamma_2(\ulcorner\chi\urcorner)\urcorner) \rightarrow \gamma_1(\ulcorner\gamma_1(\ulcorner\chi\urcorner)\urcorner) \in W$. □

Damit ist Satz 2.9 a) auf der Basis von W durchgeführt. Mit den vorhandenen Mitteln kann die Argumentation des zweiten Unvollständigkeitssatzes nicht weiter übertragen werden. Es müsste für geeignete Wahl von $\gamma_2(x)$ von $\gamma_1(\ulcorner\gamma_2(\ulcorner\chi\urcorner)\urcorner) \rightarrow \gamma_1(\ulcorner\gamma_1(\ulcorner\chi\urcorner)\urcorner) \in W$ auf $\gamma_1(\ulcorner\gamma_2(\ulcorner\chi\urcorner)\urcorner) \rightarrow \gamma_1(\ulcorner\bot\urcorner) \in W$ geschlossen werden. Dann gilt über die Fixpunkteigenschaft auch $\gamma_1(\ulcorner\chi\urcorner) \rightarrow \gamma_1(\ulcorner\bot\urcorner) \in W$. Von hier müsste auf $\gamma_2(\ulcorner\bot\urcorner) \rightarrow \gamma_2(\ulcorner\chi\urcorner) \in W$ geschlossen werden.

2.4 Eine abstrakte Version des Satzes von Tarski für Terme

In Satz 1.16 in Abschnitt 1.3.2 wird der Satz von Tarski von der Frage nach der Ausdrückbarkeit der Wahrheit auf die Frage nach der Ausdrückbarkeit der

Bezeichnungsrelation übertragen. Statt der semantischen Zuordnung des Wahrheitswertes eines Satzes zu diesem Satz wird die semantische Zuordnung eines von einem Term bezeichneten Objekts des Universums zu diesem Term betrachtet. Auch Satz 1.16 kann für abstrakte Sprachen formuliert werden. Zuvor wird die Definition einer abstrakten Sprache für diesen Abschnitt entsprechend abgeändert und der Fixpunktsatz formuliert.

Definition 2.13. Eine *abstrakte Sprache* S ist ein Tupel $(T_S, T'_S, F'_S, e_S, n_S, \equiv)$ aus

- einer Menge T_S (Menge der S-Terme),
- einer Menge T'_S (Menge der geschlossenen S-Terme) mit $T'_S \subseteq T_S$,
- einer Menge F'_S (Menge der S-Sätze),
- einer Abbildung $e_S : (T_S \backslash T'_S) \times T_S \to T_S$ (Einsetzung eines Terms u in einen Term $t(x)$ mit genau einer freien Variable, $t(u)$)
  - mit $e_S[(T_S \backslash T'_S) \times T'_S] \subseteq T'_S$ (die Einsetzung eines geschlossenen Terms in einen Term mit genau einer freien Variable ist ein geschlossener Term),
  - $e_S[(T_S \backslash T'_S) \times (T_S \backslash T'_S)] \subseteq T_S \backslash T'_S$ (die Einsetzung eines Terms mit genau einer freien Variable in einen Term mit genau einer freien Variable ist ein Term mit genau einer freien Variable)
  - und $e_S(e_S(t_1, t_2), t_3) = e_S(t_1, e_S(t_2, t_3)))$ für alle $t_1, t_2 \in T_S \backslash T'_S, t_3 \in T_S$,
- einer injektiven Abbildung $n_S : T_S \to T'_S$ (Standardterm oder Standardname eines Terms t, $\ulcorner t \urcorner$),
- einer Abbildung $\equiv: T'_S \times T'_S \to F'_S$ (Identität).

Die Menge der $S_{[v_0]}$-Terme sei die Menge $T_S \backslash T'_S$.

Die folgende Defintion ist die Parallele zur Definition einer abstrakten Sprache mit Äquivalenz bzw. mit Negation in Abschnitt 2.1.

Definition 2.14. S=$(T_S, T'_S, F'_S, e_S, n_S, \equiv)$ sei eine abstrakte Sprache. W sei eine Menge von S-Sätzen.
a) (S, W) ist eine *abstrakte Sprache mit Identität* gdw.:

- für alle $t_1, t_2 \in T'_S, v(x) \in T_S \backslash T'_S$ gilt: wenn $t_1 \equiv t_2 \in W$, dann $v(t_1) \equiv v(t_2) \in W$,

- für alle S-Terme t_1, t_2 gilt: wenn $\ulcorner t_1 \urcorner \equiv \ulcorner t_2 \urcorner \in W$, dann $t_1 = t_2$,
- für alle geschlossenen S-Terme t_1, t_2 gilt: wenn $t_1 \equiv t_2 \in W$, dann $t_2 \equiv t_1 \in W$,
- für alle geschlossenen S-Terme t_1, t_2, t_3 gilt: wenn $t_1 \equiv t_2 \in W$ und $t_2 \equiv t_3 \in W$, dann $t_1 \equiv t_3 \in W$.

b) (S, W) sei eine abstrakte Sprache mit Identität. $+1 : T_S \to T_S$ sei eine Abbildung. $(S, W, +1)$ ist eine *abstrakte Sprache mit Nachfolgerfunktion* gdw.:

- für alle $S_{[v_0]}$-Terme $t(x)$ und alle geschlossenen S-Terme v gilt: $t(x) + 1 \in T_S \setminus T'_S$ und $t(v) + 1 = (t(x) + 1)(v)$.

Definition 2.15. S sei eine abstrakte Sprache und W sei ein Menge von S-Sätzen. Eine (partielle) Funktion $f : T_S \to T_S$ heißt *repräsentierbar in* W gdw. es einen $S_{[v_0]}$-Term $t(x)$ gibt, so dass für alle S-Terme v_1, v_2:

$$\begin{aligned} &v_1 \in f^{-1}[T_S] \text{ und } f(v_1) = v_2 \text{ gdw.} \\ &t(\ulcorner v_1 \urcorner) \equiv \ulcorner v_2 \urcorner \in W. \end{aligned}$$

$t(x)$ *repräsentiert* f *in* S.

Diese Definition der Repräsentierbarkeit einer partiellen Funktion ist stärker als Definition 1.6 für den konkreten Anwendungsfall in Abschnitt 1.3.1. Denn hier wird, übertragen auf den konkreten Fall, nicht nur für n, für die die partielle Funktion definiert ist, gefordert, dass $\Phi \vdash \dot{\bigvee} x \varphi(\boldsymbol{n}, x)$, sondern allgemein. Ohne diese Verstärkung hätte die Konklusion in Satz 2.17 a) eine andere Form. Hierauf wird im Anschluss an Satz 2.17 kurz eingegangen.

Lemma 2.16. (S, W) *sei eine abstrakte Sprache mit Identität.* $f : T_S \to T_S$ *sei gegeben durch*

$$f(t(x)) = \begin{cases} t(\ulcorner t(x) \urcorner), & \textit{falls } t(x) \textit{ ein } S_{[v_0]}\textit{-Term ist,} \\ t^*, & \textit{sonst.} \end{cases}$$

Dabei sei t^* *ein beliebiger* S*-Term.*
f *sei in* W *repräsentierbar. Dann gibt es zu jedem* $S_{[v_0]}$*-Term* $v(x)$ *einen Fixpunkt, d. h. einen geschlossenen* S*-Term* u*, so dass* $u \equiv v(\ulcorner u \urcorner) \in W$.

Beweis. Siehe den Beweis zu Lemma 1.15. $\square$

Satz 2.17. *(S, W) sei eine abstrakte Sprache mit Identität.*
a) Die Abbildung f aus Lemma 2.16 sei in W repräsentierbar. Wenn die partielle Funktion g von der Menge der S-Terme in sich, die gegeben ist durch

$$g(t) = v, \text{ falls } t \in T'_S \text{ und } t \equiv \ulcorner v \urcorner \in W \text{ für einen } S\text{-Term } v,$$

durch einen $S_{[v_0]}$-Term $s_1(x)$ repräsentierbar ist, so gibt es zu jedem $S_{[v_0]}$-Term $s_2(x)$ einen S-Term u, so dass:

$$\text{für alle } S\text{-Terme } q: \quad s_2(\ulcorner u \urcorner) \equiv \ulcorner q \urcorner \in W \text{ gdw. } s_1(\ulcorner u \urcorner) \equiv \ulcorner q \urcorner \in W.$$

b) p sei eine Abbildung von der Menge der S-Terme in sich, so dass $p(t) \neq t$ für alle S-Terme t. Die Funktion f von der Menge der S-Terme in sich, die gegeben ist durch

$$f(t) = \begin{cases} p(v), & \text{falls } t \in T'_S \text{ und } t \equiv \ulcorner v \urcorner \in W \text{ für einen } S\text{-Term } v \\ t^*, & \text{sonst,} \end{cases}$$

ist nicht repräsentierbar.

c) Wenn es einen $S_{[v_0]}$-Term $s_1(x)$ gibt, so dass für alle geschlossenen S-Terme t gilt

$$t \equiv s_1(\ulcorner t \urcorner) \in W,$$

so gibt es zu jedem $S_{[v_0]}$-Term $s_2(x)$ einen geschlossenen S-Term u, so dass

$$s_2(\ulcorner u \urcorner) \equiv s_1(\ulcorner u \urcorner) \in W.$$

d) Wenn es einen $S_{[v_0]}$-Term $s_2(x)$ gibt, so dass für alle geschlossenen S-Terme t gilt

$$t + 1 \equiv s_2(\ulcorner t \urcorner) \in W,$$

so gilt

$$s_2(\ulcorner u \urcorner) \equiv s_2(\ulcorner u \urcorner) + 1 \in W.$$

Beweis. Zu a) $s_1(x)$ sei ein $S_{[v_0]}$-Term, so dass für alle geschlossenen S-Terme t gilt: $t \equiv \ulcorner q \urcorner \in W$ gdw. $s_1(\ulcorner t \urcorner) \equiv \ulcorner q \urcorner \in W$, für alle S-Terme q. u sei ein S-Term nach Lemma 2.16, so dass für alle S-Terme q: $u \equiv \ulcorner q \urcorner \in W$ gdw.

$s_2(\ulcorner u \urcorner) \equiv \ulcorner q \urcorner \in W$. Nach Einsetzen von u für t ergibt sich: $s_2(\ulcorner u \urcorner) \equiv \ulcorner q \urcorner \in W$ gdw. $s_1(\ulcorner u \urcorner) \equiv \ulcorner q \urcorner \in W$, für alle S-Terme q.
Zu b) $s_2(x)$ sei ein $S_{[v_0]}$-Term, so dass für alle geschlossenen S-Terme t und alle S-Terme q gilt:

$$\begin{aligned} &[q = p(v) \text{ und } t \equiv \ulcorner v \urcorner \in W \text{ für einen } S\text{-Term } v \text{ oder} \\ &\qquad q = t^* \text{ und } t \equiv \ulcorner v \urcorner \notin W \text{ für alle } S\text{-Terme } v] \\ &\qquad\qquad \text{gdw. } s_2(\ulcorner t \urcorner) \equiv \ulcorner q \urcorner \in W. \end{aligned}$$

u sei ein S-Term nach Lemma 2.16, so dass für alle S-Terme q: $u \equiv \ulcorner q \urcorner \in W$ gdw. $s_2(\ulcorner u \urcorner) \equiv \ulcorner q \urcorner \in W$. Nach Einsetzen von u für t ergibt sich für alle S-Terme q:

$$\begin{aligned} &[q = p(v) \text{ und } u \equiv \ulcorner v \urcorner \in W \text{ für einen } S\text{-Term } v \text{ oder} \\ &\qquad q = t^* \text{ und } u \equiv \ulcorner v \urcorner \notin W \text{ für alle } S\text{-Terme } v] \\ &\qquad\qquad \text{gdw. } u \equiv \ulcorner q \urcorner \in W. \end{aligned}$$

Widerspruch.
Zu c) $s_1(x)$ sei ein $S_{[v_0]}$-Term, so dass für alle geschlossenen S-Terme t gilt: $t \equiv s_1(\ulcorner t \urcorner) \in W$. u sei ein Fixpunkt zu $s_2(x)$. Es gilt also $u \equiv s_2(\ulcorner u \urcorner) \in W$. Nach Einsetzen von u für t ergibt sich $s_2(\ulcorner u \urcorner) \equiv s_1(\ulcorner u \urcorner) \in W$.
Zu d) $s_2(x)$ sei ein $S_{[v_0]}$-Term, so dass für alle geschlossenen S-Terme t gilt: $t+1 \equiv s_2(\ulcorner t \urcorner) \in W$. u sei ein Fixpunkt zu $s_2(x)$. Es gilt also $u \equiv s_2(\ulcorner u \urcorner) \in W$. Mit der gleichen Argumentation wie in Lemma 2.16 erhält man auch $u + 1 \equiv s_2(\ulcorner u \urcorner) + 1 \in W$, da $s_2(\alpha(\ulcorner s_2(\alpha(x)) \urcorner)) + 1 \equiv s_2(\ulcorner s_2(\alpha(\ulcorner s_2(\alpha(x)) \urcorner)) \urcorner) + 1 \in W$. Nach Einsetzen von u für t ergibt sich $s_2(\ulcorner u \urcorner) \equiv s_2(\ulcorner u \urcorner) + 1 \in W$. □

Ohne die Verstärkung der Definition der Repräsentierbarkeit einer Funktion würde die Konklusion in Teil a) des obigen Satzes lauten:

$$\text{für alle } S\text{-Terme } q: \quad s_2(\ulcorner u \urcorner, \ulcorner q \urcorner) \in W \text{ gdw. } s_1(\ulcorner u \urcorner, \ulcorner q \urcorner) \in W.$$

Hierzu müssten zweistellige Formeln in die abstrakte Sprache eingeführt werden. $s_1(x_0, x_1)$, $s_2(x_0, x_1)$ wären dann zweistellige Formeln, von denen nicht angenommen werden kann, dass $\bigwedge x_0 \dot{\bigvee} x_1 s_1(x_0, x_1) \in W$, $\bigwedge x_0 \dot{\bigvee} x_1 s_2(x_0, x_1) \in W$.

Parallel zu Satz 2.10 kann auch Satz 2.17 in einer Form wiedergegeben werden, die an die Antinomie von Grelling, übertragen auf einstellige Terme, angelehnt ist. Teil a) lautet dann z. B.: Wenn die partielle Funktion g von der Menge der S-Terme in sich, die gegeben ist durch

$$g(t) = v, \text{ falls } t \in T_S \setminus T'_S \text{ und } t(\ulcorner t \urcorner) \equiv \ulcorner v \urcorner \in W \text{ für einen } S\text{-Term } v,$$

durch einen $S_{[v_0]}$-Term $r_1(x)$ repräsentiert wird, so gilt für jeden $S_{[v_0]}$-Term $r_2(x)$:

$$\text{Für alle } S\text{-Terme } q: \quad r_2(\ulcorner r_2(x) \urcorner) \equiv \ulcorner q \urcorner \in W \text{ gdw. } r_1(\ulcorner r_2(x) \urcorner) \equiv \ulcorner q \urcorner \in W.$$

Aus dem abstrakten Satz 2.17 ergeben sich ein Unvollständigkeitssatz, ein Unentscheidbarkeitssatz und Antinomien parallel zur Lügner-Antinomie und zu den Antinomien von Grelling und Richard. Der entsprechende Unvollständigkeitssatz lautet:

Satz 2.18. *PA sei aufzählbar, enthalte* PA *und sei widerspruchsfrei. Dann gibt es einen $S_{\mathrm{Ar}[v_0]}$-Term $s_2(x)$, so dass für alle $q \in \mathbb{N}$: $PA \nvdash s_2(\ulcorner u \urcorner) \equiv \boldsymbol{q}$, wobei u ein Fixpunkt zu $s_2(x)$ sei.*[39]

Beweis. Wie in Ritchie u. Young (1969) gezeigt wird, ist unter diesen Bedingungen die partiell rekursive Funktion $g : \mathbb{N} \to \mathbb{N}$, die gegeben ist durch

$$g(n) = \begin{cases} m, & \text{falls } n = \ulcorner t \urcorner \text{ für einen geschlossenen } S_{\mathrm{Ar}}\text{-Term } t \\ & \text{und } PA \vdash t \equiv \boldsymbol{m}, \\ \text{undefiniert,} & \text{sonst,} \end{cases}$$

in PA durch einen $S_{\mathrm{Ar}[v_0]}$-Term $s_1(x)$ repräsentierbar. Die Anwendung des abstrakten Satzes 2.17 a) ergibt also für $s_1(x) + 1$ als $S_{\mathrm{Ar}[v_0]}$-Term $s_2(x)$:

$$\text{für alle } q \in \mathbb{N}: \quad PA \vdash s_1(\ulcorner u \urcorner) + 1 \equiv \boldsymbol{q} \text{ gdw. } PA \vdash s_1(\ulcorner u \urcorner) \equiv \boldsymbol{q}.$$

Hierbei sei u Fixpunkt zu $s_1(x) + 1$.
Damit gilt für alle $q \in \mathbb{N}$: $PA \nvdash s_1(\ulcorner u \urcorner) + 1 \equiv \boldsymbol{q}$. □

[39] In Bezug auf die Formulierung des Satzes sei an die Konvention der Termschreibweise auf S. 30 erinnert. Eine Formulierung des Satzes ohne Termschreibweise lautet:
Es gibt eine $S_{\mathrm{Ar}[v_0,v_1]}$-Formel $\varphi(x, y)$ und $n \in \mathbb{N}$, so dass $PA \vdash \bigwedge x \bigvee y \varphi(x, y)$, jedoch für alle $q \in \mathbb{N}$: $PA \nvdash \varphi(\boldsymbol{n}, \boldsymbol{q})$.

Eine zur Lügner-Antinomie parallele Antinomie könnte lauten:

> Es wird ausgegangen von dem Term u, $u :=$ "diejenige Zahl, die dieser Term bezeichnet, plus 1", oder allgemeiner $u :=$ "ein von demjenigen Objekt, das dieser Term bezeichnet, verschiedenes Objekt".
> Es ergibt sich: Dasjenige Objekt, das der Term u bezeichnet, ist das, was er bezeichnet, also ein von demjenigen Objekt, das er bezeichnet, verschiedenes Objekt.

Kapitel 3

Diagonalargumente in ontologischen und semantischen Antinomien

In diesem Kapitel werden die strukturellen Gemeinsamkeiten der im letzten Kapitel dargestellten ontologischen und semantischen Antinomien untersucht. Dazu wird ein mengentheoretisches Schema entwickelt – das Schema einer Diagonalstruktur –, das wesentlich von Cantors Diagonalverfahren[1] geprägt ist und von dem gezeigt werden soll, dass auf seiner Grundlage die relevanten Züge der einzelnen Antinomien dargestellt werden können. Ferner soll gezeigt werden, dass auf diese Weise die gemeinsame Struktur, aber auch die Unterschiede der Antinomien deutlich werden.

Allgemein stellt sich hier zunächst die Frage, was die Kriterien für ein Schema sein könnten, das die relevanten Züge oder die Struktur von Einzelfällen befriedigend darstellt. Sicher handelt es sich bei einem solchen Schema um eine Abstraktion, die um so weitgehender abstrahiert und um so weniger spezifische Begriffe verwendet, je unterschiedlicher die Einzelfälle sind, die darunter subsumiert werden sollen. Trotzdem soll das Schema wesentliche Aspekte jedes Einzelfalls wiedergeben. Umgekehrt werden die in das Schema aufgenommenen Eigenschaften so auch zu wesentlichen erklärt. Priest, dessen Darstellung der Struktur der Antinomien am Ende dieses Abschnitts der hier entwickelten gegenübergestellt wird, beschreibt die Schwierigkeit, klare Bedingungen für ein solches Schema zu bestimmen, wie folgt: “We want not just any old pattern, but the essential pattern. How to cash out this notion here is an interesting question. If we were dealing with a scientific, empirical, matter, what we would be after would be a pattern that is a lawlike, not just accidental, generalisation (however, in the end, one wishes to cash out the distinction). But it is not clear (at least to me) how to transfer the distinction to the present case.”[2] Diese Frage soll

[1] Vgl. Fußnote 3, S. 89.
[2] Priest (2002), S. 135.

hier nicht weiter behandelt werden. Der Begriff eines strukturbeschreibenden Schemas bleibt also vage.

Das vorgeschlagene Schema hat die Form einer mengentheoretischen Struktur. Analog zu den abstrakten Sätzen aus Abschnitt 2.1, in denen nach der Verwendung eines Wahrheits- und eines Falschheitsbegriffs unterschieden wird, werden auch hier zwei Arten von Diagonalstrukturen eingeführt. Der zu der Verwendung eines Falschheitsbegriffs analoge Begriff einer Diagonalstruktur, die durch ein 3-Tupel aus einer Menge, einer Abbildung und einem Element, dem Diagonalelement, gegeben ist, beschreibt beispielsweise, dass die Abbildung von der Menge in ihre Potenzmenge geht und das Diagonalelement genau die Elemente der Menge enthält, die nicht in ihrem Bild enthalten sind. Unter Verwendung der klassischen Logik ergibt sich dann, dass das Diagonalelement nicht im Bild der Abbildung liegen kann.

Im Anschluss an die Darstellung werden die einzelnen Antinomien und Widerspruchsbeweise auf Diagonalstrukturen bezogen. Dies geschieht durch Interpretation der Objektsprachen über zu den jeweiligen Symbolmengen passenden Strukturen. Die ontologischen Antinomien, die auf dem Klassen- und dem Elementbegriff aufbauen, werden über einem Universum – also einer Menge –, auf dem eine zweistellige Relation gegeben ist, interpretiert. Die zweistellige Relation erzeugt eine Abbildung vom Universum in die Potenzmenge des Universums, indem einem Element des Universums die Menge der mit ihm in Relation stehenden Elemente zugeordnet wird. Auf dieser Basis wird durch die in der Antinomie gemachten Voraussetzungen eine Diagonalstruktur erzeugt. In den semantischen Antinomien werden bestimmte einstellige Formeln vorausgesetzt. Eine einstellige Formel wird über dem Universum, das die Menge der einstelligen Formeln der Sprache enthält, als Menge interpretiert. Dadurch wird wiederum eine Abbildung vom Universum in die Potenzmenge des Universums erzeugt, indem einem Element – also einer einstelligen Formel – seine Extensionsmenge zugeordnet wird. Auf diesem Weg entsteht durch die in der Antinomie gemachten Voraussetzungen eine Diagonalstruktur, die direkt auch mit den durch ontologische Antinomien erzeugten Diagonalstrukturen verglichen werden kann. In Abschnitt 3.6 werden einige Formen von Antinomien, die keine Diagonalargumente im eigentlichen Sinn verwenden, wie z. B. der Lügner-Zirkel, auf ihre Strukturen hin untersucht.

Vor diesem Hintergrund werden die Voraussetzungen ontologischer und semantischer Antinomien gegenübergestellt und es wird gefragt, ob es zu den jeweiligen Antinomien Entsprechungen im anderen Antinomietyp gibt und wo Gemeinsamkeiten und Unterschiede in der Art der Voraussetzungen bestehen.

Im Anschluss werden einige alternative Schemata, die in Zusammenhang mit Antinomien stehen, vorgestellt und mit dem Diagonalschema der vorliegenden Arbeit verglichen.

Es folgt eine Übersicht und Klassifizierung von Lösungsansätzen zur Antinomienproblematik, die auf die vorangegangene Untersuchung der Voraussetzungen der einzelnen Antinomien aufbaut. Ferner wird der Frage nach der Einheitlichkeit von Lösungen nachgegangen. Das Kapitel endet mit der Feststellung, dass Antinomien Grenzen des Aussagbaren deutlich machen – unabhängig vom eingeschlagenen Lösungsweg.

3.1 Diagonalstrukturen

Der Begriff der Diagonalstruktur wird vor dem Hintergrund einer Mengenlehre unter Verwendung der klassischen Logik eingeführt.

Von einer Diagonalstruktur$_1$ bzw. einer Diagonalstruktur$_2$ soll gesprochen werden, wenn Folgendes gegeben ist:
N, N' sind Mengen, M ist die Menge von Abbildungen von N nach N'. $f : N \to M$ ist eine Abbildung. Es gibt ein Element $m_1 \in M$, so dass für alle $n \in N$

$$f(n)(n) = m_1(n),$$

bzw. ein Element $m_2 \in M$, so dass für alle $n \in N$

$$f(n)(n) \neq m_2(n).$$

m_1 bzw. m_2 heiße Diagonalelement$_1$ bzw. Diagonalelement$_2$.[3]

[3] Das Beweisverfahren, das Cantor benutzt, um zu zeigen, dass die Mächtigkeit der reellen Zahlen echt größer ist als die der natürlichen Zahlen, wird üblicherweise als Cantorsches Diagonalverfahren bezeichnet. Der Begriff der Diagonalstruktur$_2$ ist daran angelehnt. Das Beweisverfahren von Cantor hat eine Diagonalstruktur$_2$ im Sinne der vorliegenden Arbeit als Grundlage und führt die zur Satzaussage gegenteilige Annahme zum Widerspruch, indem gezeigt wird, dass das Diagonalelement$_2$ im Bild der Abbildung f liegt. Der Begriff der

Für abzählbare N kann eine Diagonalstruktur wie folgt veranschaulicht werden:

Dabei seien $x_1, x_2, x_3, ...$ die Elemente von N. $y_1, y_2, y_3, ...$ seien die Bilder von Elementen aus N unter f (y_1 sei das Bild von x_1, usw.). Die Diagonalelemente m_1 und m_2 stehen unterhalb – es ist offen gelassen, ob sie in $f[N]$ liegen oder nicht. Im rechten Teil der Darstellung sind die Elemente von $f[N]$, angewendet auf die einzelnen Argumente, notiert. Für die Diagonalelemente ausschlaggebend ist die Diagonale. Das Diagonalelement$_1$ stimmt an jeder Stelle mit der Diagonalen überein. Das Diagonalelement$_2$ unterscheidet sich an jeder Stelle von der Diagonalen, was durch das Ungleichheitszeichen angedeutet ist.

N	$f[N]$				
x_1	y_1	$\boldsymbol{y_1(x_1)}$	$y_1(x_2)$	$y_1(x_3)$	$\cdots$
x_2	y_2	$y_2(x_1)$	$\boldsymbol{y_2(x_2)}$	$y_2(x_3)$	$\cdots$
x_3	y_3	$y_3(x_1)$	$y_3(x_2)$	$\boldsymbol{y_3(x_3)}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$
	m_1	$y_1(x_1)$	$y_2(x_2)$	$y_3(x_3)$	$\cdots$
	m_2	$\neq y_1(x_1)$	$\neq y_2(x_2)$	$\neq y_3(x_3)$	$\cdots$

Die genaue Definition einer Diagonalstruktur laute:

Definition 3.1. Eine *Diagonalstruktur*$_1$ bzw. eine *Diagonalstruktur*$_2$ ist ein Tupel (N, N', f, m_1) bzw. (N, N', f, m_2), wobei N, N' Mengen sind, $f : N \to M$ eine Abbildung in die Menge aller Abbildungen M von N nach N' ist und $m_1 \in M$ bzw. $m_2 \in M$, so dass für alle $n \in N$

$$f(n)(n) = m_1(n) \quad \text{bzw.} \quad f(n)(n) \neq m_2(n).$$

Satz 3.2. *a)* (N, N', f, m_1) *sei eine Diagonalstruktur*$_1$. *Liegt* m_1 *im Bild von* f, *so gilt für jedes Urbild* n_1 *von* m_1 *und alle* $n_2 \in N$

$$f(n_2)(n_2) = f(n_1)(n_2).$$

Diagonalstruktur$_1$ ist gewissermaßen das positive Gegenstück zu dem der Diagonalstruktur$_2$, das sich auf die nicht "verschobene" Diagonale bezieht. Zur Präzisierung und Verallgemeinerung des Begriffs des Cantorschen Diagonalverfahrens vgl. Essler (1964) und die Bemerkungen auf Seite 120 der vorliegenden Arbeit.

b) (N, N', f, m_2) *sei eine Diagonalstruktur*$_2$*. Dann liegt das Diagonalelement* m_2 *nicht im Bild von* f*.*

Beweis. Zu b) Für jedes Urbild n_2 von m_2 gilt

$$f(n_2)(n_2) \neq m_2(n_2) = f(n_2)(n_2).$$

Damit liegt m_2 nicht im Bild von f. □

In der folgenden Definition wird der Spezialfall einer Diagonalstruktur für $N' = \{0, 1\}$ als Diagonalstruktur für Teilmengen von N eingeführt:

Definition 3.3. Eine *Diagonalstruktur*$_1$ *(für Teilmengen)* bzw. *Diagonalstruktur*$_2$ *(für Teilmengen)* ist ein Tupel (N, f, m_1) bzw. (N, f, m_2), wobei N eine Menge, $f : N \to P(N)$ eine Abbildung in die Potenzmenge von N und $m_1 \in P(N)$ bzw. $m_2 \in P(N)$, so dass für alle $n \in N$ gilt:

$$\begin{aligned} n \in f(n) \quad &\text{gdw.} \quad n \in m_1 \text{ bzw.} \\ n \notin f(n) \quad &\text{gdw.} \quad n \in m_2. \end{aligned}$$

Für abzählbare N lässt sich diese Art von Diagonalstruktur wie folgt veranschaulichen. Im rechten Teil der Darstellung stehen in jedem Eintrag jeweils eine Null oder eine Eins, je nachdem, ob das Element aus N in seinem Bild enthalten ist oder nicht. Das Diagonalelement$_1$ stimmt an jeder Stelle mit der Diagonalen überein. Das Diagonalelement$_2$ unterscheidet sich an jeder Stelle von der Diagonalen, was jeweils durch das Nichtelementzeichen angedeutet wird.

N	$f[N]$				
x_1	y_1	$\boldsymbol{x_1 \in y_1}$?	$x_2 \in y_1$?	$x_3 \in y_1$?	$\cdots$
x_2	y_2	$x_1 \in y_2$?	$\boldsymbol{x_2 \in y_2}$?	$x_3 \in y_2$?	$\cdots$
x_3	y_3	$x_1 \in y_3$?	$x_2 \in y_3$?	$\boldsymbol{x_3 \in y_3}$?	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$
	m_1	$x_1 \in y_1$?	$x_2 \in y_2$?	$x_3 \in y_3$?	$\cdots$
	m_2	$x_1 \notin y_1$?	$x_2 \notin y_2$?	$x_3 \notin y_3$?	$\cdots$

Satz 3.4. *a)* (N, f, m_1) *sei eine Diagonalstruktur*$_1$. *Liegt* m_1 *im Bild von* f, *so gilt für jedes Urbild* n_1 *von* m_1 *und alle* $n_2 \in N$:

$$n_2 \in f(n_2) \quad \textit{gdw.} \quad n_2 \in f(n_1).$$

b) (N, f, m_2) *sei eine Diagonalstruktur*$_2$. *Dann liegt das Diagonalelement* m_2 *nicht im Bild von* f.

Beweis. Zu b) Für jedes Urbild n_2 von m_2 gilt:

$$n_2 \notin f(n_2) \quad \text{gdw.} \quad n_2 \in f(n_2).$$

□

Eine allgemeinere Version ergibt sich, wenn für eine beliebige Abbildung $g : N \to P(N)$ an ein Diagonalelement die Anforderung gestellt wird, dass $n \in g(n)$ gdw. $n \in m_0$:

Definition 3.5. $g : N \to P(N)$ sei eine Abbildung. Eine *Diagonalstruktur relativ zu* g ist ein Tupel (N, f, m_0), wobei N eine Menge, $f : N \to P(N)$ eine Abbildung in die Potenzmenge von N und $m_0 \in P(N)$, so dass für alle $n \in N$ gilt:

$$n \in g(n) \quad \text{gdw.} \quad n \in m_0.$$

Satz 3.6. (N, f, m_0) *sei eine Diagonalstruktur relativ zu* g. *Liegt* m_0 *im Bild von* f, *so gilt für jedes Urbild* n_0 *von* m_0 *und alle* $n \in N$:

$$n \in g(n) \quad \textit{gdw.} \quad n \in f(n_0).$$

Auf den Seiten 69 ff. wird der abstrakte Satz 2.10 – dies ist die Fixpunktversion der abstrakten Sätze – als Spezialfall der abstrakten Sätze 2.3 und 2.7 – dies sind die Versionen ohne Fixpunktsatz – beschrieben. Dieser Spezialfall kann auch auf der Ebene der Diagonalstrukturen formuliert werden. Dies ermöglicht in Abschnitt 3.3 z. B. eine direkte Interpretation der Lügner-Antinomie als Anwendung von Satz 3.4 auf eine Diagonalstruktur, die ja unter den abstrakten Satz 2.10 fällt. Damit wird auch auf der Ebene von Diagonalstrukturen der Zusammenhang zwischen der Antinomie von Grelling und der Lügner-Antinomie nachvollzogen.

Zusätzlich zu den Mengen N und $P(N)$ werden die Menge N' der geordneten Paare der Form (n,n) mit $n \in N$ und die Potenzmenge $P(N')$ betrachtet. Eine Abbildung $f' : N' \to P(N')$ wird so eingeführt, dass das Bild eines Paares (n,n) unter f' genau die Elemente (x,x) mit $x \in f(n)$ enthält. Damit ist in $f[N]$ gewissermaßen die Bildung eines geordneten Paares (n,n) und die Verknüpfung dieser Abbildung mit Elementen aus $f[N']$ – und umgekehrt – möglich.

Es wird eine Bedingung aufgestellt, die die Existenz eines Diagonalelements – relativ zu einer Abbildung g – in $P(N)$ und $P(N')$ garantiert.

Satz 3.7. *N sei eine Menge, N' sei die Menge der geordneten Paare aus N mit identischen Komponenten, $N' := \{(n,n)|n \in N\}$. $f : N \to P(N)$, $g : N \to P(N)$ seien Abbildungen in die Potenzmenge von N, $f' : N' \to P(N')$, $g' : N' \to P(N')$ seien Abbildungen in die Potenzmenge von N', so dass $f'((n_1,n_1)) = \{(n_2,n_2) \in N'|n_2 \in f(n_1)\}$ und $g'((n_1,n_1)) = \{(n_2,n_2) \in N'|n_2 \in g(n_1)\}$.*

Für $m'_0 \in M'$ und alle $n \in N$ gelte: $n \in g(n)$ gdw. $(n,n) \in m'_0$.

a) $(N, f, \{n|(n,n) \in m'_0\})$ ist eine Diagonalstruktur relativ zu g.
b) (N', f', m'_0) ist eine Diagonalstruktur relativ zu g'.
c) m'_0 liegt im Bild der Abbildung f' gdw. $\{n|(n,n) \in m'_0\}$ im Bild der Abbildung f liegt.

Beweis. Zu a) Nach Voraussetzung gilt für alle $n \in N$: $n \in g(n)$ gdw. $(n,n) \in m'_0$. Ferner gilt: $(n,n) \in m'_0$ gdw. $n \in \{n|(n,n) \in m'_0\}$, woraus die Behauptung – für alle $n \in N$: $n \in g(n)$ gdw. $n \in \{n|(n,n) \in m'_0\}$ – folgt.
Zu b) Nach Voraussetzung gilt für alle $n \in N$: $n \in g(n)$ gdw. $(n,n) \in m'_0$. Ferner gilt für alle $n \in N$: $(n,n) \in g'((n,n))$ gdw. $n \in g(n)$, woraus die Behauptung – für alle $n' \in N'$: $n' \in g'(n')$ gdw. $n' \in m'_0$ – folgt.
Zu c) Nach Voraussetzung an f und f' gilt für alle $n \in N$: $f'((n,n)) = m'_0$ gdw. $f(n) = \{n|(n,n) \in m'_0\}$. □

Die Konstruktion lässt sich für abzählbare N und $g = f$ wie folgt darstellen:

N	$f[N]$			
x_1	y_1	$x_1 \in y_1$?	$x_2 \in y_1$?	$x_3 \in y_1$?
x_2	y_2	$x_1 \in y_2$?	$x_2 \in y_2$?	$x_3 \in y_2$?
x_3	y_3	$x_1 \in y_3$?	$x_2 \in y_3$?	$x_3 \in y_3$?
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$\{n \mid (n,n) \in m'_0\}$	$x_1 \in y_1$?	$x_2 \in y_2$?	$x_3 \in y_3$?
N'	$f[N']$			
(x_1, x_1)	$\{(x,x) \mid x \in y_1\}$	$(x_1, x_1) \in \{(x,x) \mid x \in y_1\}$?	$(x_2, x_2) \in \{(x,x) \mid x \in y_1\}$?	$(x_3, x_3) \in \{(x,x) \mid x \in y_1\}$?
(x_2, x_2)	$\{(x,x) \mid x \in y_2\}$	$(x_1, x_1) \in \{(x,x) \mid x \in y_2\}$?	$(x_2, x_2) \in \{(x,x) \mid x \in y_2\}$?	$(x_3, x_3) \in \{(x,x) \mid x \in y_2\}$?
(x_3, x_3)	$\{(x,x) \mid x \in y_3\}$	$(x_1, x_1) \in \{(x,x) \mid x \in y_3\}$?	$(x_2, x_2) \in \{(x,x) \mid x \in y_3\}$?	$(x_3, x_3) \in \{(x,x) \mid x \in y_3\}$?
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	m'_0	$x_1 \in y_1$?	$x_2 \in y_2$?	$x_3 \in y_3$?

3.2 Diagonalargumente in ontologischen Antinomien

Der Satz von Cantor

Der Satz von Cantor ist eine direkte Anwendung des Diagonalarguments aus Satz 3.4 b). Genauer gesagt wird in der Klassensprechweise eine Diagonalstruktur$_2$ beschrieben und Satz 3.4 angewendet.

Die Sprache der Antinomie von Cantor wird nun als Sprache erster Stufe aufgefasst und über einer Struktur interpretiert. Die Symbolmenge S_{Kl} bestehe aus der zweistelligen Relationskonstante $\in$. Die Sprache sei interpretiert über der S_{Kl}-Struktur $\mathcal{A}_{\text{Kl}} = (U, j_\in)$, wobei $j_\in$ die zweistellige Relationskonstante $\in$ auf die Relation $R_\in$ über U abbilde.

Die Abbildung $id_{R_\in} : U \to P(U)$ bilde ein Element $x \in U$ ab auf die Menge $\{y \in U \mid (y, x) \in R_\in\}$ aus $P(U)$. $id_{R_\in} : P(U) \to P(P(U))$ bilde eine Menge $x \in P(U)$ ab auf die Menge $\{y \in id_{R_\in}[U] \mid {id_{R_\in}}^{-1}(y) \in x\}$ aus $P(P(U))$.[4] $id_{R_\in}$ ist injektiv. Die Abbildung $id_{R_\in} : U \to P(U)$ bildet also das Bild einer Klasse als

[4]Die Bezeichnung "$id_{R_\in}$" wird für beide Abbildungen verwendet, da es nicht zu Missverständnissen kommen kann.

Objekt auf das – nach $R_\in$ – entsprechende Bild der Klasse als Zusammenfassung von Klassen ab. Die Abbildung $id_{R_\in} : P(U) \to P(P(U))$ bildet das Bild einer Klasse von Klassen als Objekten auf das – nach $R_\in$ – entsprechende Bild der Klasse von Klassen als Zusammenfassungen ab:

$$\begin{array}{lllll} U & \to & P(U) & \to & P(P(U)) \\ x & \mapsto & \{y \in U | (y,x) \in R_\in\} & \mapsto & \{y \in id_{R_\in}[U] | ({id_{R_\in}}^{-1}(y), x) \in R_\in\} \end{array}$$

Die objektsprachliche Annahme, dass es zu jeder Klasse eine Potenzklasse gibt, besagt in der Struktur $\mathcal{A}_{\text{Kl}}$, dass zu jeder Menge a aus $id_{R_\in}[U]$ die Menge $\{x \in id_{R_\in}[U] | x \subseteq a\}$ in $id_{R_\in}[id_{R_\in}[U]]$ liegt.

Die Annahme im Widerspruchsbeweis des Satzes von Cantor besagt, interpretiert in der Struktur $\mathcal{A}_{\text{Kl}}$, dass es zu einer Menge $a \in id_{R_\in}[U]$ eine surjektive Abbildung $f : a \to \{x \in id_{R_\in}[U] | x \subseteq a\}$ gibt. Die Annahme wird im Beweis von Cantor zum Widerspruch geführt. Hierzu wird ein Element c definiert, dem in $\mathcal{A}_{\text{Kl}}$ das Element $\{y \in a | y \notin f(y)\} \in \{x \in id_{R_\in}[U] | x \subseteq a\}$ entspricht. Der Widerspruch, der sich daraus in der Objektsprache ergibt, lautet übertragen auf $\mathcal{A}_{\text{Kl}}$: Für jedes Urbild $x \in a$ von $\{y \in a | y \notin f(y)\}$ unter f gilt:

$$x \in \{y \in a | y \notin f(y)\} \text{ und } x \notin \{y \in a | y \notin f(y)\}.$$

Mit $N = a$, $m_2 = \{y \in a | y \notin f(y)\}$ und $f : a \to \{x \in id_{R_\in}[U] | x \subseteq a\}$ ist eine Diagonalstruktur$_2$ gegeben. Der Widerspruchsbeweis im Satz von Cantor entspricht dem Beweis zu Satz 3.4 b), der die Annahme zum Widerspruch führt, dass f surjektiv ist.

Die Antinomie von Cantor

In der Antinomie von Cantor wird der Satz von Cantor auf die Allklasse angewendet. Zusätzlich benutzte Voraussetzung ist also die Existenz der Allklasse, von der gezeigt werden kann, dass sie mit ihrer Potenzklasse identisch ist. Die Existenz der Allklasse bedeutet in der Struktur $\mathcal{A}_{\text{Kl}}$, dass die Menge U im Bild der Abbildung $id_{R_\in} : U \to P(U)$ liegt. Die Abbildung $id_{R_\in} : U \to P(U)$ ist eine surjektive Abbildung von U nach $\{x \in id_{R_\in}[U] | x \subseteq U\} = id_{R_\in}[U]$. Die Anwendung des Diagonalarguments aus dem Satz von Cantor führt zum Widerspruch zur Annahme, dass U im Bild von $id_{R_\in}$ liegt, bzw. zur Existenz der Allklasse.

Der Widerspruch ergibt sich nach Satz 3.4 b) als:

$$id^{-1}_{R_\in}(\{y \in U | y \notin id_{R_\in}(y)\}) \in \{y \in U | y \notin id_{R_\in}(y)\} \quad \text{und}$$
$$id^{-1}_{R_\in}(\{y \in U | y \notin id_{R_\in}(y)\}) \notin \{y \in U | y \notin id_{R_\in}(y)\}.$$

Die Antinomie von Burali-Forti

Hier werden in der Objektsprache Klassen betrachtet, die transitiv sind und auf denen die Elementrelation eine Wohlordnung ist: Ordinalzahlen. Da jede solche Klasse nur Klassen mit dieser Eigenschaft enthält, ist jede Ordinalzahl eine Klasse von Ordinalzahlen. $ord \in P(U)$ sei das Bild des einstelligen Begriffs der Objektsprache “x ist eine Ordinalzahl” in der Struktur $\mathcal{A}_{\mathrm{Kl}}$, also die Menge aller Elemente aus U, die als Variableninterpretation den Begriff “x ist eine Ordinalzahl” wahr machen.

Die Annahme in der Objektsprache, dass es eine Klasse aller Ordinalzahlen gibt, bedeutet in der Struktur $\mathcal{A}_{\mathrm{Kl}}$, dass die Menge ord im Bild der Abbildung $id_{R_\in} : U \to P(U)$ liegt und ihr Urbild in ord liegt, da die Zusammenfassung aller Ordinalzahlen transitiv und wohlgeordnet ist. Der Potenzmenge entspricht innerhalb der Ordinalzahlen der Nachfolger einer Ordinalzahl.[5] Wird eine Ordinalzahl in $\mathcal{A}_{\mathrm{Kl}}$ als die Menge $a \in id_{R_\in}[ord]$ interpretiert, so wird ihr Nachfolger als $\{x \in id_{R_\in}[ord] | x \subseteq a\} \in id_{R_\in}[id_{R_\in}[ord]]$ interpretiert. Die Abbildung $id_{R_\in} : U \to P(U)$ bildet eine surjektive Abbildung von ord nach $\{x \in id_{R_\in}[ord] | x \subseteq ord\} = id_{R_\in}[ord]$.

In der Antinomie von Burali-Forti wird das Diagonalelement – interpretiert in $\mathcal{A}_{\mathrm{Kl}}$ – gebildet als $\{y \in ord\} = \{y \in ord | y \notin id_{R_\in}(y)\} \in id_{R_\in}[ord]$. Der Widerspruch ergibt sich nach Satz 3.4 b) als:[6]

$$id^{-1}_{R_\in}(\{y \in ord | y \notin id_{R_\in}(y)\}) \in \{y \in ord | y \notin id_{R_\in}(y)\} \quad \text{und}$$
$$id^{-1}_{R_\in}(\{y \in ord | y \notin id_{R_\in}(y)\}) \notin \{y \in ord | y \notin id_{R_\in}(y)\}.$$

[5] Vgl. S. 14.

[6] Geht man von Ordinalzahlen als Äquivalenzklassen wohlgeordneter Klassen aus, so wird statt der zweistelligen Relationskonstante “$\in$” die zweistellige Relationskonstante “$<$” in der Struktur $\mathcal{A}_{\mathrm{Kl}}$ durch $j_\in$ auf $R_\in$ abgebildet. Interpretiert man die Antinomie von Burali-Forti auf diese Weise in $\mathcal{A}_{\mathrm{Kl}}$, ergibt sich die gleiche Diagonalstruktur und der gleiche Widerspruch wie im oben dargestellten Fall, der von Ordinalzahlen als transitiven Klassen ausgeht. Vgl. Fußnote 9, S. 15.

Antinomie von Russell

In der Formulierung der Antinomie von Russell wird in der Objektsprache direkt vorausgesetzt, dass es eine Klasse aller Klassen, die sich nicht selbst enthalten, gibt. Diese Annahme bedeutet in der S_{Kl}-Struktur $\mathcal{A}_{\mathrm{Kl}}$, dass die Menge $\{x \in U | x \notin id_{R_\in}(x)\}$ im Bild der Abbildung $id_{R_\in}$ liegt. Mit $N = U$, $m_2 = \{x \in U | x \notin id_{R_\in}(x)\} \in id_{R_\in}[U]$ ist eine Diagonalkonstruktion$_2$ gegeben. Satz 3.4 b) führt zum Widerspruch:

$$id_{R_\in}^{-1}(\{x \in U | x \notin id_{R_\in}(x)\}) \in \{x \in U | x \notin id_{R_\in}(x)\} \quad \text{und}$$
$$id_{R_\in}^{-1}(\{x \in U | x \notin id_{R_\in}(x)\}) \notin \{x \in U | x \notin id_{R_\in}(x)\}.$$

3.3 Diagonalargumente in semantischen Antinomien

In Abschnitt 2.1 werden die Lügner-Antinomie, die Antinomien von Grelling und von Richard, der Satz von Tarski, der erste Gödelsche Unvollständigkeitssatz und der Unentscheidbarkeitssatz als Spezialfälle von abstrakten Sätzen, die sich auf eine abstrakte Sprache beziehen, dargestellt. Es genügt deshalb hier, diese abstrakten Sätze als Anwendungen von Diagonalargumenten auf Diagonalstrukturen darzustellen.

Darstellung von Satz 2.7 – Form ohne Fixpunktsatz

Satz 2.7 a) und b) sind die abstrakten, rein objektsprachlichen Fassungen der Antinomie von Grelling, aufbauend auf die Begriffe "autologisch" bzw. "heterologisch".

$(S, W, \leftrightarrow)$ sei eine abstrakte Sprache mit Äquivalenz. S wird in einer S-Struktur $\mathcal{A}_{\mathrm{abstr}} = (U, j_{\mathrm{abstr}})$ über dem Universum $U \supseteq F_S$ interpretiert. Dazu wird durch die Abbildung j_{abstr}

- jeder $S_{[v_0]}$-Formel aus $F_S \setminus F'_S$ ein Element aus $P(U)$ und
- jedem geschlossenen S-Term aus T'_S ein Element aus U zugeordnet.

Für alle Standardnamen einer S-Formel φ, also für alle S-Terme aus $n_S[F_S]$, sei $j_{\mathrm{abstr}}(\ulcorner\varphi\urcorner) = \varphi$.

Für alle $S_{[v_0]}$-Formeln $\varphi(x)$ und alle geschlossenen S-Terme t sei die Modellbeziehung $\mathcal{A}_{\mathrm{abstr}} \models$ definiert durch: $\mathcal{A}_{\mathrm{abstr}} \models \varphi(t)$ gdw. $j_{\mathrm{abstr}}(t) \in j_{\mathrm{abstr}}(\varphi(x))$.

Die Äquivalenz “ $\leftrightarrow$ ” der abstrakten Sprache sei klassisch interpretiert; für alle S-Sätze φ, ψ gelte: $\mathcal{A}_{\text{abstr}} \models \varphi \leftrightarrow \psi$ gdw. $[\mathcal{A}_{\text{abstr}} \models \varphi$ gdw. $\mathcal{A}_{\text{abstr}} \models \psi]$.

Parallel zur Abbildung $id_{R_\in} : U \to P(U)$ bilde die Abbildung $id_{Ext} : F_S \setminus F'_S \to P(F_S \setminus F'_S)$ eine $S_{[v_0]}$-Formel $\varphi(x)$ aus U auf ihre Extension über den $S_{[v_0]}$-Formeln, also die Menge $\{\psi(x) \in F_S \setminus F'_S | \mathcal{A}_{\text{abstr}} \models \varphi(\ulcorner\psi(x)\urcorner)\}$ ab:

$$\begin{array}{lcl} F_S \setminus F'_S & \to & P(F_S \setminus F'_S) \\ \varphi(x) & \mapsto & j_{\text{abstr}}(\varphi(x)) \cap F_S \setminus F'_S \end{array}$$

In Satz 2.7 a) wird eine $S_{[v_0]}$-Formel $\vartheta_1(x)$ vorausgesetzt, so dass für jede $S_{[v_0]}$-Formel $\varphi(x)$ gilt $\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow \vartheta_1(\ulcorner\varphi(x)\urcorner) \in W$. Übertragen auf die Struktur $\mathcal{A}_{\text{abstr}}$ bedeutet dies, dass die Menge der $S_{[v_0]}$-Formeln, die in ihrer Extension enthalten sind, die Extension der $S_{[v_0]}$-Formel $\vartheta_1(x)$ bildet: $\{\varphi(x) \in F_S \setminus F'_S | \varphi(x) \in id_{Ext}(\varphi(x))\} = id_{Ext}(\vartheta_1(x))$. Mit $N = F_S \setminus F'_S$, $f = id_{Ext}$ und $m_1 = id_{Ext}(\vartheta_1(x))$ ist also eine Diagonalstruktur$_1$ gegeben, wobei das Diagonalelement im Bild von f liegt. Die Anwendung von Satz 3.4 a) ergibt für alle $S_{[v_0]}$-Formeln $\vartheta_2(x)$ nach Bestimmung von n_2 als $n_2 = \vartheta_2$:

$$\vartheta_2 \in id_{Ext}(\vartheta_2) \text{ gdw. } \vartheta_2 \in id_{Ext}(\vartheta_1).$$

Zur Darstellung von Satz 2.7 b) sei die Negation $\neg$ der abstrakten Sprache zunächst klassisch interpretiert; für alle S-Sätze φ gelte: $\mathcal{A}_{\text{abstr}} \models \neg\varphi$ gdw. $\mathcal{A}_{\text{abstr}} \not\models \varphi$.

In Satz 2.7 b) wird eine $S_{[v_0]}$-Formel $\vartheta_2(x)$ vorausgesetzt, so dass für jede $S_{[v_0]}$-Formel $\varphi(x)$ gilt $\neg\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow \vartheta_2(\ulcorner\varphi(x)\urcorner) \in W$. Übertragen auf die Struktur $\mathcal{A}_{\text{abstr}}$ bedeutet dies, dass die Menge der $S_{[v_0]}$-Formeln, die nicht in ihrer Extension enthalten sind, die Extension der $S_{[v_0]}$-Formel $\vartheta_2(x)$ bildet: $\{\varphi(x) \in F_S \setminus F'_S | \varphi(x) \notin id_{Ext}(\varphi(x))\} = id_{Ext}(\vartheta_2(x))$. Wie oben, jetzt jedoch mit $m_2 = id_{Ext}(\vartheta_2(x))$, ergibt sich nun eine Diagonalstruktur$_2$, wobei das Diagonalelement im Bild von f liegt. Die Anwendung von Satz 3.4 b) ergibt den Widerspruch:

$$\vartheta_2(x) \in id_{Ext}(\vartheta_2(x)) \text{ und } \vartheta_2(x) \notin id_{Ext}(\vartheta_2(x)).$$

Geht man nicht von einer klassischen Interpretation der Negation der abstrakten

Sprache aus, so erhält man keinen Widerspruch, sondern die Äquivalenz:

$$\vartheta_2(x) \in id_{Ext}(\neg\vartheta_2(x)) \text{ gdw. } \vartheta_2(x) \in id_{Ext}(\vartheta_2(x)).$$

Das zum Teil metasprachliche Pendant zu Satz 2.7 a) und b) – Satz 2.3 a) und b) – lässt sich auf die gleiche Weise als Diagonalargument auffassen wie Satz 2.7 a) und b). Der Unterschied in der Bedeutung des Ergebnisses für die Satzteile a) besteht in folgender Weise: Der auf eine Diagonalstruktur (N, f, m_1) angewendete Satz 3.4 a) besagt, dass unter den Voraussetzungen des Satzes für alle $n_2 \in N$: $n_2 \in f(n_2)$ gdw. $n_2 \in f(n_1)$. Diese Aussage hat nun eine unterschiedliche Bedeutung, je nachdem, was für Mengen sich in $f(N)$ befinden. Geht man z. B. vom Satz von Tarski aus, so befindet sich mit $id_{Ext}(\neg w(\alpha(x)))$ immer das Komplement von $id_{Ext}(w(\alpha(x)))$ im Bild von f. Geht man vom Unvollständigkeitssatz aus, so befindet sich mit $id_{Ext}(\neg bew(\alpha(x)))$ eine Menge im Bild von f, die disjunkt zur Menge $id_{Ext}(bew(\alpha(x)))$ ist. Es handelt sich nicht – unbedingt – um das Komplement, da die Negation relativ zur metasprachlichen Negation und bezüglich $\vdash$ als Wahrheitsbegriff nicht klassisch verwendet wird.

Die Sätze 2.4 und 2.5, die sich im Vergleich zu Satz 2.3 auf repräsentierbare Teil- oder Obermengen A von W beziehen, können als Diagonalargumente bezüglich einer Diagonalstruktur relativ zu einer Abbildung g aufgefasst werden. Für eine abstrakte Sprache S und Mengen W, A von S-Sätzen ist durch $N = F_S \setminus F'_S$, $f : N \to P(N)$ mit $f(\varphi(x)) = \{\psi(x) \in F_S \setminus F'_S | \varphi(\ulcorner\psi\urcorner) \in W\}$, $g : N \to P(N)$ mit $g(\varphi(x)) = \{\psi(x) \in F_S \setminus F'_S | \varphi(\ulcorner\psi\urcorner) \in A\}$ und $m_0 = \{x \in N | x \in g(x)\}$ eine Diagonalstruktur gegeben. Satz 3.6 besagt, dass, unter der Bedingung $m_0 \in f[N]$, für jedes Urbild n_0 von m_0 und alle $n \in N$: $n \in g(n)$ gdw. $n \in f(n_0)$. Dies ist der Schluss, der auch in den Beweisen der Sätze 2.4 und 2.5 für Teil- und Obermengen A von W gezogen wird: Es gibt eine Formel $\vartheta(x)$, so dass für alle $S_{[v_0]}$-Formeln $\varphi(x)$: $\varphi(\ulcorner\varphi(x)\urcorner) \in A$ gdw. $\vartheta(\ulcorner\varphi(x)\urcorner) \in W$.

Darstellung von Satz 2.10 c), d) – Form mit Fixpunktsatz

Die Menge N' sei definiert als $\{\varphi(\alpha(\ulcorner\varphi(\alpha(x))\urcorner)) | \varphi(x) \in F_S \setminus F'_S\}$.[7] Die Abbildung $id_{Ext} : F_S \setminus F'_S \to P(F_S \setminus F'_S)$ wird ergänzt um eine Abbildung $f' : N' \to P(N')$, die einen S-Satz der Form $\varphi(\alpha(\ulcorner\varphi(\alpha(x))\urcorner))$ auf die Menge

[7] Die Definition ist als reine Abkürzung zu verstehen. Diese Menge wird in der Interpretation als Diagonalstruktur die Rolle der dortigen Menge N' übernehmen.

$\{\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner)) \in N' | \mathcal{A}_{\text{abstr}} \models \varphi(\ulcorner\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))\urcorner)\}$ abbildet:

$$\begin{array}{lcl} F_S \setminus F_S' & \to & P(F_S \setminus F_S') \\ \varphi(x) & \mapsto & j_{\text{abstr}}(\varphi(x)) \cap F_S \setminus F_S' \end{array}$$

$$\begin{array}{lcl} N' & \to & P(N') \\ \varphi(\alpha(\ulcorner\varphi(\alpha(x))\urcorner)) & \mapsto & j_{\text{abstr}}(\varphi(x)) \cap N' \end{array}$$

In Satz 2.10 c) wird eine $S_{[v_0]}$-Formel $\gamma_1(x)$ vorausgesetzt, so dass für alle S-Sätze φ gilt $\varphi \leftrightarrow \gamma_1(\ulcorner\varphi\urcorner) \in W$. Insbesondere gilt also für alle $S_{[v_0]}$-Formeln $\psi(x)$: $\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner)) \leftrightarrow \gamma_1(\ulcorner\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))\urcorner) \in W$. Übertragen auf die Struktur $\mathcal{A}_{\text{abstr}}$ bedeutet dies, dass die Menge der Selbsteinsetzungen der $S_{[v_0]}$-Formeln $\varphi(\alpha(x))$, die in ihrer Extension enthalten sind, die Menge $f'(\gamma_1(\alpha(\ulcorner\gamma_1(\alpha(x))\urcorner)))$ bildet: $\{\varphi(\alpha(\ulcorner\varphi(\alpha(x))\urcorner)) \in N' | \varphi(\alpha(x)) \in id_{Ext}(\varphi(\alpha(x)))\} = f'(\gamma_1(\alpha(\ulcorner\gamma_1(\alpha(x))\urcorner)))$.

Mit $N = \{\varphi(\alpha(x)) | \varphi(x) \in F_S \setminus F_S'\}$, $f = id_{Ext}$ und $N' = \{\varphi(\alpha(\ulcorner\varphi(\alpha(x))\urcorner)) | \varphi(x) \in F_S \setminus F_S'\}$, f', $m_1' = f'(\gamma_1(\alpha(\ulcorner\gamma_1(\alpha(x))\urcorner)))$ sind also die Voraussetzungen von Satz 3.7 gegeben.

Satz 2.10 c) baut weiter darauf auf, dass $\chi \leftrightarrow \gamma_2(\ulcorner\chi\urcorner) \in W$. Im konkreten Fall ist χ z. B. der Lügner-Satz oder der selbstbezügliche Satz, der im Fixpunktsatz konstruiert wird. Diesem Schritt entspricht in der Struktur $\mathcal{A}_{\text{abstr}}$ die Äquivalenz im Beweis zu Satz 3.7 b): $(n,n) \in g'((n,n))$ gdw. $n \in g(n)$; $n \in g(n)$ ist hierbei χ und $(n,n) \in g'((n,n))$ ist $\gamma_2(\ulcorner\chi\urcorner)$ zuzuordnen. Angewendet auf $n = \gamma_2(\alpha(x))$ und $g = f$, $g' = f'$ lautet sie: $\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in f'(\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)))$ gdw. $\gamma_2(\alpha(x)) \in f(\gamma_2(\alpha(x))$.[8] Da m_1' im Bild der Abbildung f' liegt, ergibt Satz 3.4 a) für alle $S_{[v_0]}$-Formel $\gamma_2(x)$ nach Bestimmung von n_2' als $n_2' = \gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))$:

$$\begin{array}{l} \gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in f'(\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))) \quad \text{gdw.} \\ \gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in f'(\gamma_1(\alpha(\ulcorner\gamma_1(\alpha(x))\urcorner))). \end{array}$$

Dies entspricht der Konklusion von Satz 2.10 c): $\gamma_2(\ulcorner\chi\urcorner) \leftrightarrow \gamma_1(\ulcorner\chi\urcorner) \in W$.

[8] Der Fixpunktsatz wird häufig als “Diagonalization Lemma” bezeichnet, was den Bezug zu Cantors Diagonalverfahren andeutet.

In Satz 2.10 d) wird eine $S_{[v_0]}$-Formel $\gamma_2(x)$ vorausgesetzt, so dass für alle S-Sätze φ gilt $\neg\varphi \leftrightarrow \gamma_2(\ulcorner\varphi\urcorner) \in W$. Insbesondere gilt also für alle $S_{[v_0]}$-Formeln $\psi(x)$: $\neg\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner)) \leftrightarrow \gamma_2(\ulcorner\psi(\alpha(\ulcorner\psi(\alpha(x))\urcorner))\urcorner) \in W$. Übertragen auf die Struktur $\mathcal{A}_{\text{abstr}}$ bedeutet dies, dass die Menge der Selbsteinsetzungen der $S_{[v_0]}$-Formeln $\varphi(\alpha(x))$, die nicht in ihrer Extension enthalten sind, die Menge $f'(\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)))$ bildet: $\{\varphi(\alpha(\ulcorner\varphi(\alpha(x))\urcorner)) \in N' | \varphi(\alpha(x)) \notin id_{Ext}(\varphi(\alpha(x)))\} = f'(\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)))$.

Wie oben, jetzt mit $m'_2 = f'(\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)))$, ergibt sich nun eine Diagonalstruktur$_2$, wobei das Diagonalelement im Bild von f' liegt. Die Anwendung der Selbstbezüglichkeit des Lügner-Satzes oder des Fixpunktsatzes zeigt sich wieder in der Äquivalenz im Beweis zu Satz 3.7 b): $(n,n) \in g'((n,n))$ gdw. $n \in g(n)$. Die Selbstbezüglichkeit wird hier benutzt in der Form $\neg\chi \leftrightarrow \neg\gamma_2(\ulcorner\chi\urcorner) \in W$.

Die Anwendung des Beweises von Satz 3.4 b) ergibt die Äquivalenz:

$$\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \in f'(\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))) \quad \text{gdw.}$$
$$\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \notin f'(\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))).$$

Dies entspricht der Konklusion von Satz 2.10 c): $\gamma_2(\ulcorner\chi\urcorner) \leftrightarrow \neg\gamma_2(\ulcorner\chi\urcorner) \in W$.

3.4 Vergleich ontologischer und semantischer Antinomien

In den Abschnitten 3.2 und 3.3 werden ontologische und semantische Antinomien als Anwendungen von Diagonalargumenten in Diagonalstrukturen dargestellt. Über diese Beziehung entstehen Analogien zwischen einzelnen Voraussetzungen in den ontologischen Antinomien und entsprechenden Aussagen für den semantischen Fall, und umgekehrt zwischen Voraussetzungen, die in semantischen Antinomien gemacht werden, und Aussagen für den ontologischen Fall. Eine Untersuchung dieser Analogien erfolgt in Abschnitt 3.4.1. Die in den Antinomien implizit oder explizit getroffenen Annahmen werden im Anschluss in Abschnitt 3.4.2 in ihrem Verhältnis zu Diagonalargumenten und -strukturen betrachtet und nach drei verschiedenen Typen kategorisiert.

3.4.1 Analogien in den beiden Antinomietypen

Die ontologischen Antinomien sind so aufgebaut, dass direkt über Klassen als Elemente gesprochen wird, die gleichzeitig auch als Zusammenfassungen von Elementen auftreten. Die Klassentheorie wird dabei extensional verstanden. Zwei Klassen sind identisch, wenn sie die gleichen Elemente haben. Die Abbildung $id_{R_\in} : U \to P(U)$ ist also injektiv.[9] Dem stehen in den semantischen Antinomien Formeln als Objekte gegenüber. Die Abbildung id_{Ext} von einer Teilmenge von U in die Potenzmenge ist also in der Regel nicht injektiv. Im Folgenden werden einige Analogien aufgeführt.

Die Klasse aller Klassen
In der Antinomie von Cantor spielt die Existenz einer Klasse aller Klassen eine wesentliche Rolle. Betrachtet man diese Annahme innerhalb der Diagonalstruktur, die sich wie in Abschnitt 3.2 beschrieben aus der Interpretation der objektsprachlichen Argumentation in der Struktur $\mathcal{A}_{\mathrm{Kl}}$ ergibt, so zeigt sich, dass die Klasse aller Klassen der Menge U in $P(U)$ entspricht und die Forderung ihrer Existenz der Tatsache, dass U im Bild von $id_{R_\in}$ liegt, entspricht. Überträgt man dies auf eine Diagonalstruktur, wie sie aus den semantischen Antinomien entsteht, also durch Interpretation der Objektsprache in der Struktur $\mathcal{A}_{\mathrm{abstr}}$, ergibt sich als Entsprechung zur Klasse aller Klassen die Menge aller einstelligen Formeln der Objektsprache. Der Forderung der Existenz einer Klasse aller Klassen entspricht also die Forderung, dass die Menge aller einstelligen Formeln im Bild von id_{Ext} liegt bzw. dass eine $S_{[v_0]}$-Formel $\gamma(x)$ existiert, so dass für alle $S_{[v_0]}$-Formeln $\psi(x)$: $\gamma(\ulcorner\psi(x)\urcorner) \in W$.

Eine solche einstellige Formel ist in den meisten konkreten Sprachen vorhanden und führt nicht wie die Klasse aller Klassen zu einer Antinomie. Die Klasse aller Klassen wird in der Antinomie von Cantor ihrer Potenzklasse gegenübergestellt. Der Potenzklasse einer Klasse entspricht in der Struktur $\mathcal{A}_{\mathrm{Kl}}$ die Menge $\{x \in id_{R_\in}[U] | x \subseteq a\}$ zu einer Menge a. Dem Potenzklassenaxiom könnte im semantischen Fall also die Forderung entsprechen, dass es zu jeder $S_{[v_0]}$-Formel $\varphi(x)$ eine $S_{[v_0]}$-Formel $p_{\varphi(x)}(x)$ gibt, so dass für alle $S_{[v_0]}$-Formeln $\psi(x)$: $p_{\varphi(x)}(\ulcorner\psi(x)\urcorner) \in W$ gdw. $[\psi(\ulcorner\vartheta(x)\urcorner) \in W$ impliziert $\varphi(\ulcorner\vartheta(x)\urcorner) \in W$ für alle

[9] In einer intensionalen Variante werden Eigenschaften von Eigenschaften betrachtet und gebildet. Die entsprechende Abbildung $id_{R_\in} : U \to P(U)$ ist dann nicht mehr injektiv. Vgl. S. 119 zur Russellschen Antinomie für Eigenschaften.

$S_{[v_0]}$-Formeln $\vartheta(x)$]. Für die $S_{[v_0]}$-Formel $\gamma(x)$, die auf alle $S_{[v_0]}$-Formeln zutrifft, und die zugehörige $S_{[v_0]}$-Formel $p_{\gamma(x)}(x)$ würde damit gelten: $p_{\gamma(x)}(\ulcorner\psi(x)\urcorner) \in W$ gdw. $\gamma(x)(\ulcorner\psi(x)\urcorner) \in W$ für alle $S_{[v_0]}$-Formeln $\psi(x)$.

Die Anwendung des Aussonderungsschemas in der Antinomie von Cantor bewirkt, dass zu jeder Klasse jede in der Sprache S_{Kl} beschreibbare Zusammenfassung von Elementen der Klasse selbst eine Klasse, also Element der Potenzklasse ist. Insbesondere ist das Diagonalelement$_2$ im Bild von $id_{R_\in}[U]$ enthalten, da U selbst im Bild liegt, was zur Antinomie führt. Das Aussonderungsschema ist so nicht auf den semantischen Fall übertragbar, da die Abbildung $id_{R_\in}[U]$ in der Sprache der Klassentheorie ausdrückbar ist, nicht jedoch in den Sprachen der semantischen Antinomien. Jedenfalls würde zu einer Entsprechung gehören, dass es zu jeder $S_{[v_0]}$-Formel $\psi(x)$ und jedem metasprachlichen $\{\ulcorner\cdot\urcorner\}_{[v_0]}$-Term $F(x)$[10] eine $S_{[v_0]}$-Formel $\varphi(x)$ gibt, so dass für alle $S_{[v_0]}$-Formeln $\vartheta(x)$ gilt: $[\psi(\ulcorner\vartheta(x)\urcorner) \in W$ und $F(\vartheta(x)) \in W]$ gdw. $\varphi(\ulcorner\vartheta(x)\urcorner) \in W$. Speziell gäbe es also eine einstellige Formel $het(x)$, für die gilt: $[\gamma(\ulcorner\vartheta(x)\urcorner) \in W$ und $\neg\vartheta(\ulcorner\vartheta(x)\urcorner) \in W]$ gdw. $het(\ulcorner\vartheta(x)\urcorner) \in W$ für alle $S_{[v_0]}$-Formeln $\vartheta(x)$. Also: $\neg\vartheta(\ulcorner\vartheta(x)\urcorner) \in W$ gdw. $het(\ulcorner\vartheta(x)\urcorner) \in W$, woraus sich parallel zur Antinomie von Cantor: $\neg het(\ulcorner het(x)\urcorner) \in W$ gdw. $het(\ulcorner het(x)\urcorner) \in W$ ergibt. Im klassentheoretischen Fall besteht ein gravierender Unterschied zwischen einem uneingeschränkten Abstraktionsschema und einem – eingeschränkten – Aussonderungsschema. In den Sprachen der semantischen Antinomien ist die Situation eine andere, da aus der klassischen Verwendung der Negation und dem – eingeschränkten – Aussonderungsschema die uneingeschränkte Abstraktionsmöglichkeit folgt. Denn zu jeder $S_{[v_0]}$-Formel $\psi(x)$ gibt es die dazu negierte Formel $\neg\psi(x)$. Setzt man diese an Stelle von $\psi(x)$ und verknüpft die sich ergebenden Formeln disjunktiv, so erhält man eine Formel der uneingeschränkten Abstraktion.

Es besteht also im semantischen Fall nicht der Unterschied in den Voraussetzungen einer solchen Antinomie und der Antinomie von Grelling wie im ontologischen Fall zwischen den Voraussetzungen der Antinomie von Cantor und der Antinomie von Russell. Denn das übertragene – eingeschränkte – Aussonderungsschema fällt mit dem uneingeschränkten Abstraktionsschema zusammen. Eine einstellige Formel, die auf alle einstelligen Formeln zutrifft, das Pendant zur Allklasse also, ist damit für sich genommen unproblematisch.

[10] $F(x)$ bilde eine $S_{[v_0]}$-Formel – eventuell unter Benutzung von $\ulcorner\cdot\urcorner$ – auf einen S-Satz ab.

Die Klasse aller Klassen, die sich nicht selbst enthalten

Die Existenz einer Klasse aller Klassen, die sich nicht selbst enthalten, ist eine der wesentlichen Voraussetzungen der Antinomie von Russell. Interpretiert in der Struktur $\mathcal{A}_{\mathrm{Kl}}$ und innerhalb der entsprechenden Diagonalstruktur betrachtet, bewirkt diese Annahme, dass die Menge aller Mengen aus U, die nicht in ihrem Bild unter $id_{R_\in}$ enthalten sind, im Bild von $id_{R_\in}$ liegt. Übertragen auf die Struktur $\mathcal{A}_{\mathrm{abstr}}$ bedeutet dies, dass die Menge aller Mengen aus $F_S \setminus F'_S$, die nicht in ihrem Bild unter id_{Ext} enthalten sind, in id_{Ext} liegt; dass es also eine $S_{[v_0]}$-Formel $het(x)$ gibt, die auf die Standardnamen genau der einstelligen Formeln zutrifft, die nicht auf ihre Standardnamen zutreffen. Die genaue Parallele zur Antinomie von Russell ist damit die Antinomie von Grelling.

Der $S_{[v_0]}$-Formel $aut(x)$ entspricht im ontologischen Fall die Klasse aller Klassen, die sich selbst enthalten. Der zum Widerspruch führenden, einzusetzenden Formel $\neg aut(x)$ entspricht das Komplement der Klasse aller Klassen, die sich selbst enthalten. Der Voraussetzung, dass zu jeder Menge $G \in id_{Ext}[F_S \setminus F'_S]$ das Komplement $(F_S \setminus F'_S) \setminus G$ in $id_{Ext}[F_S \setminus F'_S]$ liegt, entspricht die Forderung, dass zu jeder Klasse die Zusammenfassung aller Elemente, die nicht in der Klasse liegen, als Klasse existiert. Geht man im Spezialfall von der Existenz einer Klasse aller Klassen aus, die sich nicht selbst enthalten, so lässt sich folgende Parallele zur Antinomie von Grelling in der "autologisch"-Form bilden: Die Klasse aller Klassen, die sich nicht selbst enthalten, ist in der Klasse aller Klassen, die sich nicht selbst enthalten, enthalten gdw. sie in der Klasse aller Klassen, die sich selbst enthalten, enthalten ist – $\{x|x \notin x\} \in \{x|x \notin x\}$ gdw. $\{x|x \notin x\} \in \{x|x \in x\}$.

Das Wahrheitsprädikat

In der Formulierung der Lügner-Antinomie wird eine $S_{[v_0]}$-Formel $w(x)$ vorausgesetzt, die auf Standardnamen von Sätzen der Form $\varphi(\alpha(\ulcorner\varphi(\alpha(x))\urcorner))$ genau dann zutrifft, wenn $\varphi(\alpha(x))$ auf seinen eigenen Standardnamen zutrifft. In Abschnitt 3.3 wird diese Formel in der Struktur $\mathcal{A}_{\mathrm{abstr}}$ interpretiert und als Diagonalelement$_1$ gedeutet. Die zur Diagonalstruktur gehörende Menge N' besteht aus allen S-Sätzen der Form $\varphi(\alpha(\ulcorner\varphi(\alpha(x))\urcorner))$. Sätzen dieser Form entsprechen in der Struktur $\mathcal{A}_{\mathrm{Kl}}$ geordnete Paare (x, x) mit zwei gleichen Einträgen von Elementen aus $x \in U$.

Die Abbildung f' als Bestandteil der Diagonalstruktur bildet einen S-Satz $\varphi(\alpha(\ulcorner\varphi(\alpha(x))\urcorner))$ ab auf die Menge $\{\psi(\alpha\ulcorner\psi(\alpha(x))\urcorner) \in F'_S|\mathcal{A}_{\mathrm{abstr}} \models \varphi(\ulcorner\psi(\alpha\ulcorner\psi(\alpha(x))\urcorner)\urcorner)\}$. Ein geordnetes Paar (x, x), mit $x \in U$, würde entsprechend auf die Menge $\{(y,y)|y \in id_{R_\in}(x)\}$ abgebildet werden. Diagonalelement$_1$ ist in diesem Fall die Menge $\{(y,y)|y \in id_{R_\in}(y)\}$. Diese Menge liegt im Bild der Abbildung f', wenn es eine entsprechende objektsprachliche Klasse $a = \{x|x \in x\}$ gibt. Dann wird $(j_\in(a), j_\in(a))$ durch f' auf $\{(y,y)|y \in id_{R_\in}(y)\}$ abgebildet. Der $S_{[v_0]}$-Formel $w(x)$ entspricht also die Klasse $\{(x,x)|x \in x\}$. Die Anwendung des Diagonalarguments auf diese Diagonalstruktur$_1$ bzw. die Anwendung von Satz 3.4 a) ergibt die Äquivalenz: Für alle (x,x) mit $x \in U$: $(x,x) \in \{(y,y)|y \in id_{R_\in}(x)\}$ gdw. $(x,x) \in \{(y,y)|y \in id_{R_\in}(y)\}$. Oder in klassentheoretischer Objektsprache: $\bigwedge x : (x,x) \in \{(y,y)|y \in x\} \leftrightarrow (x,x) \in \{(y,y)|y \in y\}$.

Setzt man für $\{(y,y)|y \in x\}$ das Komplement der Klasse $\{(y,y)|y \in y\}$ ein, für x also die Klasse $\{y|y \notin y\}$, erhält man als Parallele zur Lügner-Antinomie: Das geordnete Paar mit dem zweifachen Eintrag der Klasse aller Klassen, die sich nicht selbst enthalten, ist in der Klasse aller geordneten Paare mit identischen Einträgen, die sich nicht selbst enthalten, enthalten gdw. es in der Klasse aller geordneten Paare mit identischen Einträgen, die sich selbst enthalten, enthalten ist: $(\{y|y \notin y\}, \{y|y \notin y\}) \in \{(y,y)|y \notin y\} \leftrightarrow (\{y|y \notin y\}, \{y|y \notin y\}) \in \{(y,y)|y \in y\}$.

Ist zusätzlich zu einer Formel $w(x)$ die Möglichkeit der Einsetzung von Termen in einstellige Formeln in der Sprache gegeben, z. B. durch einen zweistelligen Term $\alpha(y,z)$, so ist mit $w(\alpha(y,z))$ die Möglichkeit gegeben, zwischen Formeln als Objekten und Formeln als Zusammenfassungen hin- und herzuwechseln, wie dies in der klassentheoretischen Sprechweise möglich ist. Die metasprachliche Abbildung id_{Ext} wird quasi objektsprachlich ausdrückbar. Insofern ist mit einer Formel $w(x)$ nicht nur die Entsprechung zur Klasse $\{(x,x)|x \in x\}$ gegeben, sondern auch eine Art uneingeschränkte Abstraktionsmöglichkeit. Das Wahrheitsschema – für alle $S_{\text{Lü}}$-Sätze φ: $\mathcal{A}_{\text{Lü}} \models \varphi \leftrightarrow w(\ulcorner\varphi\urcorner)$ – entspricht in diesem Sinn dem uneingeschränkten Abstraktionsschema der naiven Klassentheorie: $\bigvee x \bigwedge y : y \in x \leftrightarrow \psi(y)$ für alle $S_{\text{Kl}[v_0]}$ Formeln $\psi(y)$.

3.4.2 Voraussetzungen in den beiden Antinomietypen

Die Voraussetzungen exemplarischer Antinomien werden daraufhin untersucht, in welchen Bereich des Gesamtkomplexes von Sprache sie fallen und was ihre Funktion bei der Bildung einer Diagonalstruktur und der Anwendung eines Diagonalarguments ist. Hierbei wird zurückgegriffen auf die Abschnitte 3.2 und 3.3, in denen die ontologischen bzw. semantischen Antinomien als Anwendung eines Diagonalarguments in Form von Satz 3.4 auf eine Diagonalstruktur dargestellt werden.

In der Formulierung der Antinomie von Cantor wird vorausgesetzt, dass eine Klasse aller Klassen existiert. Dies ist eine ontologische Voraussetzung, denn es wird innerhalb der Objektsprache die Existenz einer bestimmten Zusammenfassung von Objekten als Klasse gefordert. Zusammen mit dem Potenzklassenaxiom und dem Aussonderungsschema, die ebenfalls ontologische Voraussetzungen darstellen, ergibt sich die objektsprachliche Äquivalenz

$$\{x|x \notin x\} \notin \{x|x \notin x\} \leftrightarrow \{x|x \notin x\} \in \{x|x \notin x\}.$$

Was nun aus dieser Äquivalenz folgt, ob sie zum Problem wird, weil jeder Satz logisch folgt, hängt von der in der Objektsprache verwendeten Logik ab. Es kommt also eine die Logik betreffende Voraussetzung hinzu, bevor die Antinomie entsteht.

Innerhalb der zugehörigen Diagonalstruktur bewirken die ontologischen Voraussetzungen die Möglichkeit der Bildung einer Diagonalstruktur – in diesem Fall einer Diagonalstruktur$_2$ –, deren Diagonalelement im Bild der zugehörigen Abbildung f liegt.[11] Die die Logik betreffende Voraussetzung, die in der Annahme einer klassischen Logik besteht, bewirkt die Ableitbarkeit jedes objektsprachlichen Satzes aus der obigen Äquivalenz. Oder, auf die Diagonalstruktur bezogen, bewirkt sie, dass die Wahrheitswertbestimmung eines Satzes klassisch geschehen kann – ausgehend von einer zweiwertigen Semantik und klassischer metasprachlicher Logik – und sich der Widerspruch so im metasprachlichen Diagonalargument ergibt.

Die Antinomie von Russell enthält ganz analog zur Antinomie von Cantor ontologische und die Logik betreffende Voraussetzungen. Die ontologische Voraussetzung besteht hier in der direkten Annahme einer Klasse aller Klassen, die sich nicht selbst enthalten.

In der Formulierung der Lügner-Antinomie wird vorausgesetzt, dass es in der Sprache eine einstellige Formel gibt, die genau auf die Standardnamen aller wah-

[11]Vgl. Definition 3.3 und Abschnitt 3.2.

ren Sätze zutrifft. Dies ist eine Voraussetzung bezüglich der Ausdrucksfähigkeit der Sprache. Diese Voraussetzung führt zu der objektsprachlichen Äquivalenz

$$\neg w(\ulcorner\chi\urcorner) \leftrightarrow w(\ulcorner\chi\urcorner).$$

Was aus dieser Äquivalenz folgt, ob sie zum Problem wird, weil jeder Satz logisch folgt, hängt von der Logik bzw. der Semantik der Objektsprache ab. Denn im Vergleich beispielsweise zu den ontologischen Antinomien scheint hier die Voraussetzung nicht nur als Voraussetzung bezüglich der objektsprachlichen Logik interpretierbar zu sein, sondern auch als Voraussetzung bezüglich der objektsprachlichen Semantik, genauer bezüglich der in der Objektsprache ausgedrückten Semantik der Objektsprache. Identifiziert man $\neg w$ mit "falsch", so sagt obige Äquivalenz über einen bestimmten Satz, dass er falsch ist genau dann, wenn er wahr ist. Verwendet man nun z. B. statt einer zweiwertigen eine dreiwertige Semantik, ergibt sich kein Widerspruch. Von einer solchen Änderung bleibt die Logik der Objektsprache jedoch nicht unberührt. Es muss festgelegt werden, wie sich die logischen Operatoren semantisch verhalten. Selbst bei Verwendung der klassischen Logik für alle Sätze mit den Wahrheitswerten "wahr" oder "falsch" gelten logische Gesetze wie $\varphi \vee \neg\varphi$ nicht für Sätze φ mit Wahrheitswert "unbestimmt". Die Bereiche Logik und Semantik werden daher hier zu einem Bereich zusammengefasst.

Innerhalb der zugehörigen Diagonalstruktur bewirkt die Voraussetzung bezüglich der Ausdrucksstärke der Sprache die Möglichkeit der Bildung einer Diagonalstruktur – in diesem Fall einer Diagonalstruktur$_1$ –, deren Diagonalelement im Bild der zugehörigen Abbildung f' liegt.[12] Die Voraussetzungen bezüglich Logik und Semantik, die in der Verwendung einer klassischen Logik und Semantik bestehen, bewirken, dass jeder Satz logisch folgt. Auf die Diagonalstruktur bezogen bewirken sie einen Widerspruch im metasprachlichen Diagonalargument, da die einstellige Formel $\neg w(x)$ durch das Komplement der Interpretation der einstelligen Formel $w(x)$ interpretiert wird.

Analog zur Lügner-Antinomie enthält die Antinomie von Grelling Voraussetzungen bezüglich der Ausdrucksstärke der Sprache und bezüglich der Logik und Semantik der Sprache. Die Annahme bezüglich der Ausdrucksfähigkeit der

[12] Vgl. Satz 3.7 und Abschnitt 3.3.

Sprache bildet hier die Voraussetzung eines Prädikats “heterologisch” bzw. “autologisch”, je nach Variante dieser Antinomie.

Es treten also in den dargestellten Antinomien drei verschiedene Bereiche von Sprache[13] auf, in Bezug auf die für das Zustandekommen der Antinomien relevante Voraussetzungen getroffen werden; und zwar: Ontologie, Ausdrucksstärke der Sprache, Logik und Semantik.[14] Weitere Voraussetzungen, wie die Möglichkeit durch Standardnamen auf die Formeln einer Sprache Bezug zu nehmen, werden als unproblematische Voraussetzungen angesehen und nicht weiter gesondert behandelt.

Die logisch-semantischen Voraussetzungen entsprechen sich in beiden Antinomietypen. In den ontologischen Antinomien werden zusätzlich ontologische Voraussetzungen und in den semantischen Antinomien Voraussetzungen bezüglich der Ausdrucksstärke der Sprache gemacht. In ihrer Funktion zur Erzeugung einer Diagonalstruktur und Anwendung eines Diagonalarguments entsprechen sich die ontologischen Annahmen der ontologischen Antinomien und die Annahmen bezüglich der Ausdrucksfähigkeit der Sprache der semantischen Antinomien.

In der klassentheoretischen Sprache ist die zur Diagonalstruktur gehörende Abbildung f durch die Identität ausdrückbar, da die Elementbeziehung in dem einstufigen System Elemente als Objekte mit Elementen als Zusammenfassungen verbindet. So ist die Bildung einer einstelligen Formel möglich, die das Diagonalelement darstellt. Die Existenz eines zugehörigen Objekts, also einer Klasse, die genau die Klassen enthält, auf die die Formel zutrifft, muss jedoch explizit gefordert werden. Genau dies leisten die ontologischen Voraussetzungen in den ontologischen Antinomien.

In den Sprachen der semantischen Antinomien besteht nicht schon vom Aufbau der Sprache her die Möglichkeit, die jeweilige Abbildung f innerhalb der Objektsprache auszudrücken. Einstellige Formeln als Objekte und einstellige Formeln als Zusammenfassungen sind von unterschiedlichem Typ, es gibt keine Abbildungsmöglichkeit zwischen ihnen. Verfolgt man einen typentheoretischen

[13] Sprache wird dabei in einem sehr allgemeinen Sinn verstanden, so dass auch die durch sprachliche Voraussetzungen dargestellte Ontologie einer Sprache dazu zählt.

[14] Diese Unterteilung entspricht der Unterteilung der Erklärungen für das Zustandekommen von Antinomien in Bromand (2001), wo allerdings die Bereiche Logik und Semantik getrennt behandelt werden. Vgl. auch Abschnitt 3.8 der vorliegenden Arbeit.

Ansatz einer Logik höherer Stufe mit Abstraktionsregeln, kann vorausgesetzt werden, dass es eine Abbildung gibt, die eine einstellige Formel als Objekt nullter Stufe auf die Klasse aller Klassen, auf die die Formel zutrifft, als Objekt erster Stufe abbildet. Die einstellige Formel $het(x)$ kann dann mit Hilfe dieser Abbildung konstruiert werden. Die Voraussetzung, dass es eine solche Abbildung gibt, ist allerdings auch keine objektsprachlich formulierbare, sondern eine metasprachliche Voraussetzung bezüglich der Ausdrucksfähigkeit der Objektsprache. In dem hier verfolgten Ansatz muss metasprachlich explizit eine einstellige Formel gefordert werden, die das Diagonalelement repräsentiert, die also ein bestimmtes Formelschema erfüllt, wie in der Antinomie von Grelling beispielsweise die einstellige Formel $het(x)$ das Schema $\neg\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow het(\ulcorner\varphi(x)\urcorner)$ für alle einstelligen Formeln $\varphi(x)$ erfüllt. Ist eine Formel gegeben, die das Formalschema erfüllt und damit das Diagonalelement repräsentiert, dann existiert sie auch als Objekt, auf das mit Hilfe des Standardnamens Bezug genommen werden kann. Diese Existenz muss also im Gegensatz zu den ontologischen Antinomien nicht explizit gefordert werden. Ebenso ergibt sich auch zu einer Formel die Existenz der verneinten Formel aus dem Aufbau des Systems heraus, während im ontologischen Fall eine gesonderte ontologische Annahme gemacht werden muss.

3.4.3 Abgrenzung gegen triviale Fälle von Diagonalargumenten

Der Begriff der Diagonalstruktur ist ein sehr allgemeiner Begriff und die Anwendung von Satz 3.4 hat ebenfalls einen sehr allgemeinen Charakter. So ist es nicht erstaunlich, dass sich auch als trivial empfundene Fälle darunter subsumieren lassen. Es kann sogar jeder Widerspruch, gegeben beispielsweise in der mengentheoretischen Form $a \in b \wedge a \notin b$, als Ergebnis eines Diagonalarguments gedeutet werden. Hierzu sei $N = b$, $f : N \to P(N)$, so dass $f(x) = \{a\}$ für alle $x \in N\setminus\{a\}$ und $f(a) = b$, $m_2 = b = \{x \in N | x \notin f(x)\}$. Dann bildet (N, f, m_2) eine Diagonalstruktur$_2$ mit einem Diagonalelement, das im Bild von f liegt. Satz 3.4 liefert also den Widerspruch $a \in b \wedge a \notin b$.

Triviale Fällen wie dieser unterscheiden sich von den Anwendungen der letzten Abschnitte in der Art, wie sie auf das Schema bezogen werden. Als Beispiel soll hier die Antinomie von Russell dienen. Das Diagonalelement wird darin *gebildet* *als* Element, das alle Klassen enthält, die sich nicht selbst enthalten. Es wird also gerade als Diagonalelement bestimmt und diese Eigenschaft wird benutzt, um einen Widerspruch abzuleiten. Da durch die Konstruktion automatisch das Diagonalelement auch im Bild der Abbildung f liegt, ergibt sich der Widerspruch. Im oben geschilderten trivialen Fall wird jedoch der Widerspruch dazu benutzt, die Diagonalstruktur erst zu erzeugen. Das Verfahren ist also genau umgekehrt. Es ist allerdings schwer, hier eine genaue Abgrenzung zu schaffen, da ein Bezug auf intensionale Aspekte der Ableitung des Widerspruchs nötig ist. Das Argument hat deshalb nur hinweisenden Charakter.

Priest weist auf das Problem der trivialen Fälle für sein *inclosure schema*, nach dem er die Antinomien darstellt, ebenfalls hin.[15] Das Diagonalelement zu einer Menge wird im Schema von Priest durch die Funktion δ gebildet. Die nichttrivialen Fälle sind in seiner Erklärung von den trivialen dadurch zu unterscheiden, dass in der Beschreibung der Funktion δ ein Zusammenhang von Argument und Wert gegeben ist, dass das Argument zur Berechnung des Werts benutzt wird. Wie Priest weiter bemerkt, ergeben sich bei einer Präzisierung des Begriffs des Zusammenhangs von Argument und Wert jedoch ebenfalls Schwierigkeiten.

Die Eigenschaft des Diagonalelements, dass es gerade als solches gebildet ist und sich nicht erst nachträglich durch einen sich ergebenden Widerspruch herausstellt, dass es die Züge eines Diagonalelements trägt, scheint für die Antinomie von Burali-Forti unter den hier geschilderten Antinomien am fragwürdigsten. Insbesondere, wenn Ordinalzahlen nicht als transitive Klassen, sondern als Äquivalenzklassen von wohlgeordneten Klassen aufgefasst werden, drängt sich der Verdacht auf, dass die sich ergebende Diagonalstruktur nur eine konstruierte ist.[16] Die Frage lässt sich reduzieren auf die Frage, ob sich das folgende Problem als ein durch ein Diagonalargument auf *natürliche* Weise lösbares Problem darstellt: a sei eine durch $<$ wohlgeordnete Klasse, die ein maximales Element b hat. Gibt es eine bijektive, ordnungserhaltende Abbildung von der Klasse $\{x \in a | x < b\}$ in die Klasse a? Nach dem Rekursionssatz, angewendet auf $<$, gibt es genau eine ordnungserhaltende Abbildung von $\{x \in a | x < b\}$ nach a, nämlich die Identität. Das Element b ist nicht im Bild der Identität, also in $\{x \in a | x < b\}$, enthalten, da es größer ist als alle Elemente dieser Klasse und

[15] Siehe Priest (2002), Kap. 9.5.

[16] Vgl. zur Antinomie von Burali-Forti auf der Basis von Äquivalenzklassen wohlgeordneter Mengen die Fußnoten 9, S. 15 und 7, S. 12.

kein Element dieser Klasse kleiner als es selbst ist. Das heißt aber nichts anderes, als dass b Diagonalelement bezüglich der Relation $<$ ist: b ist größer als alle Elemente aus $\{x \in a | x < b\}$, die nicht kleiner als ihr Bild sind.

Diese Situation stellt einen Extremfall von Diagonalstruktur dar, da kein Element mit seinem Bild in Relation steht oder – nach Interpretation der Relation – da kein Element in seinem Bild enthalten ist.

3.5 Alternative Schemata

Im Folgenden werden einige alternative Schemata wiedergegeben, die in Zusammenhang mit der Struktur von Antinomien und damit verwandten Widerspruchsbeweisen stehen. Was genau durch das jeweilige Schema erfasst werden soll, stellt sich dabei in den einzelnen Ansätzen als verschieden heraus.

Russells Darstellung der Struktur der Antinomien [17]
Priest stellt Russells Schema der Struktur der Antinomien, auf das er seine eigene Darstellung aufbaut, folgendermaßen dar: Es seien eine Eigenschaft φ und eine Funktion δ gegeben. Aus den folgenden Bedingungen ergibt sich ein Widerspruch:

1. $\Omega = \{y | \varphi(y)\}$ existiert,
2. für alle $x \subseteq \Omega$: (a) $\delta(x) \notin x$ und (b) $\delta(x) \in \Omega$.[18]

Priests Darstellung der Struktur der Antinomien[19]
Priests Darstellung ist eine Verallgemeinerung der Darstellung von Russell, mit der auch semantische Antinomien erfasst werden sollen. In seiner eigenen Formulierung lautet das Schema: Es seien Eigenschaften φ und ψ und eine Funktion δ gegeben. Aus den folgenden Bedingungen ergibt sich ein Widerspruch:

1. $\Omega = \{y | \varphi(y)\}$ existiert und $\psi(\Omega)$,

[17] Das Schema zur Struktur der Antinomien von Russell ist der Wiedergabe in Priest (2002), S. 129 ff. entnommen. Priest bezieht sich dabei auf Russell (1905).

[18] Die Bezeichnungen und Formulierungen sind die aus Priest (2002). Eine sinngemäße Übertragung in die Sprechweise der vorliegenden Arbeit könnte lauten: Aus folgenden Bedingungen ergibt sich ein Widerspruch: Ω sei eine Menge und $\delta : P(\Omega) \to \Omega$ eine Abbildung. Für alle $x \subseteq \Omega$ gelte $\delta(x) \notin x$.

[19] Vgl. Priest (2002), S. 133 ff.

2. für alle $x \subseteq \Omega$ mit $\psi(x)$: (a) $\delta(x) \notin x$ und (b) $\delta(x) \in \Omega$.[20]

Priest nennt 1) die Existenz-, 2)(a) die Transzendenz-, 2)(b) die Abgeschlossenheitsbedingung. Alle Bedingungen zusammen bezeichnet er als Inclosure-Schema. Der Widerspruch ergibt sich, da sowohl $\delta(\Omega) \in \Omega$ als auch $\delta(\Omega) \notin \Omega$.

Einige der von Priest innerhalb seines Schemas rekonstruierten Antinomien sollen nun in dieser Rekonstruktion dargestellt werden. Zur Rekonstruktion der ersten beiden Antinomien reicht das Schema von Russell aus.

Die Antinomie von Russell: Priest fasst die Antinomien von Russell und Cantor als eine einzige auf, indem er wie in Abschnitt 1.1.4 die Antinomie von Russell mit der Anwendung des Beweises des Satzes von Cantor auf die Klasse aller Klassen und ihre mit ihr identische Potenzklasse identifiziert.[21] Zur Darstellung der Antinomie innerhalb des Inclosure-Schemas sei Ω die Klasse aller Klassen. δ bilde eine Klasse von Klassen x auf die Klasse $\{y \in x | y \notin y\}$ ab. Für alle $x \subseteq \Omega$ gilt $\{y \in x | y \notin y\} \notin x$, da aus $\{y \in x | y \notin y\} \in x$ sowohl $\{y \in x | y \notin y\} \in \{y \in x | y \notin y\}$ als auch $\{y \in x | y \notin y\} \notin \{y \in x | y \notin y\}$ folgt. Für alle $x \subseteq \Omega$ gilt ferner $\{y \in x | y \notin y\} \in \Omega$. Es folgt also der Widerspruch: $\{y \in \Omega | y \notin y\} \in \Omega$ und $\{y \in \Omega | y \notin y\} \notin \Omega$.[22]

In einer Fußnote[23] gibt Priest eine weitere Möglichkeit der Rekonstruktion dieser Antinomie innerhalb des Schemas an, ohne auf die Beziehung dieser beiden Rekonstruktionen einzugehen: Ω sei die Klasse aller Klassen, die sich nicht selbst enthalten. δ sei die identische Abbildung. Für alle $x \subseteq \Omega$ gilt $x \notin x$, da aus $x \in x$ – nach Definition von Ω – $x \notin x$ folgt. Für alle $x \subseteq \Omega$ gilt damit ferner $x \in \Omega$. Es folgt also der Widerspruch: $\Omega \in \Omega$ und $\Omega \notin \Omega$.

Die Antinomie von Burali-Forti: Ω sei die Klasse aller Ordinalzahlen. δ bilde eine Klasse von Ordinalzahlen ab auf die kleinste Ordinalzahl, die echt größer ist als alle Ordinalzahlen der Klasse. Für alle $x \subseteq \Omega$ gilt offensichtlich $\delta(x) \notin x$

[20]Eine sinngemäße Übertragung in die Sprechweise der vorliegenden Arbeit könnte lauten: Aus folgenden Bedingungen ergibt sich ein Widerspruch: Ω und $\Psi \subseteq P(\Omega)$ seien Mengen, so dass $\Omega \in \Psi$. $\delta : \Psi \to \Omega$ eine Abbildung, so dass für alle $x \in \Psi$ gilt $\delta(x) \notin x$.

[21]Siehe Priest (2002), S. 130.

[22]Wie in Abschnitt 1.1 werden, abweichend von Priests Darstellung, die ontologischen Antinomien in einer Klassensprechweise formuliert. Priest bezieht die bei der Bildung der Antinomien verwendeten Klassen und Abbildungen direkt auf das Inclosure-Schema und nicht über die Interpretation in einer Struktur wie in den Abschnitten 3.2 und 3.3. Daher werden auch die Zusammenfassungen des Inclosure-Schemas jetzt als Klassen bezeichnet.

[23]Siehe Priest (2002), Fußnote 8, S. 130.

und $x \in \Omega$. Der Widerspruch ergibt sich dann, da $\delta(\Omega) = \Omega$, als: $\Omega \in \Omega$ und $\Omega \notin \Omega$.

Die Antinomie von Berry: Hier spielt zum ersten Mal die zusätzliche Eigenschaft ψ eine Rolle. Ω sei die Menge der durch weniger als 20 Worte definierbaren natürlichen Zahlen. ψ sei die Eigenschaft, durch weniger als 11 Worte definierbar zu sein. δ bilde eine durch weniger als 11 Worte definierbare Menge $x \subseteq \Omega$ ab auf die kleinste natürliche Zahl, die nicht in x enthalten ist. Für alle $x \subseteq \Omega$ gilt $\delta(x) \notin x$ und $\delta(x) \in \Omega$, da "die kleinste natürliche Zahl, die nicht in x enthalten ist" 9 Wörter und "x" enthält. Der Widerspruch ergibt sich als: Die kleinste nicht mit weniger als 20 Worten definierbare natürliche Zahl gehört zu Ω und gehört nicht zu Ω.

Die Lügner-Antinomie: Ω sei die Menge aller wahren Sätze einer Sprache. Die Eigenschaft ψ besage für eine Menge, dass sie in der Sprache definierbar ist. δ bilde eine Menge von wahren Sätzen, die durch die einstellige Formel $\gamma(x)$ definiert wird (hier muss für jede Menge eine Definition ausgewählt werden), auf den Satz $\neg\gamma(\alpha(\ulcorner\neg\gamma(\alpha(x))\urcorner))$[24] ab, wobei der einstellige Term $\alpha(x)$ die Selbsteinsetzungsfunktion repräsentiere. Für jede Menge von wahren Sätzen z, die durch $\gamma(x)$ definiert wird, gilt $\neg\gamma(\alpha(\ulcorner\neg\gamma(\alpha(x))\urcorner)) \notin z$, da aus $\neg\gamma(\alpha(\ulcorner\neg\gamma(\alpha(x))\urcorner)) \in z$ insbesondere die Wahrheit von $\neg\gamma(\alpha(\ulcorner\neg\gamma(\alpha(x))\urcorner))$ und damit $\neg\gamma(\alpha(\ulcorner\neg\gamma(\alpha(x))\urcorner)) \notin z$ folgt, da z durch $\gamma(x)$ definiert ist und $\alpha(x)$ die Selbsteinsetzungsfunktion repräsentiert. Für jede solche Menge von wahren Sätzen z gilt damit ferner $\neg\gamma(\alpha(\ulcorner\neg\gamma(\alpha(x))\urcorner)) \in \Omega$, was aus $\neg\gamma(\alpha(\ulcorner\neg\gamma(\alpha(x))\urcorner)) \notin z$ folgt.

Der Widerspruch ergibt sich hier also als: $\neg w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner))$ ist wahr und ist nicht wahr, wobei $w(x)$ die Menge der wahren Sätze definiere.

Die Antinomie von Grelling: Ω sei die Menge aller einstelligen Formeln, die nicht auf sich selbst – auf ihren Namen – zutreffen. Die Eigenschaft ψ sei die Eigenschaft einer Menge, in der Sprache definierbar zu sein. δ bilde eine durch die einstellige Formel $\vartheta(x)$ definierte Menge von einstelligen Formeln (hier muss für jede Menge eine Definition ausgewählt werden), die nicht auf sich selbst zutreffen, auf die Formel $\vartheta(x)$ ab. Für jede Menge $x \subseteq \Omega$, die durch $\vartheta(x)$ definiert wird, gilt $\vartheta(x) \notin x$, da aus $\vartheta(x) \in x$ auf Grund der Definition von Ω ja $\vartheta(x) \notin x$

[24]Dieser Satz ist ein Fixpunkt zu $\neg\gamma(x)$. Priest selbst gibt diesen Satz nicht an. Er schreibt, dass durch eine geeignete Diagonalisierungstechnik ein Satz gebildet werden kann, der aussagt, nicht zur durch $\gamma(x)$ definierten Menge zu gehören. Siehe Priest, S. 144.

folgt. Ferner gilt damit für jede solche Menge x auf Grund der Defnition von Ω auch $\vartheta(x) \in \Omega$.

Der Widerspruch ergibt sich hier also als: die einstellige Formel $het(x)$ trifft auf sich selbst zu und trifft nicht auf sich selbst zu, wobei $het(x)$ die Menge der Formeln definiere, die nicht auf sich selbst zutreffen.

Kritik des Schemas von Priest und Vergleich mit dem Schema einer Diagonalstruktur

Zur Beantwortung der Frage, in welchem Verhältnis das Schema von Priest und die Einordnung der Antinomien darunter zu dem Schema einer Diagonalstruktur im Sinne der vorliegenden Arbeit stehen, ist zunächst eine nähere Betrachtung dieses Schemas unter Zuhilfenahme der Erläuterungen von Priest hierzu und eine Einordnung in den Gesamtzusammenhang bei Priest sinnvoll.[25]

Priests Thema umfasst deutlich mehr als die Analyse ontologischer und semantischer Antinomien. Er beschäftigt sich mit Grenzen des Gedankens[26] im Allgemeinen. Die üblicherweise als ontologische und semantische Antinomien bezeichneten Widersprüche bilden nur einen Teil dieser Grenzen. Es gehören ferner Berkeleys Paradox, die Antinomien von Kant, Hegels Unendlichkeiten und weitere Grenzphänomene der Philosophiegeschichte dazu. Diese werden von Priest zunächst eingeteilt nach der Art, nach der die Grenzen des Gedankens erreicht werden. Für die Antinomien sind dabei relevant: die Grenze der Iteration, woraus die Paradoxien der absoluten Unendlichkeit entstehen – hierzu zählt er Kants Antinomien[27] und die Antinomien von Burali-Forti, Cantor und Russell –, die Grenze der Definierbarkeit – hierzu werden Berkeleys Paradoxie und die Antinomien von Richard und Berry gezählt – und die Grenze der Erkenntnis – hierzu werden die Lügner-Antinomie und die Antinomie von Grelling gezählt. Die übliche Einteilung in ontologische und semantische Antinomien wird von

[25] Für das Folgende vgl. Priest (2002), vor allem Kap. 9 und 10.

[26] Der Gedanke wird hierbei im Sinn von Frege verstanden. Siehe Priest (2002), S. 3.

[27] Priest entwirft eine Abwandlung der vierten Antinomie Kants und stellt sie folgendermaßen dar: Es wird der Gedanke an a gebildet, dann der Gedanke an den Gedanken an a, ..., dann der Gedanke an alle diese Gedanken, der Gedanke daran, und so weiter mit jedem Gedanken und jeder unbegrenzten Folge von Gedanken. Betrachtet man dann diese Reihe von Gedanken, so stellt man fest, dass sie kein letztes Element hat, da immer wieder ein Gedanke an ein Element gebildet werden kann. Also kann der Gedanke an die Zusammenfassung aller dieser Gedanken nicht mehr gebildet werden. Denn er käme in der Reihe der Gedanken vor – als Gedanke an eine unbegrenzte Folge von Gedanken – und müsste dann ein letztes Element der Reihe sein. Andererseits können wir den Gedanken an die Zusammenfassung bilden. Es wurde ja gerade getan. (Siehe Priest (2002), S. 100 f.)

Priest komplett abgelehnt und als irreführend bezeichnet, da es sich zum einen nicht um eine strukturelle Unterscheidung handelt und zum anderen z. B. Berkeleys Paradoxie und Kants Antinomien nicht berücksichtigt werden.[28] Hierzu soll nur kurz bemerkt werden, dass zum einen durch die Einteilung nach der Art, wie eine Grenze des Gedankens erreicht wird, wesentliche strukturelle Gemeinsamkeiten verwischt werden; denn die Antinomien von Richard und von Grelling fallen hierbei in unterschiedliche Bereiche, obwohl sie nahezu als identisch angesehen werden können.[29] Zum anderen zeigt sich in der Unterteilung in ontologische und semantische Antinomien bei aller struktureller Gemeinsamkeit doch auch ein relevanter Unterschied. Dieser bezieht sich auf die Art, wie die Voraussetzungen in der Antinomie gegeben werden: als ontologische oder als sprachliche Voraussetzungen, als logische oder als semantische Voraussetzungen.

Als gemeinsame Beschreibung, als strukturelles Schema für alle behandelten Grenzphänomene gibt Priest das oben beschriebene Inclosure-Schema an. Die Abbildung δ wird von Priest Diagonalisator genannt. Er bemerkt zu dieser Bezeichnung, dass δ nicht durch Diagonalisierung im eigentlichen Sinn definiert sein muss. Jede solche Abbildung δ ist jedoch so definiert, dass sichergestellt ist, dass der Wert eines Arguments kein Element des Arguments sein kann; und Diagonalisierung ist ein paradigmatisches Beispiel für eine solche Abbildung.[30] Unter Diagonalisierung versteht Priest das Cantorsche Diagonalverfahren.[31] In einer Diagonalstruktur$_2$ für Teilmengen entspricht dem Diagonalisator δ damit in der naheliegendsten Weise die Zuordnung des Diagonalelements m_2 zum Bild $f[N]$ unter f. Bei Priest wird der Diagonalisator auf bestimmte Teilmengen verallgemeinert: Einer Menge $f[N'] \subseteq f[N]$ ordnet der Diagonalisator die Menge $\{x \in N' | x \notin f(x)\}$ zu. Ein Diagonalisator in diesem Sinn ist also ein typisches Beispiel für eine Abbildung, die einer Menge von Mengen (und weiter jeder Teilmenge von ihr) eine Menge zuordnet, die nicht in ihr (bzw. der Teilmenge) enthalten ist. Priest spricht auch von "herausdiagonalisieren".

Der Diagonalisator muss aber bei Priest nicht diese Form haben, es kann sich um eine beliebige Abbildung handeln, die einen Zusammenhang zwischen Argument und Wert herstellt.[32] Das Schema von Priest setzt in seiner Strukturbeschrei-

[28] Siehe Priest (2002), S. 142 f.
[29] Siehe Abschnitt 2.1.1.
[30] Siehe Priest (2002), S. 130.
[31] Siehe Priest (2002), Kap. 8.3.
[32] Siehe Priest (2002), Kap. 9.5.

bung nach der Diagonalisierung an, die in ihrem inneren Aufbau nicht durch das Schema festgelegt wird. Geht man von der oben beschriebenen Entsprechung des Diagonalisators δ und der Zuordnung des Diagonalelements in einer Diagonalstruktur$_2$ aus, beschreibt das Schema von Priest, dass einerseits m_2 bzw. allgemein $\{x \in N' | x \notin f(x)\}$ nicht in $f[N]$ bzw. $f[N']$ liegen kann – auf Grund des Widerspruchs, der sich durch das Vorliegen einer Diagonalstruktur$_2$ mit Diagonalelement im Bild der Abbildung f ergeben würde –, andererseits m_2 und $\{x \in N' | x \notin f(x)\}$, aus Gründen, die in der jeweiligen Definition des Elements liegen, immer in $f[N]$ liegen. Der Widerspruch ergibt sich an der *Grenze* $f[N]$ als: $m_2 \in f[N]$ und $m_2 \notin f[N]$.

Diesem Gedanken entsprechen Priests Rekonstruktionen der Antinomie von Russell bzw. Cantor – und zwar der ersten seiner Rekonstruktionen dieser Antinomie – und der Antinomien von Burali-Forti. Nur im Fall der Antinomie von Russell wird die Abbildung δ von Priest als eine aus einer Diagonalisierung hervorgegangene Abbildung dargestellt, also als ein Diagonalisator im eigentlichen Sinn.

Die Rekonstruktionen der Lügnerantinomie und der Antinomie von Grelling und die zweite, in einer Fußnote von Priest beschriebene Rekonstruktion der Antinomie von Russell fügen sich nicht so recht diesen eigenen Erklärungen Priests zu seinem Schema. Betrachtet man z. B. die zweite Darstellung der Antinomie von Russell und vergleicht die darin vorkommenden Argumentationen mit denen der ersten, so stellt man fest, dass sie ähnliche Bestandteile enthalten. Genauer kann sogar die zweite Darstellung als aus der ersten abgeleitet angesehen werden. Allgemein lässt sich dieser Zusammenhang so beschreiben: (N, f, m_2) sei eine Diagonalstruktur$_2$ mit injektiver Abbildung f. Dann sind mit $\Omega := f[N]$ und $\delta(x) := \{z \in f^{-1}[x] | z \notin f(z)\}$ für $x \subseteq \Omega$, die Teile 1) und 2)(a) des Inclosure-Schemas erfüllt. 2)(a) ist dabei die Verallgemeinerung der Anwendung von Satz 3.4 b) auf eine Diagonalstruktur$_2$. Als $\delta(\Omega)$ ergibt sich $m_2 = \{z \in N | z \notin f(z)\}$. Teil 2)(b) des Inclosure-Schemas entspricht die Eigenschaft: $\{z \in f^{-1}[x] | z \notin f(z)\} \in f[N]$, für alle $x \subseteq f[N]$. Der Widerspruch im Schema von Priest ergibt sich als: $\{z \in N | z \notin f(z)\} \in f[N]$ und $\{z \in N | z \notin f(z)\} \notin f[N]$.

Priests Darstellung der Antinomien von Russell bzw. Cantor (die erste Darstellung) und Burali-Forti legt diesen Zusammenhang zwischen dem Inclosure-Schema und dem Schema einer Diagonalstruktur nahe. Eine weitere Möglichkeit, aus einer Diagonalstruktur einen Einsetzungsfall des Inclosure-Schemas zu erzeugen, ergibt sich wie folgt: (N, f, m_2) sei eine Diagonalstruktur$_2$ mit injektiver Abbildung f. Mit $\Omega := m_2 = \{z \in N | z \notin f(z)\}$, ψ definiert als Eigenschaft, im Bild von f zu liegen, und $\delta := f^{-1}$ sind der erste Punkt von Teil 1) und die Teile 2)(a) und (b) des Inclosure-Schemas erfüllt. 2)(a) und (b) sind dabei zusammen eine Verallgemeinerung der Anwendung von Satz 3.4 b) auf eine Diagonalstruktur$_2$. $\psi(\Omega)$ bedeutet in diesem Fall $\{z \in N | z \notin f(z)\} \in f[N]$, stellt also die Voraussetzung für Satz 3.4 b) dar. Der Widerspruch im Schema von Priest ergibt sich dann als: $f^{-1}(\{z \in N | z \notin f(z)\}) \in \{z \in N | z \notin f(z)\}$ und $f^{-1}(\{z \in N | z \notin f(z)\}) \notin \{z \in N | z \notin f(z)\}$.

Priests zweite Darstellung der Antinomie von Russell bzw. Cantor legt diesen Zusammenhang zwischen dem Inclosure-Schema und dem Schema einer Diagonalstruktur nahe. Auf diese zweite Art kann auch die Rekonstruktion der Antinomie von Grelling innerhalb einer Diagonalstruktur gedeutet werden.

Auch Priests Rekonstruktion der Lügner-Antinomie geht im Wesentlichen so aus ihrer Darstellung als Diagonalstruktur hervor. Nur handelt es sich in diesem Fall um eine Diagonalstruktur$_1$ als Grundlage. Aus einer solchen geht, parallel zu der oben geschilderten Entstehung, wie folgt ein Einsetzungsfall des Inclosure-Schemas hervor: Mit $\Omega := m_1 = \{z \in N | z \in f(z)\}$, ψ definiert als Eigenschaft, dass das Komplement im Bild von f liegt, und $\delta(x) := f^{-1}(N \setminus x)$ für alle $x \subseteq m_1$ mit $\psi(x)$, sind der erste Punkt von Teil 1) und die Teile 2)(a) und (b) des Inclosure-Schemas erfüllt. 2)(a) und (b) sind dabei zusammen eine Verallgemeinerung der Anwendung von Satz 3.4 a) auf eine Diagonalstruktur$_1$. $\psi(\Omega)$ bedeutet in diesem Fall $N \setminus \{z \in N | z \in f(z)\} \in f[N]$, stellt also die Voraussetzung dafür dar, dass sich in Satz 3.4 a) ein Widerspruch ergibt. Der Widerspruch im Schema von Priest ergibt sich dann als: $f^{-1}(N \setminus \{z \in N | z \in f(z)\}) \in \{z \in N | z \in f(z)\}$ und $f^{-1}(N \setminus \{z \in N | z \in f(z)\}) \notin \{z \in N | z \in f(z)\}$.

Das Schema von Priest ist insofern umfangreicher, als dass nicht nur die Äquivalenz: $f^{-1}(m_2) \notin m_2$ gdw. $f^{-1}(m_2) \in m_2$ betrachtet wird, wie dies in Satz 3.4 b) der Fall ist, sondern auch eine Entsprechung für Teilmengen[33] von m_2. Der eigentliche Widerspruch der Antinomie entsteht aber auch bei Priest aus der Äquivalenz, die sich auf das Diagonalelement selbst bezieht. Und auch das Argumentationsprinzip für die Implikationen, die sich auf Teilmengen des Dia-

[33]In Fällen, die über das Schema von Russell hinausgehen, gilt dies nur für Teilmengen von m_2, auf die die Eigenschaft ψ zutrifft.

gonalelements beziehen, entspricht dem für die Äquivalenz und bringt insofern nichts Neues.

Auf den Unterschied in der Art und Weise, wie die Antinomien unter das Inclosure-Schema fallen, der deutlich wird in der zweifachen Subsumierung der Antinomie von Russell bzw. Cantor unter das Schema, geht Priest nur indirekt ein. Es scheint für ihn jedenfalls nicht relevant zu sein, welche Abbildung z. B. in der Antinomie von Russell bzw. Cantor die Rolle des Diagonalisators übernimmt: die Abbildung, die eine Klasse abbildet auf die Klasse aller Klassen daraus, die sich nicht selbst enthalten, oder die Identität. Im Zusammenhang mit der Unterscheidung zwischen ontologischen und semantischen Antinomien betont er, dass der einzige strukturelle Unterschied zwischen ihnen darin besteht, wie sich Eigenschaft 2)(a) des Inclosure-Schemas – $\delta(x) \notin x$ – ergibt. Genauer unterteilt er die semantischen Antinomien weiter in solche, in denen Definitionen eine Rolle spielen, und die verbleibenden, Lügner-Antinomie und Antinomie von Grelling. Während sich Eigenschaft 2)(a) für die ontologischen und die auf Definitionen bezogenen Antinomien einfach daraus ergibt, dass die Abbildung δ aus dem Argument herausdiagonalisiert[34], wird in der Lügner-Antinomie und der Antinomie von Grelling indirekt vorgegangen. Es wird von der Annahme, dass $\delta(x) \in x$, auf $\delta(x) \notin x$ geschlossen und hierbei auf ein Wahrheits- bzw. Erfüllbarkeitsschema als *Brücken-Prinzip* Bezug genommen.[35] Für die Lügner-Antinomie und die Antinomie von Grelling gibt er nur die Variante über ein Brücken-Prinzip an. Wie oben gezeigt, lässt sich aber auch hier direkt vorgehen.

Diese Unterscheidung von Priest betrifft also keinen strukturellen Unterschied einiger Antinomien, sondern seine eigene Art, die Antinomien in unterschiedlicher Weise innerhalb seines Schemas darzustellen. Betrachtet man die zweite Rekonstruktion der Antiomie von Russell, so stellt man fest, dass hier ebenso indirekt von der Annahme, dass $\delta(x) \in x$, auf $\delta(x) \notin x$ geschlossen wird und auf ein Klassenbildungsschema als Brücken-Prinzip Bezug genommen wird. Betrachtet man die erste Rekonstruktion der Antinomie von Russell, so kann man feststellen, dass das direkte Herausdiagonalisieren der Abbildung δ aus dem Argument auf eben diesem indirekten Schluss zusammen mit dem Umkehrschluss

[34] Zu beachten ist, dass bei Priest mit Herausdiagonalisieren kein klassisches Diagonalargument gemeint sein muss.

[35] Siehe Priest (2002), S. 146.

beruht. Dies ist die oben beschriebene erste Art, von einer Diagonalstruktur zu einem Einsetzungsfall des Inclosure-Schemas zu gelangen.

Zusammenfassend kann also gesagt werden, dass eine Inkonsequenz in der Art besteht, wie Priest die Antinomien in sein Schema einordnet. Diese Inkonsequenz macht ein solches Schema unplausibel, dessen Eigenschaft, das Wesentliche der Antinomien wiederzugeben, von Natur aus schwer zu begründen ist. Ein Grund für diese Inkonsequenz ist, dass die Abbildung δ, das Herausdiagonalisieren, nicht weiter analysiert wird. Als Folge davon werden wesentliche Parallelen, die in der Darstellung als Diagonalstruktur hervortreten, nicht deutlich. Hier sind z. B. die Entsprechung einer Klasse aller Klassen, die sich nicht selbst enthalten, und eines Begriffs "heterologisch" zu nennen. Es wird weiter nicht berücksichtigt, von welcher Art die Voraussetzungen sind. Ob es sich um ontologische Voraussetzungen oder um Voraussetzungen an die Sprache, um logische oder um semantische Voraussetzungen handelt. Wesentliche Unterschiede zwischen den Antinomien werden also auch nicht deutlich. Davon betroffen ist z. B. die Frage, wie eine einheitliche Lösung von Antinomien aussieht.

Die Darstellung von Nelson und Grelling

Nelson und Grelling geben ein Diagonalschema an, das sie auf die Russellsche Antinomie, die Russellsche Antinomie für Eigenschaften und die Antinomien Burali-Forti und Grelling anwenden.[36] Das Schema entspricht im Wesentlichen dem einer Diagonalstruktur$_2$ bezogen auf Teilmengen. Die Anwendung auf die Antinomien von Russell und Grelling verläuft analog zu der hier beschriebenen.[37] Die Anwendung des Schemas auf die Russellsche Antinomie für Eigenschaften ist eine Parallele zur Anwendung auf die Russellsche Antinomie für Klassen. Die Elemente des Universums sind statt Klassen Eigenschaften. Statt einer Klasse sich selbst zuzuordnen, wird einer Eigenschaft die Klasse der Eigenschaften, auf die sie zutrifft, zugeordnet. Statt der Klasse aller Klassen, die sich nicht selbst enthalten, wird die Eigenschaft einer Eigenschaft, nicht auf sich selbst zuzutreffen, als Diagonalelement gebildet.

[36] Siehe Nelson u. Grelling (1908), S. 100 ff. Die Russellsche Antinomie für Eigenschaften bezieht sich statt auf Mengen auf Eigenschaften. Sie ist eine intensionale Variante der Antinomie von Russell. Es wird die Eigenschaft einer Eigenschaft gebildet, nicht auf sich zuzutreffen. Die Antinomie von Grelling wird an dieser Stelle erstmals formuliert.

[37] Allerdings werden genau genommen Argumente aus Satz 3.4 a), bezogen auf eine Diagonalstruktur$_1$, und Argumente aus Satz 3.4 b), bezogen auf eine Diagonalstruktur$_2$, vermischt. Ebenso wird in der Darstellung der Antiomie von Grelling sowohl auf den Begriff "heterologisch" als auch auf den Begriff "autologisch" Bezug genommen.

Nelson und Grelling versuchen bereits, auch die Antinomie von Burali-Forti mit der Antinomie von Russell in Verbindung zu bringen. Ihr Schema wenden sie folgendermaßen auf diese Antinomie an:[38] Ausgangspunkt sind die Ordinalzahlen in ihrer ursprünglichen Form als Äquivalenzklassen.[39] W sei die Klasse aller Ordinalzahlen. Nelson und Grelling wählen als Teilklasse von $W \times P(W)$ eine Relation, die nicht voreindeutig, also keine Abbildung ist. Ein Element $\alpha \in W$ steht mit einer Teilklasse $A \in P(W)$ in Relation gdw. α die kleinste Ordinalzahl ist, die größer ist als alle Ordinalzahlen aus A. Bernays verändert in einer Anmerkung diese Relation zu einer Abbildung $f : W \to P(W)$, indem er einer Ordinalzahl die Klasse aller kleineren Ordinalzahlen zuordnet.[40] In der Antinomie wird davon ausgegangen, dass W selbst, da es wohlgeordnet ist, einen Ordnungstyp besitzt. W liegt damit im Bild der Abbildung. Da kein Element in seinem Bild enthalten ist, ist W die Klasse aller Ordinalzahlen, die nicht in ihrem Bild enthalten sind. W ist also Diagonalelement. Diese Darstellung der Antinomie von Burali-Forti innerhalb des Schemas von Nelson und Grelling entspricht in ihren wesenlichen Zügen der Darstellung der Antinomie als Diagonalstruktur, wie sie in Abschnitt 3.2 gegeben wird.

Bernays bemerkt hierzu, dass die Darstellung insofern etwas Unbefriedigendes hat, als dass der Widerspruch $f^{-1}(W) \in W$ und $f^{-1}(W) \notin W$ unabhängig von der Definition von W als Diagonalelement ist.[41] Dies ist ein Einwand, auf den in Abschnitt 3.4.3 eingegangen wird.

Der Begriff des Diagonalverfahrens bei Essler

Essler entwickelt eine Präzisierung und Verallgemeinerung des intuitiven Begriffs des Cantorschen Diagonalverfahrens.[42] Das Cantorsche Diagonalverfahren wird als ein in den Antinomien von Cantor, Russell, Grelling und Richard und der Lügner-Antinomie sowie in Beweisen des Unvollständigkeits- und des Unentscheidbarkeitssatzes wesentlich verwendetes Verfahren untersucht. Es ist also nicht primäres Ziel der Arbeit, eine gemeinsame Struktur möglichst vieler Antinomien und verwandter Widerspruchsbeweise anzugeben. Der Fokus liegt von Anfang an schon auf dem Diagonalverfahren. Da dieses jedoch nach den Ausführungen dieses Kapitels gerade den wesentlichen Aspekt in der Struktur

[38] Die Darstellung ist allerdings nicht ausführlich und daher interpretationsbedürftig.
[39] Vgl. die Fußnoten 7, S. 12 und 9, S. 15.
[40] Siehe die Ergänzungen von Bernays in Nelson (1974), S. 103, Fußnote 8.
[41] Siehe die Ergänzungen von Bernays in Nelson (1974), S. 103, Fußnote 8.
[42] Für das Folgende vgl. Essler (1964).

der Antinomien ausmacht, besteht eine enge Beziehung zu der Aufgabenstellung der vorliegenden Arbeit.

Essler gibt zwei Definitionen eines Cantorschen Diagonalverfahrens an, eine, die sich auf Mengen[43] bezieht, und eine, die sich auf einstellige Formeln[44] bezieht. Sie lassen sich, übertragen auf die Sprechweise der vorliegenden Arbeit, sinngemäß wie folgt wiedergeben:

(I) [45]Vom Vorliegen des Cantorschen Diagonalverfahrens in Bezug auf eine Menge m_2 wird gesprochen, wenn mit (N, f, m_2) eine Diagonalstruktur$_2$ (für Teilmengen) gegeben ist und durch Anwendung von Satz 3.4 b) ein Widerspruch abgeleitet wird. D. h. es gilt für alle $n \in N$: $n \notin f(n)$ gdw. $n \in m_2$. Wenn nun, wie in Satz 3.4 b), angenommen wird, dass m_2 im Bild von f liegt, ergibt sich: $n_2 \notin f(n_2)$ gdw. $n_2 \in f(n_2)$.

(II) [46]Vom Vorliegen des Cantorschen Diagonalverfahrens in Bezug auf eine einstellige Formel der Form $\gamma_2(\alpha(x))$, wobei $\alpha(x)$ die Selbsteinsetzungsfunktion (der Standardterm einer einstelligen Formel wird in die Formel eingesetzt) repräsentiert und $\gamma_2(x)$ eine einstellige Formel ist, wird gesprochen, wenn $\neg\varphi(\ulcorner\varphi(x)\urcorner) \leftrightarrow \gamma_2(\alpha(\ulcorner\varphi(x)\urcorner))$ für alle einstelligen Formeln $\varphi(x)$ und durch Einsetzen von $\gamma_2(\alpha(x))$ für $\varphi(x)$ der Widerspruch $\neg\gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner)) \leftrightarrow \gamma_2(\alpha(\ulcorner\gamma_2(\alpha(x))\urcorner))$ abgeleitet wird.

Die Antinomie von Cantor, die Antinomie von Russell, die auf die vorige zurückgeführt wird, und die Antinomien von Grelling und Richard werden Schema (I) in einer Weise zugeordnet, die ihrer Auffassung als Diagonalstruktur$_2$ in den Abschnitten 3.2 und 3.3 entspricht. Die Antinomie von Burali-Forti wird von Essler nicht als Anwendung des Diagonalverfahrens gesehen, auch wenn er den Hinweis gibt, dass die Klasse aller Ordinalzahlen, die als transitive Klassen aufgefasst werden, eine Klasse von Klassen ist, die sich nicht selbst enthalten. In der Darstellung der Antinomien von Grelling (in der *heterologisch*-Form) und Richard wird wieder eine objektsprachliche Bezeichnungsrelation verwendet, die

[43]Genau genommen bezieht sich diese Definition auf Eigenschaften, bei Essler als Attribute oder Prädikate bezeichnet.

[44]Bei Essler werden diese als Aussageformen bezeichnet.

[45]Siehe Essler (1964), S. 33.

[46]Siehe Essler (1964), S. 36. Im Unterschied zu dem hier dargestellten Diagonalverfahren in Bezug auf einstellige Formeln benutzt Essler eine objektsprachliche Bezeichnungsrelation, die Sätze und von ihnen beschriebene Sachverhalte in Relation setzt.

Bezeichnungen von Eigenschaften mit Eigenschaften in Relation setzt.[47] Diese Relation entspricht der metasprachlichen Abbildung, die den Standardnamen einer einstelligen Formel auf die Formel abbildet.

Die Lügner-Antinomie wird unter das Schema (II) gefasst. Die Darstellung der Antinomie erfolgt zunächst durch die Konstruktion eines Fixpunkts der einstelligen Formel $\neg w(x)$. Diese Rekonstruktion ähnelt der Darstellung in Abschnitt 1.2 dieser Arbeit. Es folgt jedoch eine zweite Darstellung, die der Antinomie von Grelling ähnelt. Nur diese zweite Form wird weiter betrachtet und dem Schema (II) zugeordnet.[48] Auch der Unvollständigkeits- und der Unentscheidbarkeitssatz werden im Rahmen von Schema (II) behandelt.

Die getrennten Ansätze zum Diagonalverfahren für Mengen und für einstellige Formeln werden bei Essler nicht weiter vereinheitlicht. Schema (II) kann jedoch als Spezialfall von Schema (I) angesehen werden, in dem die Diagonalformel die spezielle Form $\gamma_2(\alpha(x))$ hat, N als Menge der einstelligen Formeln definiert ist und f eine einstellige Formel auf ihre Extension abbildet.

3.6 Lügner-Zirkel & Co

Goddard und Johnston geben in ihrem Artikel folgende Variante der Antinomie von Russell an, die Zirkel mit mehr als einem Glied zulässt.

Anstatt der Russellschen Klasse werden die Klassen $r_1 := \{x|\bigvee y : x \in y \wedge y \in x\}$ und $r_2 := \{x|\neg\bigvee y : x \in y \wedge y \in x\}$ gebildet.

Dann gilt: Wenn $r_2 \in r_1$, dann gibt es eine Klasse a, so dass $r_2 \in a$ und $a \in r_2$. Da $a \in r_2$, gibt es keine Klasse, die gleichzeitig a enthält und in a enthalten ist. Dies ist ein Widerspruch, da r_2 diese Bedingung erfüllt. Es gilt also $r_2 \notin r_1$ und damit $r_2 \in r_2$. Dann gibt es keine Klasse, die gleichzeitig r_2 enthält und in r_2 enthalten ist. Dies ist ein Widerspruch, da r_2 diese Bedingung erfüllt.[49]

---

[47] Vgl. S. 109.

[48] Die erste Darstellung der Lügner-Antinomie kann nicht unter das Schema subsumiert werden, da eine Variante, die sich statt auf ein Falschheitsprädikat auf ein Wahrheitsprädikat bezieht, also eine "autologisch"-Variante, darin nicht vorgesehen ist.

[49] Siehe Goddard u. Johnston (1983). Diese Antinomie kann verallgemeinert werden auf Zirkel beliebiger Länge oder unendliche Folgen $x_1, x_2, x_3...$ mit der Eigenschaft $x_1 \in x_2 \wedge x_2 \in x_3...$ Die Klassen r_1, r_2 werden dann also gebildet als Klassen aller Klassen x_1, die (nicht) erstes Glied einer unendlichen Folge $x_1, x_2, x_3...$ der obigen Art sind. Vgl. Herzberger (1970).

Der Begriff der Diagonalstruktur lässt sich analog verändern: Hierzu seien N und f wie in einer Diagonalstruktur gegeben. Für $m_1 \in P(N)$ und $m_2 \in P(N)$ gelte:

$$\begin{array}{ll} \text{Es gibt } p \text{ mit } n \in f(p) \text{ und } p \in f(n) & \text{gdw. } n \in m_1, \\ \text{es gibt kein } p \text{ mit } n \in f(p) \text{ und } p \in f(n) & \text{gdw. } n \in m_2. \end{array}$$

Liegt m_2 im Bild von f, ergibt sich ein Widerspruch analog zum obigen klassentheoretischen Fall.

Eine umgangssprachliche Übertragung auf den semantischen Fall könnte wie folgt lauten:

> Es wird ausgegangen von den Begriffen "zirkelautologisch" und "zirkelheterologisch". Ein einstelliger Begriff heißt *zirkelautologisch* gdw. es einen einstelligen Begriff gibt, der auf den ersten zutrifft und auf den der erste zutrifft. Ein einstelliger Begriff heißt *zirkelheterologisch* gdw. es keinen einstelligen Begriff gibt, der auf den ersten zutrifft und auf den der erste zutrifft.
> Es ergibt sich: Wenn der Begriff "zirkelheterologisch" zirkelautologisch ist, dann gibt es einen einstelligen Begriff, der auf "zirkelheterologisch" zutrifft und der zirkelheterologisch ist. Da dieser einstellige Begriff zirkelheterologisch ist, gibt es keinen einstelligen Begriff, der auf ihn zutrifft und auf den er zutrifft. Dies ist ein Widerspruch, da "zirkelheterologisch" diese Bedingung erfüllt. Es gilt also, dass "zirkelheterologisch" nicht zirkelautologisch, sondern zirkelheterologisch ist. Dann gibt es keinen einstelligen Begriff, der auf "zirkelheterologisch" zutrifft und der zirkelheterologisch ist. Dies ist ein Widerspruch, da "zirkelheterologisch" diese Bedingung erfüllt.[50]

[50] Eine Übertragung dieser Antinomie auf den semantischen Fall ist nicht rein objektsprachlich möglich, da in der Bestimmung der r_1 und r_2 entsprechenden Begriffe über alle Begriffe quantifiziert werden muss. Eine analoge semantische Antinomie kann mit Hilfe einstelliger Formeln ϑ_1 und ϑ_2, die folgende Bedingungen erfüllen, gebildet werden:

$$\begin{array}{ll} \text{Es gibt } \psi(x) \text{ mit } \mathcal{A} \models \psi(\ulcorner\varphi(x)\urcorner) \wedge \varphi(\ulcorner\psi(x)\urcorner) & \text{gdw. } \mathcal{A} \models \vartheta_1(\ulcorner\varphi(x)\urcorner), \\ \text{für alle } \psi(x) \text{ gilt } \mathcal{A} \models \neg(\psi(\ulcorner\varphi(x)\urcorner) \wedge \varphi(\ulcorner\psi(x)\urcorner)) & \text{gdw. } \mathcal{A} \models \vartheta_2(\ulcorner\varphi(x)\urcorner). \end{array}$$

Hintikka benutzt eine ähnliche Konstruktion wie Goddard und Johnston zur Definition einer Klasse, die nicht das Zirkelfehlerprinzip (VCP) von Russell verletzt und dennoch zu einer Antinomie führt. Dazu werden die Definitionen von r_1 und r_2 so abgeändert, dass ein Bezug auf die definierten Klassen in den Definitionen und eine direkte Anwendung der definierten Klassen auf sich selbst ausgeschlossen wird. Die Definitionen von r_1 und r_2 lauten sinngemäß:

- Für alle Klassen $x \neq r_1$: $x \in r_1$ gdw. es eine Klasse $y \neq r_1$ gibt, mit $x \in y$ und $y \in x$,
- für alle Klassen $x \neq r_2$: $x \in r_2$ gdw. es keine Klasse $y \neq r_2$ gibt, mit $x \in y$ und $y \in x$.

Die Elementbeziehung zu den beiden Klassen wird nun nicht für die Klasse r_2 selbst untersucht – dies ist ja auch durch die Definition ausgeschlossen –, sondern für die Paarklasse $\{r_1, r_2\}$. Es gilt nach Definition von r_1: $\{r_1, r_2\} \in r_1$ gdw. $\{r_1, r_2\} \in r_2$.[51] Und nach Definition von r_2: $\{r_1, r_2\} \in r_2$ gdw. $\{r_1, r_2\} \notin r_1$.[52]

Auf Grund der Äquivalenzen:

$$\begin{aligned} \{r_1, r_2\} \in r_1 \quad &\text{gdw.} \quad \{r_1, r_2\} \in r_2, \\ \{r_1, r_2\} \in r_2 \quad &\text{gdw.} \quad \{r_1, r_2\} \notin r_1 \end{aligned}$$

kann eine semantische Entsprechung dieser Antinomie als Formalisierung des Lügner-Zirkels gedeutet werden. Der Lügner-Zirkel baut auf zwei Sätzen χ_1, χ_2 der folgenden Art auf: χ_1:=“Der Satz χ_2 ist wahr”, χ_2:=“Der Satz χ_1 ist falsch”. Mit diesen Definitionen gilt: χ_1 ist wahr gdw. χ_2 ist wahr, χ_2 ist wahr gdw. χ_1 ist nicht wahr.[53]

[51] Dies gilt für alle Klassen außer r_1 an Stelle von r_2.

[52] Siehe Hintikka (1957).

[53] Einen anderen Ansatz zur Rekonstruktion des Lügner-Zirkels stellt folgende Parallele zum Fixpunktsatz dar: $S_{\text{Lü}}$-Sätz χ_1 und χ_2 mit $\mathcal{A}_{\text{Lü}} \models \chi_1 \leftrightarrow w(\ulcorner \chi_2 \urcorner)$ und $\mathcal{A}_{\text{Lü}} \models \chi_2 \leftrightarrow \neg w(\ulcorner \chi_1 \urcorner)$ lassen sich parallel zur einfachen Lügner-Antinomie durch eine Abänderung des Fixpunktsatzes konstruieren. Vorausgesetzt sei hierzu diesmal, dass es einen $S_{\text{Lü}[v_0]}$-Term $\alpha(x)$ gibt, der, angewandt auf eine $S_{\text{Lü}[v_0]}$-Formel $\varphi(x)$, die nicht mit einer Negation beginnt, semantisch der $S_{\text{Lü}}$-Formel $\varphi(\ulcorner \neg\varphi(x) \urcorner)$ entspricht, und, angewandt auf eine $S_{\text{Lü}[v_0]}$-Formel $\neg\varphi(x)$, die mit einer Negation beginnt, semantisch der $S_{\text{Lü}}$-Formel $\neg\varphi(\ulcorner \varphi(x) \urcorner)$ entspricht. Es gelte also für alle $S_{\text{Lü}[v_0]}$-Formeln $\varphi(x)$, die nicht mit einer Negation beginnen, und alle $S_{\text{Lü}[v_0]}$-Formeln $\psi(x)$: $\mathcal{A}_{\text{Lü}} \models \psi(\alpha(\ulcorner \varphi(x) \urcorner)) \leftrightarrow \psi(\ulcorner \varphi(\ulcorner \neg\varphi(x) \urcorner) \urcorner)$ und $\mathcal{A}_{\text{Lü}} \models \psi(\alpha(\ulcorner \neg\varphi(x) \urcorner)) \leftrightarrow \psi(\ulcorner \neg\varphi(\ulcorner \varphi(x) \urcorner) \urcorner)$.

Dann sind für eine $S_{\text{Lü}[v_0]}$-Formel $w(x)$, die nicht mit einer Negation beginnt, $\chi_1 :=$

3.7 Berrys Antinomie

Die Antinomie von Berry kann als einzige der im Eingangskapitel geschilderten semantischen Antinomien nicht als Speziallfall eines abstrakten Satzes aus Abschnitt 2.1 aufgefasst werden. Parallel zu den anderen sematischen Antinomien kann die Antinomie von Berry aber auch als Grundlage für Beweise des Satzes von Tarski, des ersten Gödelschen Unvollständigkeitssatzes und des Unentscheidbarkeitssatzes dienen. Die Übertragung der Antinomie von Berry auf eine semantische Form des Beweises des Unvollständigkeitssatzes wird von Boolos dargestellt.[54] Für eine syntaktische Form des Beweises des Unvollständigkeitssatzes, für den Satz von Tarski und den Unentscheidbarkeitssatz wird die Übetragung der Antinomie von Berry von Serény ausgeführt.[55] Zur genauen Durchführung sind leichte Abänderungen in der Formalisierung der Antinomie nötig. Ein Einblick kann jedoch auch auf der Basis der vorliegenden Formalisierung gegeben werden:[56]

- Die Existenz einer $S_{\mathrm{Be}[v_0,v_1]}$-Formel $def(x,y)$, so dass für alle $S_{\mathrm{Be}[v_0]}$-Formeln $\varphi(x)$ und alle $n \in \mathbb{N}$:

  $$\mathcal{A}_{\mathrm{Be}} \models \bigwedge x(\varphi(x) \leftrightarrow x \equiv \boldsymbol{n}) \quad \text{gdw.} \quad \mathcal{A}_{\mathrm{Be}} \models def(\ulcorner\varphi(x)\urcorner, \boldsymbol{n})$$

  führt in der Antinomie von Berry zum Widerspruch. Insbesondere ist also für die Symbolmenge S_{Ar} und die Struktur $\mathcal{N}$ die Menge der Gödelnummern von in $\mathcal{N}$ wahren S_{Ar}-Sätzen nicht in $\mathcal{N}$ repräsentierbar. Es ergibt sich also der Satz von Tarski.
- Wählt man statt der Wahrheit in der Struktur $\mathcal{N}$ die Ableitbarkeit aus einer aufzählbaren Menge Φ von S_{Ar}-Sätzen, ist eine Repräsentierung möglich:

  $$\Phi \vdash \bigwedge x(\varphi(x) \leftrightarrow x \equiv \boldsymbol{n}) \quad \text{gdw.} \quad \mathcal{N} \models def(\ulcorner\varphi(x)\urcorner, \boldsymbol{n})$$

$w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner))$ und $\chi_2 := \neg w(\alpha(\ulcorner w(\alpha(x))\urcorner))$ Sätze der gesuchten Art, denn es folgt

$$\mathcal{A}_{\mathrm{Lü}} \models w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner)) \leftrightarrow w(\ulcorner\neg w(\alpha(\ulcorner w(\alpha(x))\urcorner))\urcorner) \text{ und}$$
$$\mathcal{A}_{\mathrm{Lü}} \models \neg w(\alpha(\ulcorner w(\alpha(x))\urcorner)) \leftrightarrow \neg w(\ulcorner w(\alpha(\ulcorner\neg w(\alpha(x))\urcorner))\urcorner).$$

[54] Vgl. Boolos (1989).
[55] Vgl. Serény (2004).
[56] Vgl. die Darstellung der Antinomie von Berry auf S. 27.

für alle $S_{\mathrm{Ar}[v_0]}$-Formeln $\varphi(x)$ und alle $n \in \mathbb{N}$. Verfolgt man die Anintomie weiter, so erhält man

$$\begin{aligned}\mathcal{N} \models \neg \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y, \boldsymbol{m}))\\ \wedge \bigwedge z(z < \boldsymbol{m} \rightarrow \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y, z))).\end{aligned}$$

Anstatt einer Antinomie ergibt sich jedoch

$$\begin{aligned}\Phi \not\vdash \neg \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y, \boldsymbol{m}))\\ \wedge \bigwedge z(z < \boldsymbol{m} \rightarrow \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y, z))).\end{aligned}$$

Auf diese Weise gelangt man also zu einer semantischen Form des Unvollständigkeitssatzes.

- Ist PA eine aufzählbare Menge von S_{Ar}-Sätzen, die die Menge PA enthält, so erlaubt PA Repräsentierungen und es folgt analog zu der obigen semantischen Form

$$PA \not\vdash \neg \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y, \boldsymbol{m})).$$

Ist PA zudem ω-widerspruchsfrei, so gilt analog zu der obigen semantischen Form

$$PA \not\vdash \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y, \boldsymbol{m})).$$

Dies stellt eine syntaktische Form des Unvollständigkeitssatzes dar.

- Ist PA wie oben und ist zusätzlich $PA^{\vdash}$ entscheidbar, so gilt im Widerspruch zu Obigem

$$PA \vdash \neg \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y, \boldsymbol{m})),$$

da PA Repräsentierungen erlaubt. $PA^{\vdash}$ ist also unentscheidbar.

Die Antinomie von Berry lässt sich nicht als Spezialfall eines abstrakten Satzes aus Abschnitt 2.1 und ebenfalls auch nicht als Diagonalargument darstellen. Allerdings ist mit der Voraussetzung einer $S_{\mathrm{Be}[v_0,v_1]}$-Formel $def(x, y)$, so dass für alle $S_{\mathrm{Be}[v_0]}$-Formeln $\varphi(x)$ und alle $n \in \mathbb{N}$:

$$\mathcal{A}_{\mathrm{Be}} \models \bigwedge x(\varphi(x) \leftrightarrow x \equiv \boldsymbol{n}) \quad \text{gdw.} \quad \mathcal{A}_{\mathrm{Be}} \models def(\ulcorner \varphi(x) \urcorner, \boldsymbol{n}),$$

eine der Wahrheitskonvention entsprechende Voraussetzung zu machen. Diese wird nur nicht in Form eines Diagonalarguments genutzt. Mit Hilfe von $def(x.y)$ wird die den Widerspruch erzeugende Definition

$$\begin{aligned}\psi(x) :=& nat(x) \wedge \neg \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y,x)) \\ & \wedge \bigwedge z(nat(z) \wedge z < x \rightarrow \bigvee y(fm(y) \wedge l(y) < \mathbf{20} \wedge def(y,z)))\end{aligned}$$

gebildet. In ihren Voraussetzungen bezüglich der Ausdrucksfähigkeit der Sprache und bezüglich der Logik der Sprache ähnelt die Antinomie von Berry damit stark den anderen semantischen Antinomien.

3.8 Lösungen für Antinomien

Abschnitt 3.4.2 zeigt mögliche Ansatzpunkte für Lösungen von Antinomien, indem die zur Ableitung der jeweiligen Antinomien nötigen Voraussetzungen vier Aspekten von Sprache in einem sehr allgemeinen Sinn zugeordnet werden. In jedem dieser vier Bereiche können durch Einschränkung der Voraussetzungen Lösungsversuche ansetzen. Diese Bereiche sind Logik, Semantik, Ontologie (für ontologische Antinomien) und Ausdrucksstärke der Sprache (für semantische Antinomien). Logik und Semantik werden hier innerhalb eines Komplexes behandelt, da Ansätze in dieser Richtung meist Änderungen in beiden Bereichen beinhalten. Eine Semantik beinhaltet auch die Interpretation der logischen Zeichen, die Logik wird häufig durch eine Semantik festgelegt.[57] Die Aufteilung der Lösungsansätze in diese vier sprachlichen Bereiche entspricht der Aufteilung von Bromand, der sie für Erklärungen des Zustandekommens von Antinomien – sogenannte Diagnosen – verwendet. Die zum Bereich der sprachlichen Ausdrucksstärke gehörende Diagnose wird von Bromand als epistemologische Diagnose bezeichnet. Darunter versteht er jedoch in erster Linie auch eine expressive Unvollständigkeit. Bromand behandelt Logik und Semantik als getrennte Ansätze, betont aber auch deren enge Beziehung.[58]

[57] Bei einer Änderung der klassischen Semantik zu einer dreiwertigen Semantik kann allerdings die Logik – und die Semantik – für alle Sätze mit den herkömmlichen Wahrheitswerten "wahr" und "falsch" gleichbleiben. Dies ist bei Verwendung der Kleeneschen Regeln der Fall. In einem gewissen Sinn ändert sich in diesem beispielhaften Fall also die Semantik als Ganzes, nicht aber die Logik. Vgl. S. 107.

[58] Vgl. Bromand (2001). Bromand beleuchtet in seiner Arbeit hauptsächlich die Bedeutung, die Konsequenzen und den philosophiehistorischen Ursprung der Diagnose der expressiven

Im Bereich der semantischen Antinomien gehören zu den neueren, besonders häufig diskutierten Ansätzen die Lösungen von Martin und Woodruff, Kripke, Herzberger, Gupta und Belnap, Barwise und Etchemendy sowie Priest.[59] Alle diese Ansätze sind dem logisch-semantischen Bereich zuzuordnen.[60] Sie unterscheiden sich in der Art der Veränderungen von Logik und Semantik im Vergleich zur klassischen Logik und Semantik und in der Menge der als paradox einzustufenden Sätze, die keinen Wahrheitswert in klassischem Sinn erhalten. Die Lösung von Priest baut auf einer parakonsistenten Logik auf. Sowohl Martin und Woodruff als auch Kripke, deren Ansätze sich sehr ähneln, verwenden ein dreiwertiges System, das neben den klassischen Wahrheitswerten "wahr" und "falsch" eine Wahrheitswertlücke "weder wahr noch falsch" enthält. Durch wiederholte semantische Bewertung immer komplexerer, das Wahrheitsprädikat enthaltender Sätze werden Extension und Antiextension des Wahrheitsprädikats monoton *angefüllt*. Ausschlaggebend für die letztendliche semantische Bewertung ist das Verhalten der Sätze in Fixpunkten dieses Prozesses. Vom Lügner-Satz wird durch Widerspruch gezeigt, dass er in keinem Fixpunkt in Extension oder Antiextension des Wahrheitsprädikats liegen kann. Er fälllt damit in eine Wahrheitswertlücke. Bei den Ansätzen von Herzberger sowie Gupta und Belnap geht es weniger um die Bestimmung einer adäquaten Extension und Antiextension des Wahrheitsprädikats als um die Modellierung des Verhaltens von Sätzen wie dem Lügner-Satz. Es wird dabei ein Bewertungsprozess dargestellt, in dem der Wahrheitswert des Lügner-Satzes zwischen "wahr" und "falsch" oszilliert. Da die Wahrheitswerte "wahr" und "falsch" jeweils auf verschiedenen Stufen des Bewertungsprozesses angenommen werden, wird ein Widerspruch vermieden. Auch in dem situationssemantischen Ansatz von Barwise und Etchemendy

Unvollständigkeit und stärkt sie als Erklärung für das Zustandekommen der semantischen Antinomien.

[59] Vgl. Martin u. Woodruff (1975), Kripke (1975), Herzberger (1982), Gupta u. Belnap (1993), Barwise u. Etchemendy (1987), Priest (1979), Priest (1984). In Brendel (1992) werden diese Ansätze innerhalb eines einheitlichen Typs von formaler Sprache mit einer an den jeweiligen Lösungsvorschlag angepassten Semantik dargestellt und besprochen. Insbesondere wird das Verhalten des Lügner-Satzes, des Wahrsagers – "Dieser Satz ist wahr" – und des verstärkten Lügner-Satzes innerhalb der einzelnen Lösungen untersucht. Auch in Kühnberger (2002) werden diese Ansätze, mit Ausnahme des Ansatzes von Priest, ausführlich untersucht. Hier werden unter anderem die formalen Methoden, die jeweils verwendet werden, um zirkuläre Phänomene zu modellieren, detailliert dargestellt und weiterentwickelt.

[60] Der Ansatz von Barwise und Etchemendy kann allerdings auch berechtigterweise dem Bereich der Ausdrucksstärke der Sprache zugeordnet werden. Vgl. die Ausführungen in diesem Absatz und ferner Priest (2002), S. 152 und Bromand (2001), S. 92.

geht es um die Modellierung des Verhaltens von Sätzen wie dem Lügner-Satz. Der Oszillationsprozess, der sich in der Reflexion über den Lügner-Satz ergibt, wird hier durch die Einführung eines kontextsensitiven Situationsparamenters antinomiefrei aufgelöst. Der Situationsparameter ermöglicht den Bezug eines Satzes auf unterschiedliche Situationen. In diesem Ansatz erhält der Lügner-Satz den Wahrheitswert "falsch".

Der erste Unvollständigkeitssatz von Gödel kann ebenfalls als Diagonalargument aufgefasst werden. Er wurde implizit in Abschnitt 3.3 als Diagonalargument, angewendet auf eine Diagonalstruktur$_1$, beschrieben, da er unter den abstrakten Satz 2.3 a) fällt. Es kommt jedoch zu keiner Antinomie, da sich grob gesagt die objektsprachliche Negation bezüglich der Ableitbarkeitsrelation und der metasprachlichen Negation nicht klassisch verhält. Es gilt also nicht für alle Sätze φ: $PA \nvdash \varphi$ gdw. $PA \vdash \neg\varphi$. Der erste Unvollständigkeitssatz kann damit auch als ein Lösungsansatz in einer dreiwertigen Semantik für die Lügner-Antinomie betrachtet werden.

Für ontologische Antinomien sind entsprechende Lösungsansätze weniger ausgearbeitet, jedoch durchaus denkbar. Maddy überträgt beispielsweise in ihrem Ansatz den Weg Kripkes zur Bestimmung von Extension und Antiextension des Wahrheitsprädikats auf Klassen. Dabei werden Klassen entsprechende Extensionen und Antiextensionen so zugeordnet, dass Lücken möglich sind. Es gibt also Klassen, die weder in die Extension noch in die Antiextension einer bestimmten Klasse fallen.[61] Auch das Problem des verstärkten Lügners[62] überträgt sich dann allerdings auf den ontologischen Fall. Die verstärkte Lügner-Antinomie für die Antinomie von Russell soll an dieser Stelle kurz angedeutet werden. Der Metasprache und dem metasprachlichen Wahrheitsbegriff entsprechen im ontologischen Fall eine Metaklassentheorie bzw. eine Metaklasse m aus Paaren von Klassen, die auf der Metaebene die Elementbeziehung der Klassen ausdrückt. Zur genauen Festlegung dieser Metaklasse ist wie beim metasprachlichen Wahrheitsbegriff die Interpretation der logischen Verknüpfungen, wie z. B. des Komplements einer Klasse, nötig. $(x_1, x_2) \in m$ für Klassen x_1, x_2 im ontologischen Fall entspricht dann grob ausgedrückt $\mathcal{A} \models \varphi(t)$ im semantischen Fall. Um analog zur Lügner-Antinomie vorgehen zu können, wird ei-

[61] Vgl. Maddy (1983).

[62] Die verstärkte Lügner-Antinomie ist eine Variante der einfachen Lügner-Antinomie innerhalb einer dreiwertigen Semantik, die sich auf die metasprachliche Negation bezieht. Sie wird auf S. 21 geschildert.

ne Variante der Antinomie von Russell betrachtet, die von der Klasse aller Klassen, die sich selbst enthalten, ausgeht.[63] Es ergibt sich damit als Äquivalenz: $\{x|x \notin x\} \in \{x|x \notin x\}$ gdw. $\{x|x \notin x\} \in \{x|x \in x\}$. Übertragen auf die Metaklassen-Sprechweise – analog zur metasprachlichen Darstellung der Lügner-Antinomie, genauer der Antinomie von Grelling in der autologisch-Form – erhält man: $(\{x|x \notin z\}, \{x|x \notin z\}) \in m$ gdw. $(\{x|x \notin z\}, z) \in m$. Hierbei sei z eine Klasse, die die Selbsteinsetzung in die Metaklasse m in der Klassentheorie analog zu der einstelligen Formel *aut* repräsentiert: für alle Klassen x gelte: $(x, x) \in m$ gdw. $(x, z) \in m$. Die Äquivalenz $(\{x|x \notin z\}, \{x|x \notin z\}) \in m$ gdw. $(\{x|x \notin z\}, z) \in m$ stellt nun keinen Widerspruch mehr dar, wenn die Metaklassentheorie dreiwertig gestaltet wird. Allerdings tritt der Widerspruch erneut auf, sobald die metasprachliche Komplementbildung objektsprachlich ausdrückbar wird: sobald es also eine Klasse z^k gibt, so dass $(x, z) \notin m$ gdw. $(x, z^k) \in m$.

Übliche Systeme der Mathematik, die verwendet werden, um ontologische Antinomien zu umgehen, sind Mengenlehren wie ZF oder Mengen-Klassen-Theorien wie NBG. In reinen Mengenlehren wird das Abstraktionsschema durch andere Mengenbildungsaxiome wie das Paarmengen- oder das Potenzmengenaxiom und ein Aussonderungsschema, das die Abstraktion *aus einer Menge* erlaubt, ersetzt. In Mengen-Klassen-Theorien gibt es zusätzlich ein Abstraktionsschema, durch das beliebige durch eine Formel beschreibbare[64] Zusammenfassungen von *Mengen* als *Klassen* gebildet werden können. Die Bildung von beliebigen Klassen durch Quantifikation über alle Klassen ist jedoch auch in dieser Theorie nicht möglich. Mengenlehren und Mengen-Klassen-Theorien gehören damit zu den ontologischen Ansätzen. Bromand ordnet die Mengen-Klassen-Theorien dem Ansatz der sprachlichen Ausdrucksstärke zu, da beispielsweise die Zusammenfassung aller Mengen in diesem System zwar nicht als Menge, aber doch als Klasse, also als Objekt existiert. Er argumentiert, dass es sich damit nicht um eine ontologische Beschränkung, sondern um eine Beschränkung der Ausdrucksstärke handelt. Selbst wenn man in diesem Fall eine Beschränkung der Ausdrucksstärke sieht, was nicht nahe liegend scheint, so bleibt doch die ontologische Beschränkung bestehen, dass es keine Klasse aller Klassen in diesen Theorien gibt.

[63] Vgl. S. 104.

[64] Das “beliebige” wird allerdings jeweils durch die zugelassene Art von Formeln eingeschränkt. Die grobe Unterscheidung ist, ob nur prädikative oder auch imprädikative Begriffsbildungen zugelassen sind.

Als Beispiel für ontologische Ansätze zu semantischen Antinomien nennt Bromand die Auffassung, dass es keine Menge gibt, die genau die wahren Sätze enthält. Eine solche Auffassung wird von Grim als Diagnose für das Zustandekommen von semantischen Antinomien vertreten.[65] Beispiele für Ansätze im Bereich der Ausdrucksstärke der Sprache für semantische Antinomien sind die axiomatischen Systeme von Feferman.[66] In seinem System KF überträgt Feferman den semantischen Ansatz von Kripke auf ein klassisches Logiksystem. Der Wahrheitsbegriff ist dabei partiell definiert.

Russells (einfache) Typentheorie und die Sprachstufentheorie Tarskis werden von Bromand den semantischen Ansätzen zugeordnet. Genauer gesagt fasst er sie als syntaktische Ansätze und Vorläufer der semantischen auf.[67] Nach Russells Typentheorie werden nur noch die einer Stufung in der Syntax entsprechenden Sätze zugelassen. Die Sätze, die dieser Stufung nicht entsprechen, können als in eine Wahrheitswertlücke fallend gedeutet werden. Insofern besteht eine gewisse Berechtigung, Russells Typentheorie mit semantischen Ansätzen in Verbindung zu bringen. Durch die Stufung der Konstanten und Variablen wird jedoch auch die Quantifikation über Klassen im Abstraktionsschema eingeschränkt. Es wird nicht mehr über den gesamten Gegenstandsbereich quantifiziert, der durch die Stufung in der Syntax ebenfalls in verschiedene Stufen zerfällt, sondern nur noch über einen Gegenstandsbereich, der dem Typ der Quantifikationsvariable entspricht. Der Typ der Quantifikationsvariable muss eine wohlgebildete Formel entstehen lassen, falls dies überhaupt möglich ist. Eine Klasse $\{x|x \notin x\}$ kann beispielsweise für keinen Variablentyp von x gebildet werden, es wird quasi über einem leeren Bereich quantifiziert. Zur Bildung einer Klasse $\{x|x = x\}$ wird jeweils über einen Bereich von Klassen eines festen Typs quantifiziert. Insofern wird nach der Russellschen Typentheorie das Abstraktionsschema als Klassenbildungsprinzip eingeschränkt und damit die Ontologie geändert. In dieser Hinsicht ist diese Lösung also dem ontologischen Ansatz zuzurechnen.

Tarski selbst stellt seine Hierarchie als eine Stufung verschiedener Sprachen dar. Die Idee des Ansatzes lässt sich aber auch als Stufung innerhalb einer Sprache

[65] Vgl. Bromand (2001), Kap. 7, Grim (1991).

[66] Vgl. Feferman (1984) und Feferman (1991).

[67] Während Bromand diese Ansätze als nur noch historisch interessant ansieht, wird von Brendel die Sprachstufentheorie Tarskis gegen neuere Ansätze zu semantischen Antinomien verteidigt. Das Hauptargument ist dabei die verstärkte Lügner-Antinomie, die letztlich eine Sprachstufung unumgehbar macht. Vgl. Brendel (1992) und Abschnitt 3.9 der vorliegenden Arbeit.

rekonstruieren.[68] Dabei werden syntaktisch Stufen eingeführt bezüglich des Vorkommens von einstelligen Relationskonstanten w_1, w_2, w_3, ..., die die Funktion eingeschränkter Wahrheitsbegriffe erfüllen. Jedem Term und jeder Formel ist dadurch eine bestimmte Stufe zugewiesen, und ein wohlgebildeter Satz aus einem Term und einer einstelligen Formel kommt nur zustande, wenn die Stufe des Terms kleiner ist als die Stufe der einstelligen Formel. Wie im Fall der Typentheorie Russells kann die Menge der nicht wohlgebildeten Sätze als Wahrheitswertlücke aufgefasst werden, die nach rein syntaktischen Kriterien bestimmt ist. In dieser Hinsicht rückt der Ansatz in die Nähe semantischer Ansätze. Wahrheit wird nach dem Ansatz von Tarski objektsprachlich durch die Folge w_1, w_2, w_3, ... ausgedrückt. Die Wahrheitskonvention gilt für jede dieser einstelligen Relationskonstanten w_i nur bezogen auf Sätze mit einer Stufe kleiner als i. Insofern wird das Wahrheitsschema eingeschränkt und damit die Ausdrucksstärke der Sprache verändert. In dieser Hinsicht ist der Ansatz in den Bereich der Ausdrucksstärke der Sprache einzuordnen. Im Fall des Wahrheitsbegriffs ist allerdings schon der Lügner-Satz selbst nicht formulierbar, denn der einstellige Term $\alpha(x)$, der die Selbsteinsetzung repräsentiert, kann nicht mehr beliebig angewendet werden – so z. B. nicht auf $\ulcorner w(\alpha(x)) \urcorner$.

Generell kann zu Lösungen von ontologischen und semantischen Antinomien gesagt werden, dass sich im semantischen Fall insbesondere die Frage stellt, wie ein alternativer Wahrheitsbegriff aussehen kann oder wie Extension und Antiextension des Wahrheitsprädikats in Bezug auf eine alternative Semantik gebildet werden können, und im ontologischen Fall die Frage, in welcher Form überhaupt Zusammenfassungen als Objekte gebildet werden können. Die besondere Rolle des Wahrheitsbegriffs unter den semantischen Begriffen wird auch auf S. 105 geschildert.

Eine im Zusammenhang mit Lösungen von ontologischen und semantischen Antinomien häufig diskutierte Frage ist die Frage nach der Einheitlichkeit der Lösungen.[69] Hierzu gehört die Frage, wann Lösungen, insbesondere für verschiedene Typen von Antinomien, einheitlich sind, und die Frage, ob es möglich und erstrebenswert ist, einheitliche Lösungen anzuwenden. Um von der Einheitlichkeit von Lösungen zu sprechen, muss zunächst eine Basis geschaffen werden, durch die festgelegt wird, in welcher Hinsicht sie einheitlich sein sollen. Als eine

[68] Vgl. Brendel (1992), Kap. 7.
[69] Vgl. z. B. Priest (2002), Bromand (2001).

solche Basis können die Darstellungen der Antinomien als Diagonalargumente in den Abschnitten 3.2 und 3.3 dienen. Ein Beispiel für einheitliche Lösungsansätze im logisch-semantischen Bereich wird auf S. 129 durch die Einführung einer Metaklasse parallel zu einem metasprachlichen Wahrheitsbegriff und der Verwendung einer dreiwertigen Semantik in beiden angedeutet. Ferner entsprechen Ansätze im ontologischen Bereich für ontologische Antinomien auf dieser Basis Ansätzen im Bereich der Ausdrucksstärke der Sprache für semantische Antinomien. Diese Parallele gilt auch für Antinomien wie den Lügner-Zirkel und die in Abschnitt 3.6 geschilderten Antinomien, die nicht über ein Diagonalargument im eigentlichen Sinn zustande kommen. Eine weitere Parallele kann auf einer anderen Ebene zwischen Lösungen für semantische Antinomien gebildet werden, indem die Menge der Sätze ohne klassischen Wahrheitswert in einer semantischen Lösung der Menge der Sätze, auf die eine veränderte Wahrheitskonvention nicht angewendet werden kann, gegenübergestellt wird.

In Bezug auf einheitliche Lösungen für ontologische und semantische Antinomien muss die unterschiedliche Situation der beiden Fälle beachtet werden. Im ontologischen Fall geht es um den Entwurf einer Mengen- oder Klassentheorie, in der mathematische Theorien formuliert und bewiesen werden sollen. Im semantischen Fall geht es um den Entwurf einer Sprache und ihrer Semantik, insbesondere eines Wahrheitsbegriffs für die Sprache, wobei auch bestimmte Intuitionen in Bezug auf die natürliche Sprache zu berücksichtigen sind.

Priest sieht in der Einheitlichkeit von Lösungen für Antinomien ein ausschlaggebendes Kriterium, diese überhaupt als sinnvoll zu erachten.[70] Natürlicherweise macht Priest sein eigenes Antinomien-Schema zum Kriterium für die Einheitlichkeit von Lösungen. Eine Übertragung der Lösungen für semantische Antinomien von Tarski, Kripke, Barwise und Etchemedy und der meisten anderen Lösungen auf ontologische Antinomien schließt Priest mit der einfachen Bemerkung aus, dass in ihnen kein Bezug auf Gesamtheiten, das Unendliche oder andere Objekte der ontologische Antinomien genommen wird.[71] Diese wenig ausgearbeitete Ablehnung der Übertragungsmöglichkeit dieser Lösungen überrascht, da Priest bei der Gegenüberstellung der verschiedenen Antinomien den strukturellen Aspekt besonders betont, im Fall der Lösungen die Parallele jedoch an dieser Stelle nicht auf die von ihm ausgearbeitete Struktur der Antinomien

[70]Vgl. Priest (2002), S. 166 f.
[71]Siehe Priest (2002), S. 167.

bezieht. Lösungen für ontologische Antinomien wie ZF und NBG werden von ihm anhand seines Schemas übertragen – in diesem Fall also nach strukturellen Aspekten. Diese Lösungen setzen nach seiner Darstellung an Punkt 1 seines Schemas an. Im Fall der Russellschen Antinomie sagt Punkt 1 des Schemas je nach Interpretation, dass eine Klasse aller Klassen bzw. eine Klasse aller Klassen, die sich nicht selbst enthalten, existiert. Priest überträgt dies nun auf die Lügner-Antinomie – als ein Beispiel für eine semantische Antinomie –, indem er dieser Existenz die Existenz der Menge der wahren Sätze gegenüberstellt. Da diese jedoch abzählbar ist, ist sie nicht in Frage zu stellen. Damit kommt er zu dem Schluss, dass ZF, NBG und ähnliche Theorien keine Entsprechung für semantische Antinomien haben.[72] Im Fall der Lügner-Antinomie besteht Punkt 1 des Schemas jedoch noch aus einer zweiten Bedingung: die Menge der wahren Sätze ist in der Sprache definierbar. Und dies ist die für die Lügner-Antinomie ausschlaggebende Bedingung. Eine Übertragung könnte also nach dem Schema von Priest auch an dieser Stelle ansetzen.

3.9 Antinomien als Grenzen des Aussagbaren

In Abschnitt 3.8 wird grob geschildert, wie in den drei Bereichen Logik und Semantik, Ontologie sowie Ausdrucksstärke der Sprache Lösungen für Antinomien ansetzen können. In diesem Abschnitt soll nun gezeigt werden, dass die hier besprochenen Antinomien immer, welchen Lösungsweg man auch einschlägt, Grenzen des Aussagbaren deutlich machen.

Geht man von einer ontologischen Diagnose für die ontologischen Antinomien aus, sieht man also die Ursache für das Zustandekommen der Antinomien im intuitiv überschätzten ontologischen Reichtum, der sich aus Sicht der Objektsprache ergibt, so ist in einem alternativ zu entwickelnden System das unbeschränkte Abstraktionsschema einzuschränken. Das Auftreten der Antinomie mit der Implikation, dass im antinomischen System jeder Satz wahr ist, kann also – wählt man diese Diagnose – als Hinweis auf eine Unvollständigkeit bezüglich der Klassenbildung gedeutet werden. Geht man parallel dazu von einer Diagnose bezüglich der Ausdrucksstärke der Sprache für die semantischen Antinomien aus, sieht man also die Ursache für das Zustandekommen in

[72] Siehe Priest (2002), S. 167.

der intuitiv überschätzten sprachlichen Ausdrucksfähigkeit der Objektsprache, so ist in einem alternativ zu entwickelnden System das Wahrheitsschema, die Wahrheitskonvention einzuschränken. Das Auftreten der Antinomie mit der Implikation, dass im antinomischen System jeder Satz wahr ist, kann also – wählt man diese Diagnose – als Hinweis auf eine Unvollständigkeit bezüglich der Ausdrückbarkeit semantischer Begriffe der Sprache in der Sprache selbst gedeutet werden.

Geht man primär von einer logisch-semantischen Diagnose aus, sieht man also die Ursache für das Zustandekommen der Antinomien in einem verkehrten Verständnis von Logik oder Semantik, so kann durch eine Änderung dieser in einem alternativen System das Abstraktionsschema – in einer übertragenen Form – erhalten bleiben bzw. der Wahrheitsbegriff in der Objektsprache repräsentiert werden. Diesbezüglich kann also von Vollständigkeit gesprochen werden. Die Unvollständigkeit tritt jedoch an einer anderen Stelle wieder zu Tage: Die metasprachliche Negation, der Nichtwahrheitsbegriff, das Komplement einer Klasse sind nicht ausdrückbar. Für die genannten logisch-semantischen Ansätze zu semantischen Antinomien wird diese Unvollständigkeit von Brendel jeweils in Form einer verstärkten Lügner-Antinomie deutlich gemacht.[73] Die verstärkte Lügner-Antinomie bezieht sich ja gerade auf die metasprachliche Negation bzw. den Nichtwahrheitsbegriff, weswegen durch sie gezeigt werden kann, dass sich nicht die gesamte Semantik der Objektsprache in der Objektsprache ausdrücken lässt. Dieselbe Form von Unvollständigkeit wird auch im Gödelschen Unvollständigkeitssatz und im Satz von der Unentscheidbarkeit des Halteproblems deutlich.

Die Unvollständigkeit in ontologischer oder expressiver Hinsicht wird also durch keinen der Lösungsansätze umgangen. Insofern machen die Antinomien Grenzen des Aussagbaren deutlich. Ausschlaggebend für das Zustandekommen der Antinomie – oder der verstärkten Antinomie – ist jeweils die Existenz einer bestimmten Klasse, im Fall der ontologischen Antinomien, und die objektsprachliche Ausdrückbarkeit einer bestimmten Eigenschaft von Formeln, im Fall der semantischen Antinomien. Zur Bildung dieser Objekte – der Klasse bzw. der Formel – ist zum einen die Quantifikation über den gesamten Objektbereich, also über alle Klassen bzw. über alle Formeln, relevant[74] und zum anderen

[73]Vgl. Brendel (1992).

[74]Dies darf allerdings nicht so verstanden werden, dass eine Einschränkung des Quantifika-

der Bezug von Klassen als Objekten zu Klassen als Zusammenfassungen bzw. von einstelligen Formeln als Objekten zu einstelligen Formeln als Zusammenfassungen. Ferner ist eine Negation mit bestimmten klassischen Eigenschaften relevant. Wird die Negation im logischen System nicht klassisch verwendet, so kommt über die Anforderung, den metasprachlichen Nichtwahrheitsbegriff oder das metatheoretische Komplement objektsprachlich repräsentieren zu können, eine solche zur Hintertür herein.

Diese Möglichkeiten zusammengenommen ergeben Antinomien; die Grenzen, die durch die Antinomien angedeutet werden, liegen davor. Auch die in Abschnitt 3.6 geschilderten Antinomien, die nicht über ein Diagonalargument im eigentlichen Sinn zustande kommen, zeigen diese Grenzen auf, denn zu ihrer Bildung wird ebenfalls auf die oben geschilderten Möglichkeiten zurückgegriffen. Es muss also keinesfalls immer ein Diagonalargument im eigentlichen Sinn vorliegen – dies ist z. B. auch in der Antinomie von Berry nicht der Fall –, Diagonalisierung ist aber eine naheliegende Möglichkeit, ein Objekt so zu bestimmen, dass ein Widerspruch entsteht.

So unterschiedlich also die Art der Voraussetzungen in den einzelnen Antinomien ist, so unterschiedlich auch die Voraussetzungen in ihrer Funktion innerhalb des Diagonalarguments und der Diagonalstruktur sind, gemeinsam ist diesen Antinomien, dass ein Bezug auf die Gesamtheit als Umgrenzung des Systems als Ganzem zur Bildung von Objekten ermöglicht wird.[75] Bei der Gesamtheit kann es sich, wie im Fall der Klassentheorie, um die Gesamtheit aller Klassen handeln oder, wie im Fall der Ordinalzahltheorie, um die Gesamtheit aller Ordinalzahlen oder um die Gesamtheit aller einstelligen Begriffe. Der Bezug auf die Gesamtheit, der in Form einer Quantifikation über die Gesamtheit erfolgt, sieht ebenfalls unterschiedlich aus. In der Antinomie von Russell wird direkt

tionsbereichs automatisch eine Antinomie vermeidet. Vgl. Abschnitt 3.6.

[75] Vgl. hierzu auch Priest (2002). Für Priest stellen Umgrenzungen oder Umhüllungen, auf die man sich bei der Bildung einer Antinomie so bezieht, dass ein Widerspruch entsteht, die relevante strukturelle Gemeinsamkeit der Antinomien dar. In Zusammenhang mit der Antinomie von Burali-Forti beschreibt er dies wie folgt: "... the limit is defined "from below"; but the contradiction is produced by considering it "from above": that is, in each case we take the limit to be itself a unity and note its properties." (Priest (2002), S. 120.)

Priest wählt jedoch in einigen Antinomien etwas Anderes als Grenze, als es in der vorliegenden Arbeit geschieht. Er untersucht außerdem nicht weiter, in welcher Form jeweils auf die Grenze Bezug genommen wird, was unablässlich ist, will man die relevanten Parallelen und Unterschiede zwischen ontologischen und semantischen Antinomien erkennen. Eine Auseinandersetzung mit der Arbeit von Priest erfolgt in Abschnitt 3.5.

die Bildung einer problematischen Klasse, die aus der Gesamtheit aller Klassen abstrahiert wird, zugelassen. In der Antinomie von Cantor erfolgt dies über die Existenz der Gesamtheit aller Klassen als Klasse. In den semantischen Antinomien erfolgt die Quantifikation metasprachlich. Es zeigt sich aber auch, dass der Bezug auf die Gesamtheit allein nicht zu einer Antinomie führt, dass zusätzlich die Abbildung von Objekten als Elementen auf Objekte als Zusammenfassungen in irgendeiner Form ausdrückbar sein muss. Ferner muss eine klassisch verwendete Negation vorhanden sein oder müssen entsprechende metasprachliche Begriffe oder Klassen wie Nichtwahrheit oder das Komplement einer Klasse repräsentierbar sein. Antinomien bringen so ein allgemeines Unbehagen auf den Punkt, das man dabei empfindet, wenn man sich innerhalb eines Ganzen auf das Ganze bezieht – wie auch immer dies genau aussieht. Sie stellen Fälle dar, in denen dieses Unbehagen gerechtfertigt ist.

An dieser Stelle lässt sich auch der gesteckte Rahmen der vorliegenden Arbeit über ontologische und semantische Antinomien hinaus erweitern. Es können z. B. Kants Antinomien unter diesem Gesichtspunkt untersucht werden.[76] Auch hier wird eine Umgrenzung zum Objekt innerhalb der Umgrenzung gemacht, indem z. B. – innerhalb der Betrachtung von Gegenständen in Raum und Zeit – Raum und Zeit selbst als Gegenstände angesehen und ihre Eigenschaften untersucht werden.[77]

[76] Vgl. Kant (1990), “Die transzendentale Dialektik”. Essler nennt diese Antinomien metaphysische Antinomien, da sie sich auf metaphysische Annahmen über die Natur beziehen. Vgl. Essler (1972), S. 129, Fußnote 23.

[77] Zu Kants Antinomien unter diesem Gesichtspunkt vgl. Priest (2002) und die Bemerkungen zu Priest auf S. 114 der vorliegenden Arbeit. Eine frühe Bemerkung zu einer Parallele zwischen ontologischen Antinomien und den Antinomien von Kant findet sich bei Hessenberg: die Antinomie von Burali-Forti “... erinnert an diejenigen Antinomien, die nach Kant entstehen, wenn wir die Natur als ein abgeschlossenes Ganzes betrachten.” (Hessenberg (1906), S. 633.) In dieser Parallele spielt allerdings der in der vorliegenden Arbeit irrelevante Begriff der Abgeschlossenheit eine wichtige Rolle.

Literaturverzeichnis

[Barwise u. Etchemendy 1987] Barwise, Jon ; Etchemendy, John: *The Liar. An Essay on Truth and Circularity.* New York, Oxford, 1987

[Berka u. Kreiser 1983] Berka, Karel ; Kreiser, Lothar: *Logik-Texte.* Berlin, 1983

[Boolos 1989] Boolos, Georg S.: A New Proof of the Gödel Incompleteness Theorem. In: *Notices Amer. Math. Soc.* 36 (1989), S. 388–390

[Boolos u. Jeffrey 2002] Boolos, George S. ; Jeffrey, Richard C.: *Computability and Logic.* 4. Cambridge, 2002. – Erste Auflage: 1974

[Brendel 1992] Brendel, Elke: *Die Wahrheit über den Lügner: eine philosophisch-logische Analyse der Antinomie des Lügners.* Berlin, New York, 1992

[Bromand 2001] Bromand, Joachim: *Philosophie der semantischen Pardoxien.* Paderborn, 2001

[Burali-Forti 1897] Burali-Forti, Cesare: Una questione sui numeri transfiniti. In: *Rendiconti der Circolo matematico di Palermo* 11 (1897), S. 154–164. – Wiederabgedruckt in van Heijenoort (1967), S. 104–111.

[Cantor 1890] Cantor, Georg: Über eine elementare Frage der Mannigfaltigkeitslehre. In: *Jahresber. der Dt. Math.-Verein.* 1 (1890/91), S. 75–78

[Carnap 1934] Carnap, Rudolph: *Logische Syntax der Sprache.* Wien, 1934

[Champlin 1988] Champlin, T. S.: *Reflexive Paradoxes.* London, 1988

[Church 1936] Church, Alonzo: An Unsolvable Problem of Elementary Number Theory. In: *American Journal of Mathematics* 58 (1936), Nr. 2, S. 345–363

[Curry 1942] Curry, Haskell B.: The Inconsistency of Certain Formal Logics. In: *The Journal of Symbolic Logic* 7 (1942), S. 115–117

[Ebbinghaus u. a. 1992] Ebbinghaus, Heinz-Dieter ; Flum, Jörg ; Thomas, Wolfgang: *Einführung in die mathematische Logik.* Mannheim, Leipzig, Wien, Zürich, 1992

[Ehrenfeucht u. Feferman 1960] Ehrenfeucht, Andrzej ; Feferman, Solomon: Representability of Recursively Enumerable Sets in Formal Theories. In: *Arch. Math. Logik Grundlagenforsch.* 5 (1960), S. 37–41

[Essler 1964] Essler, Wilhelm K.: *Aufzählbarkeit und Catorsches Diagonalverfahren*, Ludwig-Maximilians-Universität München, Diss., 1964

[Essler 1972] Essler, Wilhelm K.: *Analytische Philosophie I.* Stuttgart, 1972

[Essler u. Brendel 1993] Essler, Wilhelm K. ; Brendel, Elke: *Grundzüge der Logik II.* Frankfurt am Main, 1993

[Essler u. a. 2001] Essler, Wilhelm K. ; Martínez, Rosa F. ; Labude, Joachim: *Grundzüge der Logik I.* Frankfurt am Main, 2001

[Feferman 1984] Feferman, Solomon: Toward Useful Type-Free Theories. I. In: *The Journal of Symbolic Logic* 49 (1984), Nr. 1, S. 75–111. – Wiederabgedruckt in Martin (1984).

[Feferman 1991] Feferman, Solomon: Reflecting on Incompleteness. In: *The Journal of Symbolic Logic* 56 (1991), Nr. 1, S. 1–49

[Feferman u. a. 1986] Feferman, Solomon (Hrsg.) u. a.: *Gödel, Kurt. Collected Works.* Bd. 1. New York, 1986

[Goddard u. Johnston 1983] Goddard, Leonard ; Johnston, Mark: The Nature of Reflexive Paradoxes: Part I. In: *Notre Dame Journal of Formal Logic* 24 (1983), Nr. 4, S. 491–508

[Gödel 1931] Gödel, Kurt: Über formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme I. In: *Monatshefte für Mathematik und Physik* 38 (1931), S. 173–198. – Wiederabgedruckt in Feferman u. a. (1986).

[Grim 1991] Grim, Patrick: *The Incomplete Universe.* Cambridge (MA), 1991

[Gupta u. Belnap 1993] Gupta, Anil ; Belnap, Nuel: *The Revision Theory of Truth.* Cambridge (MA), 1993

[van Heijenoort 1967] Heijenoort, Jean van (Hrsg.): *From Frege to Gödel. A Source Book in Mathematical Logic, 1879-1931.* Cambridge (MA), London, 1967

[Herzberger 1970] Herzberger, Hans G.: Paradoxes of Grounding in Semantics. In: *The Journal of Philosophy* 67 (1970), Nr. 6, S. 145–167

[Herzberger 1982] Herzberger, Hans G.: Notes on Naive Semantics. In: *Journal of Philosophical Logic* 11 (1982), S. 61–102. – Wiederabgedruckt in Martin (1984).

[Hessenberg 1906] Hessenberg, Gerhard: Grundbegriffe der Mengenlehre. In: *Abhandlungen der Fries'schen Schule, Neue Folge* I (1906), Nr. 4, S. 479–706

[Hilbert u. Bernays 1939] Hilbert, David ; Bernays, Paul: *Grundlagen der Mathematik*. Berlin, 1939

[Hintikka 1957] Hintikka, K. Jaakko J.: Vicious Circle Principle and the Paradoxes. In: *The Journal of Symbolic Logic* 22 (1957), Nr. 3, S. 245–249

[Kahle 2007] Kahle, Reinhard: Die Gödelschen Unvollständigkeitssätze. In: *Mathematische Semesterberichte* 54 (2007), S. 1–12

[Kant 1990] Kant, Immanuel: *Kritik der reinen Vernunft*. Hamburg, 1990

[Kleene 1986] Kleene, Stephen C.: Introductory Note to 1930b, 1931 and 1932b. In: **(Feferman u. a., 1986)**, S. 126–141

[König 1905] König: *On the Foundations of Set Theory and the Continuum Problem*. 1905. – Wiederabgedruckt in van Heijenoort (1967), S. 145–149.

[Kripke 1975] Kripke, Saul A.: Outline of a Theory of Truth. In: *The Journal of Philosophy* 72 (1975), Nr. 19, S. 690–716. – Wiederabgedruckt in Martin (1984).

[Kühnberger 2002] Kühnberger, Kai-Uwe: *Formal Frameworks for Circular Phenomena: Possibilities of Modeling Pathological Expressions in Formal and Natural Languages*, Universität Tübingen, Diss., 2002. – http://w210.ub.uni-tuebingen.de/dbt/volltexte/2002/477/

[Maddy 1983] Maddy, Penelope: Proper Classes. In: *Journal of Symbolic Logic* 48 (1983), Nr. 1, S. 113–139

[Martin 1984] Martin, Robert L. (Hrsg.): *Recent Essays on Truth and the Liar Paradox*. Oxford, New York, 1984

[Martin u. Woodruff 1975] Martin, Robert L. ; Woodruff, Peter W.: On Representing "True-in-L" in L. In: *Philosophia* 5 (1975), S. 213–217. – Wiederabgedruckt in Martin (1984).

[Mostowski 1964] Mostowski, Andrzej: *Sentences Undecidable in Formalized Arithmetic*. Amsterdam, 1964

[Nelson 1974] Nelson, Leonard ; Bernays (Hrsg.) u. a.: *Leonard Nelson: Gesammelte Schriften in neun Bänden: Bd. 3. Die kritische Methode in ihrer Bedeutung für die Wissenschaft*. Hamburg, 1974

[Nelson u. Grelling 1908] Nelson, Leonard ; Grelling, Kurt: Bemerkungen zu den Paradoxien von Russell und Burali-Forti. In: *Abhandlungen der Fries'schen Schule, Neue Folge* II (1908), Nr. 3, S. 301–334. – Zitiert nach Nelson (1974).

[Peano 1906] Peano, Giuseppe: Additione. In: *Rivista di Mathematica* 8 (1906), S. 143–157

[Priest 1979] Priest, Graham: The Logic of Paradox. In: *Journal of Philosophical Logic* 8 (1979), S. 219–241

[Priest 1984] Priest, Graham: The Logic of Paradox Revisited. In: *Journal of Philosophical Logic* 13 (1984), S. 152–180

[Priest 2002] Priest, Graham: *Beyond the Limits of Thought*. Oxford, 2002

[Ramsey 1925] Ramsey, Frank P.: The Foundations of Mathematics. In: *Proceedings of the London Mathematical Society* Bd. 25, 1925 (2), S. 338–384. – Wiederabgedruckt in Ramsey (1931), S. 1–61.

[Ramsey 1931] Ramsey, Frank P. ; Braithwaite, Richard B. (Hrsg.): *The Foundations of Mathematics and Other Logical Essays*. London, New York, 1931. – RP New York, 1950

[Richard 1905] Richard, Jules: Les principes des mathématiques et le problème des ensembles. In: *Revue générale des sciences pures et appliquées* 16 (1905), S. 541. – Wiederabgedruckt in van Heijenoort (1967), S. 142–144.

[Ritchie u. Young 1969] Ritchie, Robert W. ; Young, Paul R.: Strong Representability of Partial Functions in Arithmetic Theories. In: *Information Sciences* 1 (1969), Nr. 2, S. 189–204

[Rosser 1936] Rosser, Barkley: Extensions of Some Theorems of Gödel and Church. In: *The Journal of Symbolic Logic* 1 (1936), Nr. 3, S. 87–91

[Rosser 1939] Rosser, Barkley: An Informal Exposition of Proofs of Gödel's Theorems and Church's Theorem. In: *The Journal of Symbolic Logic* 4 (1939), Nr. 2, S. 53–60

[Russell 1903] Russell, Bertrand: *The Principles of Mathematics*. 1. Cambridge, 1903. – (RP der 2. Aufl. London, 1972.)

[Russell 1905] Russell, Bertrand: On Some Difficulties in the Theory of Transfinite Numbers and Order Types. In: *Proceedings of the London Mathematical Society* Bd. 4, 1905 (2), S. 29–53

[Russell 1908] Russell, Bertrand: Mathematical Logic as Based on the Theory of Types. In: *American Journal of Mathematics* 30 (1908), S. 222–262. – Zitiert nach van Heijenoort (1967), S. 150–182.

[Sainsbury 2001] Sainsbury, Richard M.: *Paradoxien*. Stuttgart, 2001

[Serény 2003] Serény, György: Gödel, Tarski, Church, and the Liar. In: *The Bulletin of Symbolic Logic* 9 (2003), Nr. 1, S. 3–25

[Serény 2004] Serény, György: Boolos-Style Proofs of Limitative Theorems. In: *Mathematical Logic Quarterly* 50 (2004), Nr. 2, S. 211–216

[Smorynski 1978] Smorynski, Craig: The Incompleteness Theorems. 1978, S. 821–865

[Smullyan 1992] Smullyan, Raymond M.: *Gödel's Incompleteness Theorems*. New York, Oxford, 1992

[Stegmüller 1957] Stegmüller, Wolfgang: *Das Wahrheitsproblem und die Idee der Semantik*. Wien, New York, 1957

[Stegmüller 1973] Stegmüller, Wolfgang: *Unvollständigkeit und Unentscheidbarkeit*. 3. Wien, New York, 1973. – Erste Auflage: 1959

[Tarski 1935] Tarski, Alfred: Der Wahrheitsbegriff in den formalisierten Sprachen. In: *Studia Philosophica* (1935). – Wiederabgedruckt in Berka u. Kreiser (1983).

[Tarski u. a. 1953] Tarski, Alfred ; Mostowski, Andrzej ; Robinson, Raphael M.: *Undecidable Theories*. Amsterdam, 1953

Begriffs- und Symbolverzeichnis

Ich erkläre, dass ich die Dissertation selbständig verfasst und alle in Anspruch genommenen Hilfsmittel in der Dissertation angegeben habe.

Frankfurt, den 27.03.2008

Stefanie Ucsnay

Lebenslauf

Stefanie Maria Ucsnay
geb. 12.05.1971 in Bonn
Staatsangehörigkeit: deutsch

Schulbildung

08/81 – 05/90	**Gymnasium Bonn** Abitur

Studium/Promotion

10/90 – 09/96	**Universität Bonn** Diplom Mathematik
04/97 – 09/98	**Universität Mainz** 1. Staatsexamen Mathematik und Philosophie
10/98 – 09/00	**Universität Frankfurt, Institut für Philosophie** Magisterstudium Philosophie
10/00 – heute	**Universität Frankfurt, Institut für Philosophie** Doktorarbeit bei Prof. Wilhelm K. Essler

Berufstätigkeit

10/00 – 09/05	**Universität Frankfurt, Institut für Philosophie** Wissenschaftliche Mitarbeiterin
04/06 – 09/07	**R+V-Versicherung, Wiesbaden**
seit 02/08	**Studienseminar Frankfurt** Referendariat am Gymnasium

Frankfurt, den 27.03.2008