MIT META-MATHEMATIK GEGEN DEN REDUKTIONISMUS

Marcos Kiesekamp
marcoskiesekamp(at)gmail.com
20. Januar 2020

Zusammenfassung. Angesichts der Unvollständigkeitssätze Gödels gab Stephen Hawking seinen Reduktionismus auf. Auch John Lucas findet in ihnen Gründe, um reduktionistische Tendenzen in der Wissenschaft abzulehnen. Wir werden hier zeigen, dass das falsch ist. Wir zeigen aber weiter, dass andere Ergebnisse aus der Metamathematik sich durchaus dazu eignen. Im Einzelnen sind das:

(1) Eine Serie von Sätzen aus der Modelltheorie, die wir hier als *die Sätze Skolems* bezeichnen und die Gödel größtenteils selbst unabhängig von Skolem bewies.
(2) Eine Serie von Sätzen Tarskis.
(3) Die genannten Sätze von Skolem und Tarski zusammen.
(4) Die berühmten Sätze Turings.

Wir erhalten vier Ergebnisse aus der Meta-Mathematik, aus denen sich die Unmöglichkeit vier verschiedener Positionen ableiten lässt, die man als reduktionistisch bezeichnen könnte. Im Einzelnen sind das:

(1) Die Behauptung, wissenschaftliche Begriffe ließen sich in einem logischen System eindeutig definieren (*Reduzierbarkeit auf logische Systeme*).
(2) Die Behauptung, jede wissenschaftliche Erkenntnis ließe sich in einer universellen Sprache ausdrücken (*Reduzierbarkeit von Wissen auf sprachliche Aussagen*).
(3) Die Behauptungen, eine vollständige Theorie der Mikrowelt würde ausreichen, um die Phänomene der Makrowelt logisch herzuleiten, bzw. eine vollständige Theorie von Individuen würde ausreichen, um die Phänomene bei Gruppen von Individuen logisch herzuleiten (*Emergenzphänomene*).
(4) Die Behauptung, es wäre möglich, ein insofern vollständiges Geflecht aus Theorien zu bilden, dass es die Berechenbarkeit eines jeden Ereignisses im Universum ermöglicht (*Laplacescher Dämon*).

Wir werden auch darauf eingehen, welche Ergebnisse *nicht gegen* oder im Gegenteil *nicht für* reduktionistische Positionen einsetzbar sind, zum Beispiel, warum die Informationstheorie Kolmogorovs nicht für eine *algorithmic theory of everything* angeführt werden kann.

Inhaltsverzeichnis

0. Vorbemerkungen

Vorgehensweise

Im ersten Teil präsentieren wir einige Sätze aus der Metamathematik, die relevant für die Wissenschaftsphilosophie sind. Wir unterteilen in vier Kapiteln: Die Sätze Skolems, Tarskis, Turings und Gödels. Die Sätze von Turing und Gödel sind gut bekannt, aber entscheidend für unsere Argumente sind die Sätze von Skolem und Tarski. Da für sie in der Literatur nach unserem Wissen noch keine allgemeinverständliche Darstellung vorliegt, werden wir neben den Aussagen zumindest die wichtigsten Beweisideen anführen. Wir setzen Schulkenntnisse und rudimentäre Programmierkenntnisse voraus. Es geht uns nicht darum, die Beweise zu rekonstruieren, sondern zu vermitteln, welche Argumente entscheidend sind. Darauf bauend schließen wir jedes Kapitel mit einer philosophischen Interpretation der Sätze ab. Wer sich in der Modelltheorie auskennt, braucht natürlich nur die Interpretationen zu lesen. Eine problemorientierte Einleitung zu jedem Kapitel wird helfen zu erkennen, in welchem Kontext wir die Sätze interpretieren wollen.

In einem wissenschaftsphilosophischen Teil gehen wir auf die Voraussetzungen ein, die eine Wissenschaftstheorie erfüllen muss, damit metamathematische Ergebnisse auf sie anwendbar sind. Wir werden hier eine abgrenzende Definition von Wissenschaft und Mathematik benötigen. Davon ausgehend werden wir die Bedeutung der Sätze Turings und der Unvollständigkeitssätze Gödels für die Wissenschaftsphilosophie möglichst genau einschränken und einige typische Missverständnisse ausräumen.

Erst im letzten Teil gehen wir auf reduktionistische Positionen ein, die aufgrund der dargebrachten Ergebnisse aus der Metamathematik nicht mehr vertretbar sind.

Begrifflichkeit

Es ist zu beachten, dass mathematische Begriffe niemals die Bedeutung aus der Alltagssprache haben. Auch wenn die Bezeichnungen freundlicherweise oft so gewählt werden, dass sie durch ihren Alltagsgebrauch eine Intuition ihrer Bedeutung vermitteln, ist immer Vorsicht geboten. Unter einem *Alphabet* in der Mathematik sollte man sich eher eine Menge von Wörtern vorstellen und unter einem *Wort* eher einen Satz. Die Zeichenkette

$$2 + 3 = 5$$

ist ein typisches Wort mathematischer Sprachen, alltagssprachlich „zwei plus drei ist gleich fünf“ handelt es sich um einen Satz. Folgerichtig sollte man sich unter einer *Sprache* besser nicht eine Menge von Wörtern vorstellen, sondern eine Menge von Sätzen. Man sollte trotzdem flexibel mit diesen Vorstellungen umgehen, denn die mathematischen Bezeichnungen sind nicht grundlos: Springt man in die Metasprache, wird plötzlich jeder Satz der Objektsprache zu einem Wort, dadurch wird auch jedes Wort zu einer Art Alphabetzeichen. Kurz und gut: Was ein Satz, was ein Wort und was ein Alphabetzeichen ist, hängt von der Ebene der Betrachtung ab. Wenn wir also *Wort* sagen, könnte man *Satz* lesen und umgekehrt.

Wir werden Begriffe stets in ihrer mathematischen Bedeutung verwenden, werden sie aber in der Regel nicht definieren. Für die Definitionen verweisen wir auf [Hof09] und [Hof18]. Definiert werden hier nur Begriffe, die für die Argumentation entscheidend sind und auch nur dann, wenn die Bedeutung im Alltagsgebrauch relevant abweicht.

1. Mathematik

1.1. **Die Sätze Skolems**

Wir beginnen mit einem der berühmtesten und wichtigsten Sätze der Mathematik und der Philosophie. Erst sein Beweis ermöglichte die Beweise der Sätze von Tarski, Turing und Gödel:[1]

Theorem 1 (Cantor)**.** *Sei M eine Menge und $\mathcal{P}(M)$ die Potenzmenge von M, also die Menge aller Teilmengen aus M. Dann reichen die Elemente aus M nicht aus, um die Elemente aus $\mathcal{P}(M)$ abzuzählen.*

Der Satz gehört zur Mengenlehre, ist aber auch außerhalb der Mathematik bekannt, weil er impliziert, dass immer größere Unendlichkeiten existieren. Oder wie Georg Cantor formuliert: Jeder unendlichen Menge kann eine Menge zur Seite gestellt werden, deren Unendlichkeit noch größer ist.

Definitionen. Eine Menge heißt *abzählbar*, wenn ihre Elemente mit den natürlichen Zahlen abgezählt werden können. Andernfalls heißt sie *überabzählbar*. Die abzählbare Unendlichkeit wird $\beth_0$ genannt. Die aufsteigende Reihe

$$\beth_0, \beth_1, \beth_2, \ldots$$

von Unendlichkeiten ist definiert durch $\beth_{n+1} = \mathcal{P}(\beth_n)$. Sie ist also die Reihe der Potenzmengen von $\beth_0$ und heißt *Beth*-Reihe.

$\beth_0$ ist die Unendlichkeit, die jeder in Gedanken kennt und die man sich gerade noch so vorstellen kann. Überabzählbare Unendlichkeiten sind unvorstellbar groß. Die Beth-Reihe hat eine interessante philosophische Bedeutung, die Alfred Tarski herausarbeitete: Jede Wissenschaft wohnt in den Beths; zwar breiten sich fast alle Wissenschaften auf benachbarte Beths aus, aber sie bleiben immer vollständig in der Beth-Reihe enthalten. Was das heißen soll, werden wir im Detail sehen. Um zu verstehen, warum es so ist, wird eine unmittelbare Folgerung aus Theorem 1 entscheidend sein:

Folgerung 2. Es gibt keine Menge (insbesondere auch kein Beth), die alle Elemente der eigenen Potenzmenge enthält.

Thoralf Skolem konnte zeigen, dass Theorem 1 kein logisch beweisbares Theorem ist. Trotzdem ist es mathematisch beweisbar und damit innerhalb der Mengenlehre mathematisch wahr. Wir werden zuerst eine Annäherung an diese Diskrepanz versuchen und gehen erst ab Seite 15 auf Tarski und die Beth-Reihe ein.

Die Logik der Mathematik

Skolem formulierte die Mengenlehre als eine Theorie in der „Prädikatenlogik erster Stufe“ und Bourbaki erhob die Mengenlehre zur Grundlage der Mathematik. Seitdem verwendet die Mathematik ausschließlich die Prädikatenlogik erster Stufe. Wir werden sie hier mit $\mathcal{L}_1$ bezeichnen. $\mathcal{L}_1$ hat viele Vorteile, doch gibt es andere Logiken, die man verwenden könnte, warum gerade $\mathcal{L}_1$? Es wäre wohl zu leichtfertig, diesen Usus als Tradition abzutun. Der wichtigste Grund für die Wahl von $\mathcal{L}_1$ ist, dass sie eine *Prädikaten*-Logik ist. Ein Prädikat soll eine Relation sein. Leider ist nicht jede Relation ein Prädikat, aber die Bezeichnung „Prädikatenlogik“ ist nur ein historischer Unfall, gemeint ist eine Logik, die über Relationen zwischen Individuen sprechen kann. Bourbaki (wie zuvor schon Descartes) hat die Mathematik als die Wissenschaft der Strukturen charakterisiert, die sich allein mit den Relationen zwischen Individuen beschäftigt, nicht aber mit den Individuen selbst (bei Descartes ist das wohl als Definition gemeint). Früher hießen solche Logiken passender *Funktionenkalküle*, die Bezeichnung *kartesische Logiken* wäre genauer und für Mathematiker selbsterklärend: Relationen sind mathematisch gesehen kartesische Produkte, und kartesische Produkte sind in der Mathematik überall, selbst Funktionen sind kartesische Produkte. Ganz präzise also: Eine Prädikatenlogik ist eine Logik, die über kartesische Produkte und ihre Teilmengen sprechen kann. Jede andere Logik wäre für die Mathematik zu ausdrucksarm. Aber es gibt viele andere Prädikatenlogiken. Die Sätze von Skolem beziehen sich auf die Grenzen von $\mathcal{L}_1$, es ist daher angebracht anzuführen, warum $\mathcal{L}_1$ auch unter den Prädikatenlogiken unersetzlich ist.

Definition (Logik)**.** Eine *Logik* (oder auch *Kalkül* genannt) ist eine formale Sprache mit einer Grammatik vom Typ-0, sowie eine nichtleere Menge von Deduktionsregeln, die die Grammatik respektieren.

Wer nicht zufällig weiß, was die Chomsky-Hierarchie ist, hat keine Chance, die Definition zu verstehen. Aber das muss man hier für diesen Abschnitt noch nicht, wir gehen im nächsten Abschnitt (Tarski) auf formale Sprachen und im übernächsten (Turing) auf Typ-0-Grammatiken ein. Dort wird es wichtig sein. Jetzt für die Sätze Skolems braucht man nur soviel: Eine Logik liefert

- eine Sprache, um die Theorie zu formulieren,
- eine Syntax, um neue Begriffe in die Sprache zu integrieren, sowie
- Deduktionsregeln, um Folgerungen zu ziehen, die in der Sprache formulierbar sind.

Die Erweiterung der Sprache geschieht dabei nicht nur durch neue Wörter, sondern auch durch eine besondere Erweiterung des Alphabets mit Sonderzeichen. Sie werden zur Bezugnahme auf Dinge, die in der Theorie vorkommen (Individuen, Relationen etc.), gebraucht und heißen *Symbole* der Theorie. Sie ermöglichen, die Ontologie (die Existenzbehauptungen einer Theorie) derart zu formulieren, dass sie nicht mit der logischen Sprache verwechselt wird.

Eine Logik ist also nicht nur ein Regelwerk, um Folgerungen zu ziehen, sie ist auch eine Sprache mit einer Syntax. Das ist nicht ganz das traditionelle Verständnis von Logik, aber es ist auch nicht vollkommen etwas anderes: In der Tradition Aristoteles' bestand der Wunsch, Sätze, die aus rein syntaktischen Gründen aus den Annahmen folgen, als *logisch* zu bezeichnen. Dagegen sollte ein Satz *theoretisch* heißen, wenn seine Herleitung von der Semantik der Sprache abhängt.[2] Die Semantik eines Ausdrucks ist seine Bedeutung, sie hängt von dem ab, was es gibt und was es nicht gibt (*Ontologie*). Diese Trennung spiegelt sich in der modernen Definition von Logik durch die Unterscheidung zwischen den Symbolen der Theorie und der logischen Syntax. Aristoteles machte in seiner *Analytica Priora* einen ersten Versuch herauszufinden, welche die Deduktionsregeln der Logik in diesem Sinne sind. Auch wenn sein Versuch als gescheitert gilt,[3] war die Logik, die er entwickelte, eine Millenniumsleistung, sie galt lange Zeit als *die* Logik, die einzig richtige und korrekte Logik. Es gab Ungereimtheiten und nicht wenige Denker versuchten, eine andere Logik zu entwickeln. Aber erst mit Tarskis Schrift [Tar36] kam die Wende, Tarski kehrte das traditionelle Verständnis um: Eine logische Folgerung ist korrekt, so definiert Tarski, wenn sie mathematisch wahr ist. Die mathematisch wahre Folgerung nennt man heute auch *mathematisches Schließen* oder *modelltheoretische Folgerung* und verwendet das Symbol $\models$. Die logische Folgerung wird traditionell *Deduktion* genannt und man verwendet das Symbol $\vdash$. Aber bevor wir genauer auf $\models$ und $\vdash$ eingehen, führen wir einige Begriffe ein, weil durch Tarskis ulkige Definition plötzlich das Problem auftaucht, dass manche Logiken inkorrekt sein können.

Definitionen. In der Tradition seit Tarski werden als Kriterien für die „Güte" einer Logik folgende Begriffe verwendet, die auch für unsere Argumente zentral sein werden:

- syntaktisch widerspruchsfrei (oder kurz: *widerspruchsfrei*)
- semantisch widerspruchsfrei (oder *korrekt*)
- syntaktisch vollständig
- semantisch vollständig
- entscheidbar

Es gibt Logiken, mit denen eine Aussage x und die Negation von x herleitbar ist. Solche Logiken heißen **syntaktisch widersprüchlich**. Sie können durchaus interessant sein, in der frühen Sowjetunion fanden sie viel Aufmerksamkeit, heute sind sie als sogenannte Random-Algorithmen wieder en vogue. Im Vorgriff auf das nächste Kapitel sagen wir: Für die Wissenschaft sind widersprüchliche Logiken unbrauchbar. Aber auch für die Verwendung in mathematischen Beweisen sind sie unbrauchbar, weil sie **semantisch widersprüchlich** sind (inkorrekt): Man kann mit ihnen Folgerungen ziehen, die mathematisch nicht wahr sind. Drücken wir es mit Symbolen aus, heißt das: Falls $\vdash$ mehr als $\models$ beweisen kann, ist $\vdash$ inkorrekt. Da die Mathematik widerspruchsfrei ist, folgt aus dem zweiten das erste Kriterium. Das Minimum für die Verwendung einer Logik in der Mathematik ist also das zweite Kriterium: Die Logik muss in sich selbst widerspruchsfrei sein (Kriterium 1), aber auch gegenüber der Mathematik darf sie nicht in Widerspruch stehen (Kriterium 2).

Die Logik des Aristoteles[4] erfüllt beides, aber sie ist **syntaktisch unvollständig**: Es gibt Aussagen, die sie ausdrücken kann, aber weder beweisen noch widerlegen kann. Somit ist sie erst recht **semantisch unvollständig**: Es gibt mathematisch wahre Folgerungen, die sie zwar ausdrücken, aber nicht beweisen kann. Mit unseren Symbolen ausgedrückt: $\vdash$ kann (innerhalb der Sprache) weniger beweisen als $\models$. Und das war historisch das große Problem für die Logik des Aristoteles: Sie war sehr mächtig, was ihre Ausdruckskraft anging, aber kaum brauchbar, wenn es darum ging, etwas zu beweisen. Die erste uns bekannte alternative Logik war die *Logik vom Port-Royal* von 1662. Die Motivation der Autoren Antoine Arnauld und Pierre Nicole war ausdrücklich, genau diesen Mangel zu beheben. Zwar war ihr Ansatz richtig: Sie wollten von der damals gerade geborenen modernen Mathematik lernen (zur Unterscheidung der alten wurde sie damals „Analysis“ genannt), aber sie kamen nicht weit. Erst spätere Arbeiten basierend auf der modernen Algebra führten schließlich zu $\mathcal{L}_1$. Diese ist semantisch widerspruchsfrei und semantisch vollständig: Innerhalb der Sprache stimmen $\vdash$ und $\models$ überein (hier Satz 4). Also ist jede mathematische Folgerung, die sich mit $\mathcal{L}_1$ ausdrücken lässt, innerhalb von $\mathcal{L}_1$ beweisbar. Von einem solchen Glück wagten Arnauld und Nicole nicht zu träumen.

Die **Entscheidbarkeit** ist mehr: Ein Ausdruck sei gegeben, entscheide, ob er in der Logik beweisbar ist. Alan Turing bewies, dass $\mathcal{L}_1$ unentscheidbar ist (hier Satz 23). Wie soll man sich das vorstellen: semantisch vollständig und zugleich unentscheidbar? Das ist nicht so einfach, aber wichtig: Wir gehen ab Seite 24 genauer darauf ein. Es stellt sich die Frage, warum nicht eine Logik wählen, die alle fünf Kriterien erfüllt. Derartige Logiken gibt es viele. Der Nachteil ist, dass sie extrem ausdrucksarm sind. Die sogenannte Aussagenlogik erfüllt alle Kriterien, aber sie kann nicht ausdrücken, dass 3 eine Primzahl ist. Eine Logik ist wie gesagt eine Sprache, und nicht mit jeder Sprache lässt sich alles sagen. Die Situation ist also die, dass wir in der Beweiskraft einen Kompromiss finden müssen zwischen dem Reichtum an Ausdruck bei Aristoteles und der Armut entscheidbarer Logiken. Dieser Kompromiss ist $\mathcal{L}_1$.

Aber es ist mehr als nur ein Kompromiss. Der Kompaktheitssatz (hier Satz 5) liefert einen großen Vorteil: Mit $\mathcal{L}_1$ haben wir ein Instrument in der Hand, um das Unendliche mit endlichen Mitteln zu erforschen. Der sogenannte *Satz von Lindström* besagt, dass $\mathcal{L}_1$ die einzige Prädikatenlogik ist, die semantisch vollständig ist und den Kompaktheitssatz erfüllt. Damit ist $\mathcal{L}_1$ in der Mathematik unersetzbar. Wir werden daher von nun an von *der* Logik sprechen.

1.1.1. *Die Grenzen der Logik*

Definitionen. Ist $\mathcal{L}$ eine Logik, dann ist eine $\mathcal{L}$-*Theorie* eine Menge von $\mathcal{L}$-Sätzen zusammen mit allen aus ihnen in $\mathcal{L}$ herleitbaren Sätzen. Da ist ein Unterschied zur Alltagssprache: Alle Folgerungen aus der Theorie (auch die zu einem Zeitpunkt noch nicht entdeckten) gehören zur Theorie. Mit $\mathcal{L}$-Sätzen ist gemeint, dass die Sätze in der Syntax von $\mathcal{L}$ formuliert sind. Weiter stelle man sich unter einer *Interpretation* eine „denkbare Welt“ vor. Unter einem *Modell der Theorie* stelle man sich eine „denkbare Welt, in der alle Sätze der Theorie wahr sind,“ vor. Mathematische Wahrheit ist also immer relativ zur Welt, die man sich gerade denkt. Eine Folgerung, bei der B aus A gefolgert wird, nennen wir *mathematisch wahr*, wenn in jeder denkbaren Welt, in der A wahr ist, auch B wahr ist. Symbolisch $A \models B$.

Wiederholen wir Tarskis Definition: Eine logische Folgerung, symbolisch $\vdash$, ist korrekt, wenn sie mathematisch wahr ist. Nun haben wir definiert, was das bedeutet. Ist jetzt die logische Folgerung von der Ontologie unabhängig? Es sieht so aus, aber bei genauerem Hinsehen wird es doch fraglich, da die Wahrheit in einer denkbaren Welt von den Dingen abhängt, die dort vorkommen. Mit anderen Worten: Ontologische Argumente (und nur diese) können eine logische Folgerung widerlegen. Das vernichtet das traditionelle Verständnis.

Definitionen. Eine Theorie heißt mathematisch inkorrekt, wenn sie kein Modell hat, wenn sie also „undenkbar“ ist: Man kann sich keine Welt denken, in der alle Sätze der Theorie wahr sind. Nach Tarski dürfen wir *widersprüchlich* als „undenkbar“ verstehen: Etwas sei widersprüchlich, bedeutet, es sei undenkbar.

Der folgende Satz von Skolem[5] besagt, dass wir auch *widerspruchsfrei* mit „denkbar“ identifizieren können. Jedenfalls in $\mathcal{L}_1$:

Lemma 3 (Skolem-Gödel-Theorem)**.** *Eine $\mathcal{L}_1$-Theorie ist genau dann (syntaktisch) widerspruchsfrei, wenn sie ein Modell hat.*

$\mathcal{L}_1$ garantiert also einen Zusammenhang zwischen Syntax und Semantik. Nur wenige Logiken haben diese Eigenschaft. Eine einfache, aber wichtige Folgerung ist der Gödelsche Vollständigkeitssatz.

Satz 4 (Gödelscher Vollständigkeitssatz: Skolem, Gödel, Henkin)**.** *$\mathcal{L}_1$ ist semantisch vollständig. Mit anderen Worten: Sind A und B aus $\mathcal{L}_1$, dann gilt: Aus $A \models B$ folgt $A \vdash B$.*

Weil $\mathcal{L}_1$ auch semantisch widerspruchsfrei ist, gilt zudem die Umkehrung: Aus $A \vdash B$ folgt $A \models B$. D. h. innerhalb von $\mathcal{L}_1$ stimmen $\models$ und $\vdash$ überein. Damit erhält $\mathcal{L}_1$ eine Art Zulassung für mathematische Beweise. Henkin konnte auch beweisen, dass keine Prädikatenlogik höherer Stufe diese Eigenschaft besitzt.

Satz 5 (Kompaktheitssatz: Skolem, Gödel, Malcev)**.** *Eine $\mathcal{L}_1$-Theorie hat genau dann ein Modell, wenn jede endliche Teilmenge der Theorie ein Modell hat.*

Beweis. Wir verwenden die Übereinkunft, dass ein Beweis nur endlich lang sein darf. Die Theorie hat laut dem Skolem-Gödel-Theorem genau dann kein Modell, wenn sie einen Widerspruch enthält. Der Widerspruch müsste aber dann beweisbar sein (weil $\mathcal{L}_1$ vollständig ist). $\mathcal{L}_1$ hat weiter die Eigenschaft, dass sie nur aus endlich vielen Annahmen Folgerungen zieht. Da der Beweis endlich ist, nutzt er also nur endlich viele Voraussetzungen, und so gäbe es eine endliche Teilmenge von Sätzen aus der Theorie, aus der sich der Widerspruch herleiten lässt. Die endliche Teilmenge hat dann wieder gemäß dem Skolem-Gödel-Theorem kein Modell. □

Die Sätze der Logik haben unter Mathematikern den Ruf, völlig nutzlos zu sein. Eine Ausnahme ist der Kompaktheitssatz, nicht nur, weil er in vielen Bereichen zur Anwendung kommt, sondern auch, weil Kompaktheitsbeweise ästhetisch kaum zu übertreffen sind.

Kommen wir aber nun zu den Konsequenzen dieser Sätze. Um sie zu formulieren brauchen wir den Begriff *Nicht-Standard-Modell.* Wie gesagt, eine Interpretation ist eine denkbare Welt, in der nicht unbedingt alle Sätze der Theorie wahr sind; ein Modell ist sozusagen eine „gelungene“ Interpretation: Jeder Satz der Theorie ist wahr. Wenn man sich so eine Welt denkt, kommen da Dinge vor. Nicht jede Theorie spricht über alle Dinge. Ergo: Für ein Modell braucht man nur an die Dinge zu denken, die in der Theorie vorkommen. Diese heißen *Individuen.* Spricht eine politische Theorie nur über Staaten, Institutionen und Menschen, nicht aber über Zellen und Atomen, so sind Zellen und Atomen keine Individuen, Individuen sind nur Staaten, Institutionen und Menschen. Falls aber eine Theorie nur von Atomen und Elementarteilchen spricht, dann sind Menschen dort keine Individuen.

Definition (Universum)**.** Die Gesamtheit der Individuen einer Interpretation heißt *Universum*.

Grob gesagt ist das Universum die Ontologie an die man denkt, um eine Theorie zu interpretieren.

Definition. Sei eine Theorie mit einem Modell gegeben. Ein *Nicht-Standard-Modell* ist ein anderes Modell mit einem anderen Universum.

Beispiel: Ein Realist denkt, dass wir real sind. Nun könnte man aber seine Theorien so interpretieren, dass wir nur in einem Buch vorkommen, eine höhere Vernunft uns liest und wir nur in ihrer Phantasie real sind. Das wäre ein Nicht-Standard-Modell des Realismus.[6] Solche Uminterpretationen von Theorien sind nur und nur dann möglich, wenn einige Begriffe der Theorie mehrdeutig sind.

Folgerung 6. Der Begriff der Identität ist in $\mathcal{L}_1$ nicht definierbar.

Beweis. Um zu zeigen, dass ein Begriff in einer Theorie nicht eindeutig definiert ist, reicht es, ein Nicht-Standard-Modell anzugeben, in dem der Begriff eine andere Bedeutung hat. Sei eine Theorie mit einem Modell und dem dazugehörigen Identitätsbegriff gegeben. Wir interpretieren das Modell um, indem wir uns eine Verdoppelung denken: Verdoppelung des Universums und aller Relationen. Borges hätte *Spiegelung* geschrieben. Dann hat der Identitätsbegriff eine neue Bedeutung, weil es logisch nicht auszumachen ist, ob ein Individuum oder sein Spiegelbild gemeint ist. Schließlich kann man zeigen, dass aufgrund der Eigenschaften von $\mathcal{L}_1$ die neue Interpretation immer noch ein Modell der Theorie ist.[i] □

Um das Problem zu umgehen, wird gern die Gleichheit als logischer Begriff als Teil der Sprache $\mathcal{L}_1$ definiert. Das ist suspekt, da die Gleichheit ein rein semantischer Begriff ist: $x = y$ soll nicht heißen, dass die Terme x und y identisch sind, sondern, dass x und y dasselbe bezeichnen, dass sie also semantisch gleich sind. Aber nehmen wir es hin, es gibt andere Probleme.

Folgerung 7. Der Begriff der Endlichkeit ist in $\mathcal{L}_1$ nicht definierbar.

Beweis. Nehmen wir an, es gibt eine $\mathcal{L}_1$-Theorie der Endlichkeit. Dann gilt sie für jede endliche Welt. Genauer: Jede endliche Interpretation erfüllt die Sätze der Theorie. Gemäß dem Kompaktheitssatz gibt es also eine unendliche Interpretation, die sie ebenfalls erfüllt. Also gibt es ein Nicht-Standard-Modell, in dem der Begriff der Endlichkeit auch unendliche Universen umfasst. □

---

[i] *Anm. für Wissenschaftler.* Formal: In $\mathcal{L}_1$ kann die Gleichheit immer als Äquivalenzrelation zwischen Äquivalenzklassen interpretiert werden. Könnte $\mathcal{L}_1$ Beziehungen zwischen Mengen ausdrücken, könnte sie sagen, dass genau eine Menge von Individuen existiert, die die Sätze der Theorie erfüllen, aber das kann sie nicht.

Der folgende Satz wird manchmal mit dem ersten Unvollständigkeitssatz Gödels verwechselt. Gödel selbst bemerkte, dass eine schwächere Version aus seinem Satz folgt, aber die Sätze sind nicht äquivalent.

Folgerung 8 (Skolem). Die Zahlenreihe $0, 1, 2, \ldots$ ist logisch nicht definierbar: Jede $\mathcal{L}_1$-Theorie der Arithmetik hat Nicht-Standard-Modelle.

Beweis. Wir erweitern das Universum um eine unendlichgroße Zahl z. Dazu fügen wir den Axiomen unendlich viele weitere Axiome hinzu: Zu jeder natürlichen Zahl nehmen wir an, z sei größer. Und wenden den Kompaktheitssatz an: Die Annahmen sind in jeder endlichen Teilmenge denkbar, also ist sie auch im Unendlichen denkbar. Also gibt es ein unendliches Modell mit einer Zahl z, die größer ist, als jede natürliche Zahl. Die neue Zahl muss genau dieselben Eigenschaften der natürlichen Zahlen haben: Sie hat einen Vorgänger, sie hat einen Nachfolger, sie kann prim oder nicht prim sein usw. Es ist, als ob wir die natürlichen Zahlen jenseits des Unendlichen doppelt gespiegelt hätten.[7] □

Das bedeutet: Man kann, ohne Widerspruch zu erzeugen, das Axiom hinzufügen, dass es eine unendlichgroße Zahl gibt. Man kann aber auch die Negation des Satzes als Axiom hinzufügen. Dann hätte man solche Modelle ausgeschlossen, aber keinen logischen Widerspruch erzeugt.

Theorem 9 (Skolem)**.** *Jede $\mathcal{L}_1$-Theorie der Mengenlehre hat Nicht-Standard-Modelle.*

Beweis. Wir beweisen hier den Satz nur für den Fall abzählbar vieler Symbole. Denken wir an ein Sandbuch à la Borges: Ein Sandbuch ist ein Buch mit unendlich vielen Seiten. Nehmen wir an, in diesem Buch sind alle Beweise der Mengenlehre enthalten, aber sonst nichts anderes. Da alle Beweise endlich lang sind und jeder Beweis nur endlich viele Voraussetzungen verwendet, ist das Buch nur abzählbar unendlich. Nun kommt eine besondere Eigenschaft von $\mathcal{L}_1$ ins Spiel: $\mathcal{L}_1$ kann nur Individuen untereinander in Beziehung setzen, sie kann also zwar dadurch von Gruppen von Individuen sprechen, aber sie kann keine Beziehung zwischen Gruppen ausdrücken. Jeder Name (jedes „Symbol“) steht immer nur für Individuen und niemals für eine Gruppe oder gar für eine Gruppe von Gruppen. Wenn aber das Buch nur abzählbar lang ist, kommen im Buch nur abzählbar viele Namen vor; also muss eine Welt denkbar sein, in der abzählbar viele Individuen vorkommen und in der alle Sätze der Mengenlehre wahr sind. □

Wie wir in Theorem 12 sehen werden, ist das ein Satz mit weitreichenden Folgen für die Wissenschaftsphilosophie. Skolem veranschaulichte die Aussage seines Satzes mit einem Paradoxon (einem scheinbaren Widerspruch): Er fand einen $\mathcal{L}_1$-Satz, der besagt, dass es überabzählbar viele Individuen gibt. Dann fügte er diesen Satz den Axiomen der Mengenlehre hinzu. Das ist aber immer noch eine $\mathcal{L}_1$-Theorie und hat entsprechend ein abzählbares Modell. Wie ist das möglich? Ist Cantors Beweis der Existenz überabzählbarer Mengen (Theorem 1) falsch? Nein, man könnte ja dem Axiomensystem die Negation des Satzes hinzufügen und es ergäbe sich trotzdem kein Widerspruch. Kurz anhalten: Was bedeutet das? Man hat eine Theorie, in der man eine Aussage A annehmen kann, ohne auf logische (=syntaktische) Widersprüche zu stoßen. Stattdessen könnte man aber auch die Negation von A annehmen und es gäbe trotzdem keinen syntaktischen Widerspruch. Dann ist doch A in der Theorie weder logisch beweisbar noch logisch widerlegbar! Der Beweis Cantors ist nicht falsch, der Satz ist nur nicht logisch beweisbar. Skolem gab folgende Erklärung: Der Beweis Cantors gründet auf Begriffen und das Problem ist, dass diese Begriffe nicht logisch eindeutig definierbar sind. Dadurch verwenden wir logisch nicht fassbare, versteckte Annahmen, die nur in unserer Anschauung vorkommen.

Das Löwenheim-Skolem-Tarski-Theorem ist die wichtigste Folgerung für die Mathematik. Skolem bewies es zuerst, aber Anatoli Malcev brachte es in eine fruchtbare Form. Der Name kommt daher, dass Skolem den Satz Löwenheim und Tarski in den Mund legte. Er besagt:

Theorem 10 (Skolem, Malcev). *Hat eine $\mathcal{L}_1$-Theorie ein unendliches Modell, dann hat sie Modelle beliebiger Kardinalität.*

Mit *Kardinalität* ist die Größe der Unendlichkeit gemeint (es gibt unendlich viele Beths und auch andere Unendlichkeiten).

Beweis. Der Beweis ergibt sich aus der Beweisidee mit dem abzählbaren Sandbuch und der Einführung neuer Individuen: Man findet immer ein abzählbares Modell, und dann zeigt man, dass man beliebig viele neue Individuen hinzu denken kann. □

1.1.2. *Die philosophische Bedeutung der Sätze Skolems*

Skolem bezeichnete die Konsequenzen seiner Sätze als „a priori sehr plausibel". Das sind sie mit Sicherheit nicht, sie sind überraschend und subtil. Ein einfaches Beispiel als Warnung: Dass der Endlichkeitsbegriff nicht definierbar ist, bedeutet nicht, dass man keine Theorie angeben kann, die nur endliche Modelle hat (wer es versteht: $\exists x\, \forall y\; x = y$). Auch darf man nicht folgern, dass jeder $\mathcal{L}_1$-Ausdruck mehrdeutig ist; richtig ist: Es gibt gewisse Begriffe, die nicht in $\mathcal{L}_1$ definierbar sind.

Mit Sicherheit können wir sagen: Die Sätze Skolems zeigen Grenzen und warum sie bestehen. Die Grenzen: Einige Grundbegriffe der Mengenlehre sind logisch nicht eindeutig definierbar. Insbesondere kann die Beth-Reihe $\beth_0, \beth_1, \beth_2, \ldots$ nicht als Reihe verschieden großer Unendlichkeiten logisch definiert werden. Der Grund: $\mathcal{L}_1$ kann weder definieren, welche die Gesamtheit der Individuen ist, noch was ein Individuum ist.

Die Frage ist nur: Wessen Grenzen? Skolem kam aus der Algebra, dort ist $\mathcal{L}_1$ die natürliche Logik; für Skolem war sie die unhinterfragbare Logik. Er interpretierte seine Ergebnisse daher als Hinweise dafür, dass die Mengenlehre ungeeignet sei, um als Grundlage der Mathematik zu dienen. Er bezweifelte nie die Fruchtbarkeit der Mengenlehre und bewunderte die „grandiosen" Leistungen Cantors ([Sko62]), suchte aber nach Alternativen für eine Grundlegung der Mathematik. Tarski bot einen Ausweg, der heute Konsens ist: Nicht die Logik sagt, was richtig ist, sondern die Mathematik sagt, welche Logiken richtig sind und wo ihre Grenzen liegen. So gesehen zeigen die Ergebnisse Skolems nur die Beschränktheit der Logik, nicht aber die der Mathematik. Wir übernehmen hier diese Interpretation und veranschaulichen unsere Beweggründe mit einem Beispiel: Cantors Beweis von Theorem 1.

Cantors Diagonalargument

Inwiefern ist Theorem 1 mathematisch wahr, aber logisch nicht beweisbar? Ist es nicht widersinnig zu beweisen, dass in $\mathcal{L}_1$ der logische Beweis $\vdash$ und der mathematische Beweis $\models$ übereinstimmen (das sagt ja Satz 4 aus), und jetzt folgern, dass man in der Mathematik mehr beweisen kann als mit $\mathcal{L}_1$? Um etwas zu beweisen, gibt es zwei Möglichkeiten: Man setzt in eine Aussage Bekanntes ein (eine Definition, ein Axiom oder einen bereits bewiesenen Satz) und der sich daraus ergebende Satz gilt als mathematisch bewiesen. „Geistlose Manipulation von Symbolen", hört man von Mathematikern oft sagen. Die Alternative ist der Widerspruchsbeweis. Blaise Pascal war wohl der erste, der die beiden Methoden unterschied; zur zweiten schreibt er:

> Und darum muss man jedes Mal, wenn ein Lehrsatz unvorstellbar ist, mit seinem Urteil zurückhalten und ihn nicht auf Grund dieses Kennzeichens ablehnen, sondern dessen Gegenteil prüfen; und wenn man dieses als offensichtlich falsch erkennt, kann man kühn den ursprünglichen Lehrsatz bejahen, so unbegreiflich er auch sein mag.
> *Vom geometrischen Geist und von der Kunst zu überzeugen.*

Wenn man das beherzigt, dürfte Cantors Beweis nachvollziehbar sein.

Beweis von Theorem 1. Nehmen wir an, es gibt eine Abzählung der Elemente von $\mathcal{P}(M)$ durch Elemente aus der Menge M. Betrachte jetzt ein beliebiges Element m aus M, das ein Element aus $\mathcal{P}(M)$ abzählt. Nennen wir es p: Das Element m zählt das Element p ab. Die Menge $\mathcal{P}(M)$ war die Menge aller Teilmengen aus M, das bedeutet: Jedes Element aus $\mathcal{P}(M)$ ist eine Menge aus Elementen aus M. Es gibt jetzt also genau zwei Möglichkeiten: m ist ein Element aus p oder es ist kein Element aus p. Dieses Entweder-Oder ist ein Axiom der Mengenlehre: Man ist oder man ist nicht Element einer gegebenen Menge. Nennen wir all jene Elemente m, die nicht in der Menge p, die sie abzählen, enthalten sind, *widerspenstige Elemente*. Genau an dieser Stelle scheitert die Logik, aber deine Vernunft ist hoffentlich noch dabei ist. Wenn du Wochen brauchst, um durchzublicken: keine Sorge, das ist normal. Weiter: Betrachte die Menge aller widerspenstigen Elemente. Diese Menge besteht aus Elementen aus M, also ist sie ein Element aus $\mathcal{P}(M)$. Da wir annahmen, dass es eine Abzählung gibt, gibt es ein Element x aus M, das die Menge aller widerspenstigen Elementen abzählt. Und nun fragen wir uns, ob x ein widerspenstiges Element ist. Das kann nicht sein, weil x gerade diese Menge abzählt. Also ist x kein widerspenstiges Element. Das bedeutet, dass die Bedingung, um ein widerspenstiges Element zu sein, nicht erfüllt ist: x zählt eine Menge ab, die x enthält. Widerspruch. Also ist gemäß Pascal das Gegenteil der Annahme richtig, was zu beweisen war.[i] □

[i] In Formeln ist es hübscher: Angenommen $p : M \to \mathcal{P}(M)$ sei surjektiv, dann hat $\{m \in M \mid m \notin p(m)\} \in \mathcal{P}(M)$ ein Urbild $x \in M$. Ist $x \notin p(x)$? Widerspruch.

Nun kannst du selbst entscheiden, was es bedeutet, dass ein Satz mathematisch wahr, aber mit der Logik $\mathcal{L}_1$ unbeweisbar ist. Wir schlagen vor: *Da hat* $\mathcal{L}_1$ *Pech gehabt*. Möchte jemand unbedingt einen logischen Beweis des Satzes sehen, kann er sich eine passende Logik ausdenken, mit der der Beweis aufgeht, den Beweis logisch formulieren und die Logik auf Nimmerwiedersehen in den Mülleimer werfen. In diesem Sinne gibt es Wahrheiten, die unserer Vernunft zugänglich sind, die aber jenseits unserer gerade verwendeten Logik liegen.

Es wäre denkbar, eine Art Hierarchie von Logiken aufzubauen, sodass man die neue Logik anstatt auf den Müll zu werfen in eine Schublade für spätere Wiederverwendung legen könnte. Das ist der Grund, warum $\mathcal{L}_1$ Logik „erster Stufe" heißt, sie sollte die erste Ebene einer solchen Hierarchie bilden. Die wohl einfachste Hierarchie ist die monadische Hierarchie. Die „erste Stufe" steht für die Logik, die über Individuen spricht, „zweite Stufe" für die, die *zusätzlich* über Mengen von Individuen spricht, „dritte Stufe" für die, die *zusätzlich* über Mengen von Mengen von Individuen spricht. Usw. Denken wir an die Beth-Reihe: Mit $\mathcal{L}_1$ kann man in $\beth_0$ alles erfassen; mit einer Logik der Stufe $n+1$ könnte man alle $\beth$-Stufen von $\beth_0$ bis einschließlich $\beth_n$ erfassen. Leon Henkin zeigte, dass diese Logiken semantisch unvollständig sind. Durch seinen Beweis wurde die Idee einer Hierarchie nutzlos, denn diese Logiken machen uns blind für gewisse mathematisch wahre Folgerungen. Statt dessen müssen wir also $\mathcal{L}_1$ für jede Beth-Stufe verwenden.

Zwei unmittelbare Konsequenzen: Wir können auf der n-ten Stufe nur über die Elemente in $\beth_n$ sprechen, nicht über die niedrigerer Stufe, weil $\mathcal{L}_1$ nur über Individuen spricht, sodass die Mengen in $\beth_n$ als Ur-Elemente behandelt werden müssen (damit sind Elemente gemeint, die keine Mengen sind). Wir werden also blind dafür, dass die Individuen, über die wir sprechen, in Wirklichkeit Mengen sind.

Zweitens: Da die angeführten Sätze Skolems auf jeder Stufe gelten, bleiben alle Folgerungen Skolems auf jeder Stufe erhalten.

Die ursprünglich angedachte Hierarchie hängt damit zusammen, dass $\mathcal{L}_1$ über Relationen spricht. Dann wird auf der nächsten Ebene nicht nur über (lose) Mengen von Individuen gesprochen, sondern auch über Mengen von Individuen, die in gewissen Relationen zueinanderstehen. Sind die Individuen Atome, so können wir auf der nächsten Ebene Atome, die in gewissen Beziehungen zueinander stehen, zu einen Subjekt machen, über das Dinge aussagbar sind.

Tarski zeigte die praktischen Konsequenzen der Sätze Skolems: Die Hierarchien entlang der $\beth$-Reihe sind unumgänglich für das Verstehen mathematischer Modelle, also auch für die Wissenschaft. Ferner zeigte er, wie es sein kann, dass $\models$ mehr beweist als $\vdash$. Satz 4 zum Trotz.

1.2. **Die Sätze Tarskis**

Der Ausgangspunkt für Tarski war folgende Überlegung: Nehmen wir an, wir könnten $\beth_0$-unendlich viele Aussagen beweisen:

Erste Aussage: 1 besitzt die Eigenschaft E
Zweite Aussage: 2 besitzt die Eigenschaft E
Dritte Aussage: 3 besitzt die Eigenschaft E
$\vdots$
usw.

Im Allgemeinen jeden Satz der Form

n besitzt die Eigenschaft E

Wobei n für eine natürliche Zahl steht. Dann lässt sich der Satz:

Alle natürlichen Zahlen besitzen die Eigenschaft E

mit den klassischen Deduktionsregeln nicht ableiten. Schon Aristoteles befasste sich in seiner *Analytica Posteriora* ausführlich mit dieser Frage. Er wollte die logische Folgerung als eine rein syntaktische Folgerung verstanden wissen und musste daher sagen, dass die Verallgemeinerung, dass es für alle gilt, logisch unhaltbar ist, da sie ja eine ontologische Annahme macht: Dass es nichts mehr gibt als das „Alle“. Aber auch in der modernen Logik ist die Folgerung problematisch, denn die Aneinanderreihung der Beweise wäre unendlich lang, also kein Beweis.

Induktionsbeweise gehören dazu und Aristoteles musste konsequent die mathematische Induktion ablehnen. Es dauerte bis Blaise Pascal, bis die Induktion wieder als Beweisprinzip formuliert wurde.[8] Das machte aber nur den Widerspruch offensichtlicher: In keiner Logik ist sie beweisbar, es sei denn, sie wird ausdrücklich als Deduktionsregel hinzugefügt. Das erkannte Giuseppe Peano, deswegen spricht man von *Peano-Axiomen*. In $\mathcal{L}_1$ ist die Induktion unsagbar, sie wird durch unendlich viele Axiome ersetzt, aber auch dann ist sie nicht wirklich die Induktion Pascals und Peanos: In $\mathcal{L}_1$ sind nur Folgerungen aus endlich vielen Voraussetzungen erlaubt.

Tarski gab sich damit nicht zufrieden, denn diese Folgerungen aus unendlich vielen Aussagen sind ja trotzdem wahr.

1.2.1. *Die Grenzen der Beweisbarkeit*

Definitionen. Eine Menge von Aussagen in einer formalen Sprache mit Typ-0-Grammatik heißt *Axiomensystem*. Jede dieser Aussagen heißt *Axiom*. Eine Aussage a heißt mathematisch wahr im Axiomensystem A, wenn $A \models a$ („*in jeder denkbaren Welt, in der die Axiome wahr sind, ist auch die Aussage wahr*“).

Dass die Sprache eine Grammatik vom Typ-0 haben soll, darf man jetzt vorerst verdrängen. Entscheidend ist, dass eine *formale Sprache* eine Sprache ist, deren Alphabet nur endlich viele Zeichen hat und die Wörter endlich lang sind. Dadurch hat die Sprache höchstens abzählbar viele Wörter, siehst du das? Denk an die Bibliothek von Babel: Das Alphabet ist endlich; wenn wir uns denken, dass die Bücher aus der Bibliothek beliebig lang, aber doch nur endlich lang sein dürfen, wäre jedes Buch, das die Regeln der Grammatik respektiert, ein Wort der formalen Sprache. Die Bibliothek wäre unendlich, aber nur abzählbar unendlich. Jede beliebige Menge solcher Bücher darf als Axiomensystem deklariert werden. Nimmt man sich eine Logik, die mit der Grammatik kompatibel ist, kann man Folgerungen aus den Büchern ziehen. Jede Folgerung befindet sich irgendwo in der Bibliothek (Sprache).

Da die Sprache abzählbar ist und ein Axiomensystem eine Teilmenge der Sprache ist, kann das Axiomensystem höchstens abzählbar sein. Immerhin darf es unendlich sein. Historisch gab es dagegen viel Widerstand. Wir wollen uns hier nicht daran stören, da wir die Grenzen von Axiomensystemen zeigen wollen und diese Grenzen gelten selbst dann, wenn man unendlich viele Axiome zulässt.

Tarskis Freund Lindenbaum begann damit, die logischen Beweise abzuzählen. Er zählte die Anzahl der Beweise innerhalb einer Logik, aber Tarski verstand sofort: Logik hin oder her, es geht um die Grenzen der Sprache, nicht um die der Logik! Ein Höhepunkt in der Geschichte der Philosophie.[9]

Lemma 11 (Tarski, Lindenbaum)**.** *Die Menge aller überhaupt axiomatisch beweisbaren Sätze ist abzählbar.*

Beweis. Die Deduktionsregeln der Logik respektieren die Grammatik, das heißt, jede Folgerung durch $\vdash$ führt zu Sätzen innerhalb der Sprache. Die Logik kann also per Definitionem nicht die Sprache verlassen, und man kann demnach höchstens so viele Sätze beweisen, wie in der Sprache formulierbar sind, unabhängig von der Logik, die man wählt. Halten wir nun eine formale Sprache fest. Sie ist höchstens abzählbar. Also können höchstens abzählbar viele Sätze bewiesen werden (die zu beweisende Sätze sind aus dieser Perspektive nur Wörter der Sprache und davon gibt es abzählbar viele).[i] Ein Satz von Noam Chomsky besagt, dass nur abzählbar viele formale Sprachen mit einer Typ-0-Grammatik existieren.[ii] Ein Satz Cantors besagt, dass abzählbare Vereinigungen abzählbarer Mengen wieder abzählbar sind. Wenn wir also jetzt über alle formalen Sprachen zusammenzählen, sind wir immer noch in $\beth_0$. □

Die Sprengkraft folgt weiter unten: In (fast) jedem Axiomensystem gibt es überabzählbar viele Wahrheiten. Um das zu beweisen, benötigen

---

[i] Das ist bemerkenswert: Falls die Sprache unendlich ist, gibt es überabzählbar viele Axiomensysteme und überabzählbar viele Logiken, trotzdem können nur abzählbar viele Sätze axiomatisch bewiesen werden.

[ii] Wir beweisen dies später im Beweis zu Folgerung 21, nachdem wir im Satz 20 angeben, was eine Sprache mit einer Typ-0-Grammatik ist.

wir die Semantik, denn die Frage, was wahr ist und was nicht, ist eine Frage der Semantik. Tarski betrachtete als erster die Semantik formaler Sprachen, sein Ergebnis: Die Semantik eines Begriffes ist durch eine Menge eindeutig gegeben.

Theorem 12 (Tarski). *Jeder Begriff ist eine Menge und umgekehrt.*[10]

Definition. Die *wissenschaftliche Semantik* einer Sprache definiert die Bedeutung eines jeden Wortes als Menge.

Von nun an betrachten wir Sprachen allein in einer wissenschaftlichen Semantik. Aus dieser Perspektive sind die nächsten beiden Sätzen klar.

Satz 13 (Tarski). *Die Metaebene einer Sprache enthält alle Elemente ihrer Potenzmenge.*

Beweis. Die Metasprache zu einer Sprache $\mathcal{L}$ enthält Wörter, deren Bedeutung die Wörter aus der Objektsprache $\mathcal{L}$ sind: Es sind also Mengen aus Elementen aus $\mathcal{L}$ und somit sind diese Wörter aus $\mathcal{P}(\mathcal{L})$. Möglicherweise enthält die Metasprache noch ganz andere Wörter, zum Beispiel auch Mengen von Relationen, aber entscheidend ist nur, dass zu jedem Element der Potenzmenge $\mathcal{P}(\mathcal{L})$ sich ein metasprachliches Wort definieren lässt. □

Umgekehrt kann eine Objektsprache durchaus Wörter aus ihrer Metasprache enthalten.[i] In diesem Fall sagen wir, dass die Sprache einen Teil ihrer Metasprache *formalisieren* kann. Aber gemäß Folgerung 2 kann eine Sprache niemals ihre gesamte Metaebene formalisieren.

Theorem 14 (Tarski). *Wahrheit ist auf Metaebene definierbar und in jedem Axiomensystem gibt es unbeweisbare Wahrheiten.*

Beweis. Nicht in jedem, es muss vorausgesetzt werden, dass unendlich viele mathematisch nicht äquivalente Aussagen beweisbar sind. Die Einschränkung ist aber praktisch bedeutungslos. Wir fassen alle Sätze, die sich in der Sprache des Axiomensystems ausdrücken lassen, als eine Sprache $\mathcal{L}$ auf, wobei wir mathematisch äquivalente Sätze als ein Wort in $\mathcal{L}$ nehmen. Wir definieren den Begriff W wie folgt: W ist die Menge der im Axiomensystem mathematisch wahren Sätze aus $\mathcal{L}$. Mit anderen Worten: W enthält jeden Satz, der sich in der Sprache des Axiomensystem ausdrücken lässt und der mathematisch (nicht: *logisch*!) aus den Axiomen folgt. Offensichtlich ist W ein Element aus $\mathcal{P}(\mathcal{L})$ und daher dort (semantisch) definiert. Jede Teilmenge w aus W, also jedes Element w aus $\mathcal{P}(W)$, lesen wir als: „Die Elemente dieser Menge sind wahr.“ Dann entspricht jede Teilmenge w einer metasprachlichen Aussage, die wahr im Axiomensystem ist: $W \models w$. Offensichtlich sind diese Aussagen paarweise nicht äquivalent. Nach Voraussetzung ist W abzählbar unendlich und somit ist $\mathcal{P}(W)$ überabzählbar. Aber gemäß Lemma 11 gibt es nur abzählbar viele Beweise. □

Tarskis Satz hat eine ungeheure Reichweite in philosophischen Fragen: $\vdash$ bleibt innerhalb der Sprache, während $\models$ notwendigerweise die Sprache transzendiert: Ist $\vdash$ korrekt, dann beweist $\vdash$ weniger als $\models$.

A priori wäre ein Axiomensystem denkbar, in dem zumindest W, die Menge aller in der Sprache des Systems ausdrückbaren Wahrheiten, definierbar ist. Tarski zeigte, dass selbst das unmöglich ist:

---

[i] *Anm. für Wissenschaftler.* Zu jeder Sprache $\mathcal{L}$ lässt sich zum Beispiel die Sprache $L := \mathcal{L} \cup \{x \in \mathcal{P}(\mathcal{L}) \mid |x| < |\mathcal{L}|\}$ definieren, die Teile ihrer Metasprache enthält.

Lemma 15. *Listet man alle beweisbaren Sätze auf, so steht die Aussage, dass all diese Sätze beweisbar sind, nicht auf der Liste.*

Beweis. Sei ein Axiomensystem mit einer Logik gegeben. Sei B die Menge aller im Axiomensystem logisch beweisbaren Aussagen. Sei a die Aussage, dass jeder Satz aus B im Axiomensystem logisch beweisbar ist. Semantisch ist a durch B definiert. Angenommen, a sei in B, dann gäbe es auf der Semantikebene eine Menge, die sich selbst als Element enthält. Dies ist nach dem Fundierungsaxiom der Mengenlehre ausgeschlossen.[i] □

Satz 16 (Tarski)**.** *In einem semantisch widerspruchsfreien Axiomensystem ist W nicht innerhalb des Systems definierbar.*[ii]

Beweis. Wir argumentieren zuerst syntaktisch und dann semantisch. Angenommen W sei im Axiomensystem definierbar. Das bedeutet syntaktisch 2 Sachen: Erstens gibt es einen Ausdruck in der Sprache $\mathcal{L}$ des Axiomensystems, der die Menge der wahren Aussagen W definiert (im System ist „das ist wahr“ sagbar); zweitens kann man im Axiomensystem zu jeder Aussage x beweisen, dass x in W oder dass sie nicht in W liegt (im System kann man über jede Aussage sagen, dass sie wahr ist oder nicht). Zudem folgt aus der semantischen Widerspruchsfreiheit, dass jede beweisbare Aussage wahr ist. Wir können daher die zwei Bedingungen umformulieren: Die Menge W der wahren Aussagen stimmt mit der Menge B der beweisbaren Aussagen überein und der Beweisbegriff ist in der Sprache definierbar („Beweisbarkeit und Wahrheit stimmen überein“). Man beobachte, dass zu jedem x aus $\mathcal{L}$ die Aussagen „x ist in W“ und „x ist nicht in W“ ausdrückbar sind und genau eine davon beweisbar ist. Die Idee Tarskis war es, eine Aussage zu konstruieren, die von sich selbst behauptet, nicht wahr zu sein (Lügnerparadoxon). Das gelingt ihm über zwei Aussagen a und b in $\mathcal{L}$, die folgenderweise eine Selbstreferenz induzieren:

$$a = \text{„}b \text{ ist in } W\text{“}$$
$$b = \text{„}a \text{ ist nicht in } B\text{“}$$

Ist b wahr? Unter der Bedingung $W = B$ entsteht ein Widerspruch. Aber es ist hier a priori unklar, wie man die Selbstreferenz innerhalb von $\mathcal{L}$ herstellen kann. Man müsste eine der beiden Aussagen syntaktisch unabhängig von der anderen definieren. Das gelingt mit der (laut Annahme wahren) Aussage

$$a = \text{„Für jedes } x \text{ aus } W \text{ gibt es einen Beweis, dass } x \text{ in } W \text{ ist“}$$

Aber sie gehört nicht unbedingt zur Objektsprache $\mathcal{L}$ (sie ist metasprachlich). Erst durch die Zugehörigkeit zu $\mathcal{L}$ wären a und b entweder beweisbar oder widerlegbar, und nur so erhielten wir einen Widerspruch. Die syntaktische Lösung (die auf Gödel zurückgeht) ist technisch kompliziert, aber semantisch klappt das leicht: Wir sprechen hier von einem Axiomensystem, also haben wir eine formale Sprache vor uns und sie hat eine wissenschaftliche Semantik. Aufgrund der semantischen Widerspruchsfreiheit existiert ein Modell. Dass W definierbar ist, bedeutet, dass man eine Aussage w angeben kann, sodass in jeder denkbaren Welt, in der w wahr ist, auch jede Aussage aus W wahr ist ($w \models W$), und umgekehrt ($W \models w$). Dann kann man sich überlegen, dass w semantisch genau die Aussage a von oben ist. Syntaktisch mag die Aussage anders aussehen, aber semantisch hat sie dieselbe Bedeutung. Das bedeutet weiter, dass auch die Aussage b in der Sprache $\mathcal{L}$ ist. Nach dem vorherigen Lemma ist die Aussage a nicht beweisbar und somit ist b wahr. □

Man beachte, dass $B = W$ gelten muss und der Beweisbegriff in der Sprache definierbar sein muss. Es wäre denkbar, dass in manchen Fällen nur eines von beiden erfüllt wird. Tarski fand mathematisch wichtige Beispiele dazu:

---

[i] *Anm. für Mathematiker.* Das folgern wir zwar aus dem Fundierungsaxiom, aber Tarskis Beweis gilt auch in ZF ohne Fundierung, die entscheidende Idee ist: Kommen in einer Sprache Variablen von Ordnung höchstens n vor, dann ist eine Aussage mit Variablen der Ordnung $n + 1$ nicht mehr in der Sprache.

[ii] Auch hier muss angenommen werden, dass W unendlich ist.

Beobachtung 17 (Tarski)**.** Es gibt Axiomensysteme der Kartesischen Geometrie, der reell abgeschlossenen Körper und der algebraisch abgeschlossenen Körper, in denen die Menge der mathematisch beweisbaren Sätze aus der Sprache mit der Menge der logisch beweisbaren Sätze übereinstimmt (innerhalb der Sprache stimmen $\vdash$ und $\models$ überein); aber der Beweisbegriff ist dort nicht definierbar.

Beobachtung 18 (Tarski)**.** Es gibt Axiomensysteme der natürlichen Zahlen, in denen sich der Beweisbegriff definieren lässt, W aber nicht.

Beweis. Das folgert Tarski aus einem Ergebnis von Gödel: Dieser hatte gezeigt, dass mit den Axiomen von Peano der Beweisbegriff $\vdash$ innerhalb der Objektsprache formalisierbar ist. Würde nun in der Objektsprache die Menge der beweisbaren Aussagen mit der Menge der wahren Aussagen übereinstimmen (würden Wahrheit und Beweisbarkeit übereinstimmen), könnte das Lügnerparadoxon aus dem Beweis zu Satz 16 formuliert werden und man erhielte einen Widerspruch.[i] □

In philosophischen Diskussionen ist oft das gemeint, wenn man vom ersten Unvollständigkeitssatz Gödels spricht. Auch Gödel sah darin den „wahren Grund“ für seinen Satz, aber inhaltlich sagt sein Satz etwas anderes aus. Wir wählten hier Tarskis Formulierung von Beobachtung 18 aus [Tar69, S. 421], weil sie den Unterschied verdeutlicht, aber Tarskis Satz 16 gilt allgemeiner:

Theorem 19 (Tarskis Undefinierbarkeit der Wahrheit in der Objektsprache)**.** *In einer Sprache mit wissenschaftlicher Semantik ist die objektsprachliche Wahrheit W nicht in der Objektsprache definierbar.*[ii]

Beweis. Jede Wahrheitsdefinition, die besagt, welche Sätze einer Sprache wahr sind, ist eine Menge von Sätzen, die logisch abgeschlossen ist, also eine Theorie T. Die Voraussetzung *Axiomensystem* in Satz 16 kann man daher ohne Weiteres durch *Theorie* ersetzen. In W sind jene Sätze der Sprache, die mathematisch $\models$ aus T folgen. Es ist also zu zeigen, dass kein Satz w in der Theorie existiert, sodass w genau dann wahr ist, wenn jeder Satz aus W wahr ist. Oder $W \neq T$. Falls w in der Sprache sagbar ist, ist w in W, weil $T \models w$. Falls man annimmt, dass w in T ist, lässt sich das Lügnerparadoxon wie im Beweis zu Satz 16 formulieren. Also $W \neq T$. □

Der Satz darf zweifellos als *Fundamentalsatz der Mathematikphilosophie* bezeichnet werden: Wahrheit ist nur metasprachlich definierbar.

1.2.2. *Die philosophische Bedeutung der Sätze Tarskis*

Tarski definierte nicht, was er unter einer formalen Sprache verstand, aber ein Blick auf die Beweise verrät, dass man seine Ergebnisse auf alle Sprachen mit einer wissenschaftlichen Semantik verallgemeinern kann, damit auch auf ihre Potenzmengen. Die Beweise verraten weiter, dass die Grenzen der Beweisbarkeit durch die Grenzen der Sprache determiniert werden, und diese wiederum durch die Unmöglichkeit entstehen, in einer gegebenen Sprache ihre Metasprache vollständig zu erfassen.[11]

[i] Nochmal: In $\mathcal{L}_1$ ist die Induktion unsagbar. In einer $\mathcal{L}_1$-Arithmetik ist der Beweisbegriff nicht formalisierbar.

[ii] Und wieder wird benötigt, dass W unendlich ist.

Eine häufige Fehldeutung der Ergebnisse Tarskis, die aber tatsächlich auf Tarski selbst zurückgeht, ist die, dass seine Wahrheitstheorie der aristotelischen Korrespondenztheorie entspräche. Es ist richtig, dass Tarski diese Theorie formalisiert, aber das Ergebnis ist eindeutig:[12] „Schnee ist weiß“ ist genau dann wahr, wenn eine Metasprache angegeben werden kann, in der Schnee weiß ist. Da dort der Wahrheitsbegriff für die Metasprache nicht definierbar ist, ist der metasprachliche Satz genau dann wahr, wenn eine Meta-Metasprache angegeben werden kann, in der Schnee weiß ist, usw., man springt von $\beth$ zu $\beth$ ohne jemals ein Ende erreichen zu können. Wahrheit einer Objektsprache ist nicht auf Objektebene, sondern auf Metaebene definierbar.

Wichtiger für unseren Zusammenhang ist das auf Seite 15 geschilderte Problem, das Tarski selbst „symptomatisch für eine gewisse Unvollkommenheit und Unvollständigkeit der bisher in den deduktiven Wissenschaften verwendeten Schlussregeln“ nennt ([Tar35, S. 176]). Es zeigt, dass in jeder Logik zahlreiche (für unsere Vernunft) triviale Wahrheiten gibt, die sich nicht sprachlich erfassen lassen: Nehmen wir irgendeine Theorie, die in einer Sprache mit einer wissenschaftlichen Semantik formuliert ist, und definieren wir eine unendliche Teilmenge wahrer Sätze, so ist es offenkundig wahr, dass jeder Satz aus dieser Menge wahr ist. Aber das ist eine metasprachliche Aussage und kann im Allgemeinen nicht innerhalb der Theorie bewiesen werden. Der Vorschlag, Beweise mit unendlich vielen Voraussetzungen zu erlauben, scheitert daran, dass die Logik dann semantisch unvollständig wird. Diese Beobachtung motiviert uns zur These: Wären die Grenzen der Sprache die Grenzen meiner Gedanken, wäre Mathematik undenkbar.

Die Sätze Skolems hatten gezeigt, dass dies, sofern wir uns an die Logik $\mathcal{L}_1$ halten, nicht zu vermeiden ist. Die Sätze Tarskis zeigen, dass die Logik keinen Einfluss hat, bereits die Beschränkung auf Sprachen mit einer wissenschaftlichen Semantik macht es unmöglich, Mathematik vollständig sprachlich zu erfassen. Dies relativiert nicht die Ergebnisse Skolems, es macht sie wichtiger, denn da es bei jeder Logik Grenzen gibt, wählen wir die beste aller möglichen Logiken, das ist $\mathcal{L}_1$, und sehen zu, dass wir mit den Ergebnissen Skolems zurechtkommen.

Blicken wir auf die Beth-Reihe, schmelzen die Ergebnisse zu einer neuen Erkenntnis: Ist ein $\mathcal{L}_1$-Axiomensystem gegeben, so benötigen wir Symbole aus $\beth_0$, können aber in $\mathcal{L}_1$ nicht Mengen von Individuen in Relation setzen und auch keine Relationen zwischen Individuen mit gewissen Beziehungen untereinander begrifflich fassen, dazu müssen wir in die Metaebene und benötigen Symbole aus $\beth_1$ usw. Interessant ist, dass die Wahrheit jeder Ebene sich auf der nächsten Ebene definieren lässt, also bleibt die Wahrheit vollständig auf der $\beth$-Reihe. Da wir auf jeder Ebene auf $\mathcal{L}_1$ angewiesen sind, werden wir von hier aus die Phänomene aus der unteren Ebene nicht sehen; können aber dafür Phänomene beweisen, die unten unbeweisbar waren. Wenn es dann zu einer Wechselwirkung zwischen Individuen und Mengen von Individuen kommt, sind wir von oben und von unten blind. Wir werden hier solche Sätze, die von einer unteren oder von einer oberen Ebene unbeweisbar sind, *Emergenzphänomene* nennen. Sie treten dann auf, wenn Individuen einer Theorie, die in gewissen Beziehungen zueinanderstehen, in einer anderen Theorie als Einheit betrachtet werden. Beschränken wir uns auf Sprachen mit einer wissenschaftlichen Semantik heißt das: Immer.

Und was jetzt? Emergenzphänomene zeigen, dass Logiken höherer Stufe trotz aller Unkenrufe doch nützlich sind und auch immer sein werden, nur sie können ein logisches Licht auf Emergenz-Wahrheiten werfen. Eine andere Idee sind *Antiaxiome*: Antiaxiome sind logisch nicht fassbar und können die Sprache transzendieren. Sie werfen aber trotzdem Licht. Ein unlogisches Licht, wenn man so will.

Antiaxiome

Antiaxiome nennen wir Wahrheiten der Form „aus *logisch unbeweisbar* folgt *mathematisch wahr*“. Inwiefern kann der Beweis der logischen *Un*beweisbarkeit einer Aussage a die Wahrheit von a beweisen? Sei eine Theorie gegeben, in der eine Funktion f existiert, sodass zu jedem Individuum x innerhalb der Theorie $f(x)$ logisch ableitbar ist. Dann gilt für jede Aussage der Form „für kein x ist $f(x)$ diesunddas“, dass aus ihrer logischen Unbeweisbarkeit ihre Wahrheit folgt. Dasselbe für eine Familie von Funktionen f_i, sodass $f_i(x)$ logisch ableitbar ist, und eine Aussage der Form „für kein Index i und kein Individuum x ist $f_i(x)$ diesunddas“. Ist nämlich die Unbeweisbarkeit bewiesen, heißt das, dass kein Gegenbeispiel in der Theorie nachgewiesen werden kann, woraus die Wahrheit der Aussage folgt (denk an Pascal, S. 13).

Beispiele muss man nicht lange suchen, gleich die drei wichtigsten offenen Probleme für das 21. Jahrhundert gehören dazu [Sma98]: Die Riemann-Hypothese, das **P-NP**-Problem, die τ-Vermutung. Beispiel: Eine Theorie der Menschheit sei gegeben, die erlaube, zu jedem Individuum die Verwandten festzustellen. Nehmen wir den Satz „in keiner Verwandtschaftsrelation gibt es Fixpunkte“ (ein x mit $f_i(x) = x$). Der Beweis hängt natürlich von der Beweiskraft der Theorie ab, aber: Falls der Satz unbeweisbar ist, ist er wahr. Denn wäre er falsch, gäbe es ein Gegenbeispiel, und das hätte man in der Theorie angeben können („Hava ist ihre eigene Mutter“). Die logische Unbeweisbarkeit impliziert hier die mathematische Wahrheit.

1.3. **Die Sätze Turings**

Turing hat den Begriff der axiomatischen und der logischen Beweisbarkeit restlos klären können. Daher dürfen seine Ergebnisse in keiner Diskussion über Logik fehlen. Wie fassen sie hier schnell zusammen.

1.3.1. *Die Grenzen der Berechenbarkeit*

Definitionen. Eine übliche Programmiersprache kennt die Anweisungen IF und WHILE, sie kann den Computer dazu bringen, ein beliebiges Speicherfeld zu lesen, das Feld mit 0 oder mit 1 zu überschreiben, den Wert auf dem Feld umzukehren oder den Wert stehenzulassen. Ein Programm, das nichts anderes tut, nennt man WHILE-*Programm*. Jede Aufgabe, die durch ein WHILE-Programm berechnet werden kann, heißt WHILE-*berechenbar*.

Alan Turing war der erste, der einen Computer und eine passende Programmiersprache entwarf. Jede Aufgabe, die durch eine Turing-Maschine berechnet werden kann, heißt *Turing-berechenbar*. Es stellte sich heraus, dass Turing-berechenbar und WHILE-berechenbar gleichbedeutend sind (die Elemente, die diese Mengen umfassen, sind dieselben). Man kann auch beweisen, dass kein Computer existiert, der mehr berechnen könnte. Auch die Computer, die heute entwickelt werden, werden keine Aufgabe berechnen können, die nicht von einer Turing-Maschine berechnet werden könnte (es geht nur um größere Effizienz).

Der Speicher wird als potentiell unendlich vorausgesetzt. Es ist aber wichtig, dass für Turing-Maschinen und WHILE-Programme das Alphabet endlich sein muss, ebenso müssen die Wörter endlich lang sein. Das sind genau die Bedingungen für die Definition formaler Sprachen, die wir bereits anführten. Da es hier nur noch um formale Sprachen gehen wird, lassen wir das „formal“ weg und schreiben nur „Sprache“ – so wie es in der Informatik heute üblich ist. Das bedeutet, hier geht es ausschließlich um $\beth_0$.

Wichtig für die Logik und Wissenschaftsphilosophie ist der Zusammenhang zu formalen Sprachen mit einer Typ-0-Grammatik:

Satz 20. *Jede formale Sprache, die man mit einer Turing-Maschine oder mit einem WHILE-Programm sukzessive ausgeben lassen kann, ist eine formale Sprache mit einer Typ-0-Grammatik. Umgekehrt lässt sich jede formale Sprache mit einer Typ-0-Grammatik durch eine Turing-Maschine oder ein WHILE-Programm sukzessive ausgeben.*

Mit anderen Worten: Eine formale Sprache $\mathcal{L}$ hat genau dann eine Grammatik vom Typ-0, wenn man ein Computerprogramm angeben kann, das hintereinander jedes Wort von $\mathcal{L}$ ausgibt (wenn die Sprache unendlich ist, endet das Programm nie, aber man soll beweisen können, dass jedes Wort der Sprache irgendwann ausgegeben wird).

Man kann zeigen, dass jede Programmiersprache (gedacht als die Menge aller Programme in der Sprache) eine Typ-0-Grammatik hat: Man lasse sich alle Programme der Länge n ausgeben, das ist programmierbar. Mit WHILE n von 1 bis unendlich laufen lassen – fertig. Was ist mit der deutschen Sprache? Der deutsche Wortschatz ist endlich, also ist Deutsch, als die Menge aller deutschen Sätze gedacht, eine formale Sprache. Wer wissen möchten, ob Deutsch eine Grammatik von Typ-0 hat, braucht nur zu überlegen, ob man ein Programm schreiben kann, das alle Sätze der deutschen Sprache hintereinander ausgibt.

Nun wird der Grund sichtbar, warum Axiomensysteme gerade so eine Grammatik brauchen: Haben wir irgendwelche Axiome und irgendeine Logik $\mathcal{L}$, dann suchen wir die Sätze, die sich aus den Axiomen mit den Schlussregeln aus $\mathcal{L}$ ableiten lassen, aber auch jeden Satz, der aus grammatischen Gründen „äquivalent“ ist. Nur wenn die Sprache eine Grammatik vom Typ-0 hat, können wir systematisch aus jedem Satz alle grammatisch äquivalenten Sätze gewinnen.

Der nächste Brocken ist nicht weit: die *Deduktionsmaschine*. Aus Satz 20 folgt, dass zu jedem endlichen Axiomensystem und jeder Logik mit endlich vielen Deduktionsregeln die sogenannte *Deduktionsmaschine* konstruierbar ist: Ein Programm, das aus den Axiomen sukzessive alle möglichen Folgerungen zieht. Wenn wir eine semantisch vollständige Logik wie $\mathcal{L}_1$ wählen, wird es möglich, eine Maschine zu konstruieren, die alle mathematischen Folgerungen herleitet. Wir müssen nur lange genug warten, jeder beweisbare $\mathcal{L}_1$-Satz wird irgendwann ausgegeben.

Denken wir um: Wir fragen nun, ob ein gegebener Satz aus einem Axiomensystem herleitbar ist. Man könnte die Deduktionsmaschine mit einer WHILE-Schleife laufen lassen, bis der Satz bewiesen wird. Was passiert, wenn der Satz unbeweisbar ist? Das Programm läuft unendlich lange. Ein definitives „Nein“ gibt es nicht, es gibt nur ein definitives „Ja“. Turing zeigte, dass dieses Ja-Nein-Entscheidungsproblem der Beweisbarkeit nicht besser gelöst werden kann.

Definitionen. Ein *Entscheidungsproblem* ist folgende Aufgabe: Gegeben sei eine Sprache $\mathcal{L}$ und ein Wort w, entscheide, ob w zur Sprache $\mathcal{L}$ gehört. Gibt es eine Turing-Maschine, die zu jedem Wort w korrekt das Entscheidungsproblem für die Sprache $\mathcal{L}$ löst, dann heißt $\mathcal{L}$ *entscheidbar*. Existiert eine solche Maschine nicht, heißt $\mathcal{L}$ *unentscheidbar*.

Auf den ersten Blick sieht man es der Definition vielleicht nicht an: Eine unentscheidbare Sprache ist eine Klasse von Problemen, die für Computer unlösbar sind. Was man der Definition mit Sicherheit nicht ansieht, ist, dass praktisch jedes Problem als Entscheidungsproblem formulierbar ist. Man muss nur flexibel die Sprache je nach Problem anders denken. Ein weitreichender Satz Turings besagt, dass Computer nur einen verschwindend kleinen Anteil von Problemen lösen können:

Folgerung 21 (Turing)*.* Es gibt abzählbar viele entscheidbare Sprachen, aber überabzählbar viele unentscheidbare Sprachen.

Beweis. Zuerst überlege man, dass nicht mehr WHILE-Programme geben kann als entscheidbare Sprachen. Zwar könnte ein Programm das Entscheidungsproblem für mehrere Sprachen lösen, aber man kann zu jeder entscheidbaren Sprache ein Programm programmieren, das nur für diese eine Sprache das Entscheidungsproblem löst. Also gibt es mindestens so viele entscheidbare Sprachen wie Programme. Als nächstes überlege man sich, dass es nur abzählbar viele WHILE-Programme gibt: Man kann alle Programme durchnummerieren. Jetzt kommt sozusagen routinemäßig Theorem 1 ins Spiel: Wörter (endliche Sequenzen) über ein vorgegebenes endliches Alphabet gibt es unendlich viele. Jede Teilmenge aus der Menge aller möglichen Wörtern über das Alphabet ist eine Sprache, also gibt es überabzählbar viele Sprachen. □

Ein wichtiges Beispiel, über das wohl schon jeder Programmierer gestolpert ist: Die Frage, ob die WHILE-Schleife terminiert oder das Programm niemals anhält, was dazu führt, dass das Programm nicht mehr reagiert oder gar dass der Computer „hängt“. Beim Programmieren gehört es zu den täglichen Aufgaben, sicherzustellen, dass jede WHILE-Schleife terminiert. Die Überprüfung darf man nicht ohne Weiteres dem Computer überlassen, wie der folgende Satz belegt.

Satz 22 (Turing)**.** *Das Halteproblem ist unentscheidbar.*

Beweis. Turing nutzt folgende Überlegung: Gäbe es einen Hellseher, würde ich ihn fragen, ob ich „Ja“ oder „Nein“ sagen werde, und das Gegenteil sagen. Angenommen, ein Programm existiert, das zu jedem Programm korrekt entscheidet, ob das Programm anhalten oder hängen wird. Nennen wir das Programm *Laplace*. Wir programmieren ein neues Programm mit dem Namen *Turing*. Immer dann, wenn Laplace „Ja“ sagt, gerät Turing in eine unendliche WHILE-Schleife, und falls Laplace „Nein“ sagt, gibt Turing „Ja“ aus und hält an. Nun prüfen wir, was passiert, wenn Turing nach seinem eigenen Verhalten gefragt wird. Falls Turing „Ja“ ausgibt, heißt das, dass Laplace meinte, Turing würde unendlich laufen. Also wäre Laplace im Irrtum gewesen. Daher muss Turing unendlich laufen. Aber das bedeutet, dass Laplace meinte, Turing würde „Ja“ sagen und anhalten. Auch hier hätte sich Laplace geirrt. □

In der Praxis wird die WHILE-Schleife so vereinfacht, dass man ihre Terminierung sicherstellen kann, indem man sie als FOR-Schleife schreibt. Aber man kann beweisen, dass Entscheidungsprobleme existieren, die nur mit WHILE-Schleifen lösbar sind. Die Frage nach der Beweisbarkeit in einem Axiomensystem gehört dazu, falls die Logik unentscheidbar ist: Dann muss jede Entscheidungssoftware mindestens eine WHILE-Schleife enthalten, von der man nicht wissen kann, ob sie terminiert, und falls sie terminiert, kann man nicht im Voraus wissen, wann sie terminiert. Nun können wir die eingangs auf Seite 7 geschilderte Situation verstehen, bei der eine Logik semantisch vollständig, aber unentscheidbar ist: Die Deduktionsmaschine liefert jede mathematisch korrekte Folgerung aus den Axiomen, aber stellen wir die Frage, ob ein bestimmter Satz aus den Axiomen folgt, gibt es kein definitives „Nein“, sondern nur nach ungewisser Zeit ein definitives „Ja“.

Turing gab zwei wichtige Beispiele an. Denken wir $\mathcal{L}_1$ als die Menge aller $\mathcal{L}_1$-Ausdrücke, die logisch immer wahr sind („Tautologien“) und fragen nach der Entscheidbarkeit dieser Sprache.

Satz 23 (Church, Turing)**.** *$\mathcal{L}_1$ ist unentscheidbar.*

Beweis. Turings Beweisidee ist folgende: Stellen wir uns vor, es liegen zwei Probleme A und B vor uns. Angenommen, wir können zeigen, dass Problem A unlösbar ist. Könnten wir in dieser Situation zeigen, dass Problem A lösbar wäre, wenn Problem B lösbar wäre, würden wir damit beweisen, dass auch Problem B unlösbar ist. Turings Beweis zeigt: Wäre $\mathcal{L}_1$ entscheidbar, so wäre das Halteproblem lösbar. Wir würden also „zu viel wissen“. □

Damit ist jede $\mathcal{L}_1$-Theorie (und also die gesamte Mathematik) unentscheidbar, da nicht entscheidbar ist, ob ein Satz eine tautologische Umformung (eine $\mathcal{L}_1$-äquivalente Umformung) eines anderen Satzes ist.

Aber Turings Denken ermöglicht auch andere Perspektiven: Sei die Sprache der Peano-Arithmetik die Menge der Peano-Axiome zusammen mit allen mathematischen Folgerungen aus ihnen, die sich in der Sprache der Axiome formulieren lassen. Ist diese Sprache entscheidbar?

Satz 24 (Turing)**.** *Die Peano-Arithmetik ist unentscheidbar.*

In Turings Beweis kommt keine Logik und keine Deduktionsmaschine vor. Der Beweis ist wieder: Wäre die Sprache entscheidbar, so auch das Halteproblem. Wir würden zu viel wissen.

1.3.2. *Die philosophische Bedeutung der Sätze Turings*

Die Relevanz der Sätze Turings für die Praxis und für die Philosophie kommt aufgrund der Bedeutung des Computers für die heutige Wissenschaft nicht unerwartet, aber sie ist doch tiefgreifender, als man zunächst meinen mag.

Der sogenannte *Satz von Rice* sorgt für eine allumfassende praktische Relevanz. Grob gesagt sagt er aus, dass semantische Eigenschaften von Computerprogrammen unzugänglich für Computerberechnungen sind. Aus ihm folgt zum Beispiel, dass für Computer die Überprüfung der Spezifikation eines Programms ein unlösbares Problem ist. Das heißt: Man kann nicht automatisch überprüfen lassen, ob ein Programm genau das leistet, wofür es programmiert wurde. Weil die Überprüfung für Menschen aufgrund der riesigen Datenmengen, die zu verarbeiten wären, praktisch unmöglich ist, behilft man sich damit, dass man auf gut Glück dem Programm vertraut und bloß nicht hinterfragt.

Die philosophische Bedeutung ist ernster. Man braucht sich nur zu fragen, welche die wissenschaftliche Methode ist, um neue Erkenntnisse zu produzieren. Die Methode ist die axiomatische Beweismethode:

(1) Ist die Aussage ein Axiom, ist sie bewiesen.
(2) Andernfalls zerlege die Aussage in Teilaussagen. Beweise jede Teilaussage.
(3) Nur wenn jede Teilaussage bewiesen wurde, ist die Aussage bewiesen.

Jeder Programmierer weiß: Das ist das Prinzip DC (*diviser chacune des difficultés … pour les mieux résoudre*) und es lässt sich unter gewissen Umständen als Rekursion schreiben und somit programmieren.[13] Theoretiker wissen auch: Alles, was sich überhaupt programmieren lässt, lässt sich mit dem Prinzip DC programmieren. Das Prinzip ist aber tatsächlich nur ein Prinzip und keine Methode, wie die Menschheit lange glaubte, denn aus den Beweisen von Turing geht hervor: Die Frage, ob eine Aussage beweisbar ist, muss nicht in jedem Axiomensystem entscheidbar sein, die Frage ist gemäß Folgerung 21 und Satz 23 meistens unentscheidbar – obwohl das Deduzieren methodisch machbar ist. Das ist ein fundamentales Ergebnis für die Wissenschaftsphilosophie: Auch wenn das Finden axiomatischer Beweise und somit auch die Axiomatisierung eines Modells nicht Turing-berechenbar sind und somit keine allgemeingültige „wissenschaftliche Methode“ existiert, ist das Deduzieren aus den Axiomen Turing-berechenbar, daher können wissenschaftliche Erkenntnisse methodisch überprüft und kritisiert werden.

Es wäre übereilt zu folgern, dass Computer prinzipiell zur wissenschaftlichen Tätigkeit untauglich wären: Computer lösen heute viele Aufgaben, die eigentlich nicht Turing-berechenbar sind. Die Lösungen sind Approximationen oder nicht ganz gewiss. Und das ist bei menschlichen Wissenschaftlern auch so.[i] Trotzdem zeigen die Sätze Turings, dass bei der Entwicklung mathematischer Modelle und bei der Reduktion wissenschaftlicher Aussagen auf Begründungen eine logische Grenze existiert.

Schließlich zeigen die Ergebnisse Turings, dass Computer genau das können, was der logische Beweis kann. Sie sind unfähig, mathematische Folgerungen $\models$ zu ziehen, die nicht durch $\vdash$ deduzierbar sind. Tarskis Problem (S. 15) überfordert jeden Computer. Deswegen haben sie so große Schwierigkeiten mit $\mathbb{N}$ und erst recht mit ihrer Metaebene $\mathbb{R}$.

[i] Die Schwierigkeit für Menschen ist der Speicherbedarf. Wer das Prinzip DC per Hand ausführt, muss genaue Listen und hoch komplexe Übersichten führen, um den letzten Punkt überprüfen zu können. Auch heutige Computer kommen damit an ihre Grenzen: Der Speicherbedarf wächst oft exponentiell in der Anzahl der Selbstreferenzen. Das ist der Grund, warum kein erfahrener Programmierer das Prinzip DC ohne Not anwenden würde.

1.4. **Die Unvollständigkeitssätze Gödels**

Wir verdanken Jef Paris zwei anschauliche Beispiele für den ersten Unvollständigkeitssatz Gödels. Eines davon wollen präsentieren: Die Unbeweisbarkeit des Goodsteinsatzes innerhalb der Peano-Arithmetik. Es geht wieder, nach den Ergebnissen von Tarski wenig überraschend, um eine metasprachliche Aussage, die sich aber als rein arithmetische Aussage präsentiert.

Man kann jede Zahl x zur jeder Basis b darstellen:

$$x = \sum_{k=0}^{n-1} a_k \cdot b^k$$

Dabei ist n die Anzahl der Stellen in der Darstellung. Heute kennt jeder Zahlendarstellungen zur Basis 2, Basis 10 oder Basis 60.

$$\begin{aligned} 56 &= 1 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0 \\ &= 5 \cdot 10^1 + 6 \cdot 10^0 \\ &= 56 \cdot 60^0 \end{aligned}$$

Es ist klar: Wenn wir die Basis schrittweise erhöhen, so werden die Koeffizienten der höchsten Potenz immer kleiner, die Koeffizienten der anderen Potenzen immer größer und irgendwann bleibt nur noch $a_0 = x$ übrig, während alle anderen Koeffizienten 0 geworden sind. Wenn wir aber gleichzeitig b und x erhöhen, ist es nicht mehr klar, was stärker wirkt: Das Schwinden der Koeffizienten oder das Wachstum der x. Knobelaufgabe: Die Basis wird immer nur um 1 erhöht, wie schnell muss x wachsen, damit die Koeffizienten nicht verschwinden? Wir nehmen als Start eine Zahl x_0 und zu jedem x_k bestimmen den Nachfolger, indem wir x_k zur Basis $k+2$ darstellen und dort jede Basis $k+2$ durch $k+3$ ersetzen. Das nennen wir *Aufpusten*. Wer gewinnt das Rennen? Schaffen die aufgepusteten Zahlen den Weg gegen unendlich oder schwinden die Koeffizienten, bis nur noch a_0 übrig bleibt? Beispiel: $x_0 = 56$, die Binärdarstellung lesen wir von oben ab: 111000. Jetzt aufpusten. Was kommt? 111000 gelesen als Zahl zur Basis 3. Wieder aufpusten: 111000 gelesen als Zahl zur Basis 4. Das ist also die gesuchte Folge: Die Koeffizienten schwinden nicht, die Folge der aufgepusteten Zahlen wächst gerade so schnell, dass es nicht passiert. Nun ist klar, dass für jede langsamer wachsende Folge die Koeffizienten schwinden werden und schließlich nur a_0 übrig bleiben wird. Goodstein zieht nach jedem Aufpusten 1 ab. Irgendwann wird also nur noch a_0 übrig bleiben, das Aufpusten wirkt nicht mehr und das -1 sorgt dafür, dass wir irgendwann bei 0 landen. Ist das vollkommen klar? In der populärwissenschaftlichen Literatur wird das als eine höchst überraschende Aussage dargestellt, aber um den Goodstein-Satz im Kontext der hier präsentierten Metamathematik richtig einzuschätzen, ist es wichtig einzusehen, dass die -1 *selbstverständlich* gegen das Aufpusten gewinnt: Der jeweils höchste Koeffizient wird kontinuierlich kleiner, bis er verschwindet, das Aufpusten wirkt anfangs stark, aber mit dem Verschwinden der Koeffizienten wirkt es immer schwächer, bis es gar nicht mehr wirkt, die -1 wirkt immer gleich.

Goodstein nutzt in Wirklichkeit ein beschleunigtes Aufpusten, bei dem auch die Exponenten aufgepustet werden, aber dieselbe Überlegung zeigt, dass die Folge nach einem Schuss in die Höhe am Ende bei 0 landen wird – egal wie groß der Startwert sein mag. Für den Startwert $x_0 = 56$ erhalten wir in Abhängigkeit der Basis b die beschleunigt aufgepusteten Zahlen

$$1 \cdot b^{b^b+1} + 1 \cdot b^{b^b} + 1 \cdot b^{b+1} + 0 \cdot b^b + 0 \cdot b^1 + 0 \cdot b^0$$

Wenn wir also nach jedem Aufpusten 1 abziehen, werden zwar die Basen immer größer, aber die dazugehörigen Koeffizienten der Reihe nach immer kleiner: Während der führende Koeffizient kontinuierlich schwindet, können die anderen wachsen. Schließlich verschwinden alle.

Definition. Bei einer Goodstein-Folge ist der Nachfolger des Folgegliedes x_k die beschleunigt aufgepustete Zahl x_k minus 1. Jede natürliche Zahl kann als Startwert gewählt werden. Jeder Startwert definiert eine Goodstein-Folge.

Satz 25 (Satz von Goodstein)**.** *Jede Goodstein-Folge erreicht nach endlich vielen Schritten den Wert 0.*

Die eigentliche Überraschung ist:

Satz 26 (Kirby, Paris)**.** *Der Satz von Goodstein ist innerhalb der Peano-Arithmetik ausdrückbar, aber dort unbeweisbar.*[i]

Man kann für jede konkrete Goodstein-Folge allein mit den Peano-Axiomen beweisen, dass der Wert 0 erreicht wird. Aber dass es für alle gilt, ist bereits metasprachlich. Das bedeutet zwar noch nicht, dass der Satz unbeweisbar ist, das Beispiel der Goodstein-Folgen mit bloßem Aufpusten (statt beschleunigtem Aufpusten) zeigt das: Da kann man noch mit Induktion innerhalb der Peano-Arithmetik zeigen, dass jede Folge 0 erreicht. Aber mit Blick auf den Beweis von Theorem 14 ist

[i] Kein Beispiel für Antiaxiome: Der Goodstein-Satz hat die Form *für jedes i gibt es ein x mit $f_i(x) = 0$*, wenn wir unter $f_i(x)$ das x-te Folgeglied der Goodstein-Folge mit Startwert i verstehen. Knobelaufgabe: Was fehlt für einen antiaxiomatischen Beweis des Goodstein-Satzes? Siehe Seite 36.

die Unbeweisbarkeit dann doch nicht ganz überraschend. Das Besondere ist, dass zusätzlich der metasprachliche Satz in der Objektsprache ausdrückbar ist: Es liegt eine syntaktische Unvollständigkeit vor.

Der wesentliche Punkt in Gödels Resultat ist, dass für bestimmte formale Systeme mathematischer Sätze Sätze gibt, die sich innerhalb des Systems *ausdrücken* lassen, aber aus den Axiomen des Systems *nicht entscheiden* lassen: Sie sind im Axiomensystem weder beweisbar noch widerlegbar. So fasst Gödel die Bedeutung seiner Sätze in einem Brief an Zermelo zusammen.

1.4.1. *Die Grenzen der Peano-Arithmetik*

Satz 27 (Erster Unvollständigkeitssatz Gödels: Gödel, Turing, Rosser)**.** *Jedes syntaktisch widerspruchsfreie Axiomensystem, das stark genug ist, um die Peano-Arithmetik zu formalisieren, ist syntaktisch unvollständig.*

Beweis. Gödels Beweis ist technisch, er argumentiert stets auf syntaktischer Ebene, dadurch erreicht er, dass seine Argumente unabhängig von der Semantik gelten und somit weitgehend unabhängig von der Philosophie der Mathematik sind, die man gerade hegt (gerade das macht seine Ergebnisse so beliebt bei Philosophen). Ohne weitere Erklärung: Wenn ein Axiomensystem „gödelisieren“ kann und die Substitutionsfunktion in den Aussagenfunktionen durch ein FOR-Programm berechnet werden kann, das innerhalb des System repräsentierbar ist, dann kann man innerhalb des Systems syntaktisch „ich“ sagen. Außerdem kann man den Beweisbegriff formalisieren, sodass der Satz „ich bin im System unbeweisbar“ im System ausdrückbar ist. Falls nun das System widerspruchsfrei ist, kann der Satz im System weder beweisbar noch widerlegbar sein. □

Die Peano-Arithmetik hat drei wichtige Eigenschaften: Die Summe, die Multiplikation und das Induktionsaxiom sind dabei. Bei der Presburger-Arithmetik ist keine Multiplikation dabei, und wie Mojżesz Presburger zeigen konnte, kann sie syntaktisch vollständig axiomatisiert werden: Da greift der Satz von Gödel nicht zu. Dasselbe, wenn man nur die Addition oder nur das Induktionsaxiom auslässt.

Satz 28 (Zweiter Unvollständigkeitssatz Gödels: Gödel, Hilbert, Bernays)**.** *In jedem syntaktisch widerspruchsfreien Axiomensystem, das stark genug ist, um die Peano-Arithmetik zu formalisieren, ist die Widerspruchsfreiheit des Systems ausdrückbar, aber nicht beweisbar.*

Beweis. Das sieht aus wie eine unmittelbare Folgerung aus dem ersten Unvollständigkeitssatz. Das ist nicht so, man muss zeigen, dass der erste Unvollständigkeitssatz selbst und auch sein Beweis in der Sprache des Axiomensystems ausdrückbar sind. Erst dann ist die Folgerung mit rein syntaktischen Argumenten möglich. □

1.4.2. *Die philosophische Bedeutung der Sätze Gödels*

Da der zweite eine Folgerung aus dem ersten Unvollständigkeitssatz ist, würde es im Prinzip ausreichen, nur die Konsequenzen aus dem ersten zu überlegen. Aber es lohnt sich, die Sätze getrennt zu überdenken.[14]

Der erste Unvollständigkeitssatz

Wir halten zuerst die formalen Unterschiede zu vorherigen Sätzen fest, um die spezifischen philosophischen Implikationen aus dem Gödelschen Satz bestimmen zu können.

Skolem. Es gibt zwei wesentliche Unterschiede zwischen Folgerung 8 und Satz 27. Erstens ist bei Skolem die Logik fest vorgegeben, während die Ergebnisse von Gödel für jede beliebige Logik gelten; zweitens ist bei Gödel das Axiomensystem vorgegeben, während Skolems Satz für beliebige arithmetische Theorien gilt, auch für nicht axiomatisierbare und für entscheidbare. Zum Beispiel ist die erwähnte Presburger-Arithmetik ebenfalls betroffen.

Tarski. Der Unterschied zwischen Beobachtung 18 und Satz 27 ist in der hier gewählten Formulierung klar. Ganz formal: Gödels Satz ist eine rein syntaktische Aussage für gewisse Axiomensysteme, während Tarskis Satz eine rein semantische Aussage für Sprachen ist, die auch Theorien trifft, die nicht axiomatisierbar sind. Aus Tarski lässt sich allein die „semantische Variante“ zu Satz 27 ([Hof18, S. 202]) herleiten: Da die Begriffe „Wahrheit“ und „Beweisbarkeit“ in der Peano-Arithmetik nicht übereinstimmen und dies auch für alle ausdrucksstärkeren Systeme gilt, folgt: *Jedes semantisch widerspruchsfreie Axiomensystem, das stark genug ist, um die Peano-Arithmetik zu formalisieren, ist semantisch unvollständig*. Gödels Voraussetzung der syntaktischen Widerspruchsfreiheit lässt sich ohne Verlust durch die semantische Widerspruchsfreiheit ersetzen. Und da syntaktisch unvollständige Axiomensysteme immer semantisch unvollständig sind, ist der einzige Unterschied zu Gödel, dass Gödel spezifischer in der Folgerung ist: Bei ihm geht es um syntaktische Unvollständigkeit. Ein Blick auf Seite 6 verrät: Das ist der schlimmste Fall unter den mathematisch brauchbaren Fällen.

Turing. Auch aus Satz 24 lässt sich wie bei Tarski nur eine semantische Version herleiten. Wir tun hier aber Turing historischen Unrecht, da erst durch sein Ergebnis Gödels Sätze, die ursprünglich nur für bestimmte Axiomensysteme gedacht waren, auf jedes Axiomensystem, das die Peano-Axiome formalisieren kann, verallgemeinert werden konnte. Wenn Gregory Chaitin die Ergebnisse Turings als „wesentlich tiefgreifender“ als die Gödels nennt, dann meint er diese Zusammenhänge zwischen Berechenbarkeit und Axiomatik, die erst Turing offenlegte. Auch für die tägliche Praxis sind die Sätze Turings entscheidend. Gödels Satz offenbart woanders seine Wirkung.

Gödel. Aus dem ersten Unvollständigkeitssatz Gödels lassen sich zwar schwächere Versionen von Folgerung 8 und Beobachtung 18 beweisen, aber da Gödels Satz sich nur auf gewisse Axiomensysteme bezieht, verlieren sie die nötige Allgemeinheit, um die wesentlichen Folgerungen in Theorem 10 und Satz 16 zu ziehen. Daher wäre es falsch, die dort gezeigten philosophischen Implikationen den Gödelschen Unvollständigkeitssätzen zuzuschreiben.

Umgekehrt machen die Unabhängigkeit der gewählten Logik und die syntaktische Unvollständigkeit den Unterschied aus. Man denke an den Satz von Goodstein: Innerhalb der Arithmetik von Peano ausdrückbar, aber mit den Axiomen von Peano weder beweisbar noch widerlegbar. Erst in der Mengenlehre ist der Satz beweisbar. Solche Fälle sind bei Skolem nicht ausgeschlossen, aber auch nicht notwendig.

Gödel selbst gab mit seinem zweiten Unvollständigkeitssatz die Widerspruchsfreiheit als Beispiel an. Auch sie ist innerhalb der Peano-Arithmetik ausdrückbar, aber dort weder beweisbar noch widerlegbar. Und auch sie ist in der Mengenlehre beweisbar. Ginge es nach Tarski, könnten wir immer die Sprache durch einen Sprung in die Metaebene erweitern, um Wahrheiten der Objektsprache beweisen zu können.

Aber genau da wird es unbehaglich: Bei Gödel geht es doch um Sätze, die innerhalb der Sprache ausdrückbar sind – um sie zu beweisen, braucht man nicht die Sprache zu erweitern, man muss neue Axiome finden. In der Arithmetik kann man zur Mengenlehre greifen. Aber in der Mengenlehre gilt der erste Gödelsche Unvollständigkeitssatz weiterhin und dort gibt es keine Ober-Theorie, auf die man zugreifen könnte. In dieser Situation gibt es zwei Möglichkeiten:

Erstens könnte man eine Ober-Theorie zur Mengenlehre entwickeln. Darauf können wir nur mit einem Gefühl antworten, das aber sicherlich schon viele Mathematiker empfunden haben: Die Mengenlehre ist mächtig im Ausdruck (übrigens deutlich mächtiger als Aristoteles' Begriffslehre), mit ihr lassen sich sehr beeindruckende Begriffssysteme angeben und Strukturen sehr schön formalisieren. Aber mit diesen Begriffen lässt sich wenig beweisen. Wie die Ergebnisse Skolems zeigen, können selbst grundlegende Theoreme der Mengenlehre nicht mit $\mathcal{L}_1$ bewiesen werden, woraus Skolem folgerte, dass die Mengenlehre nur sehr beschränkt dazu beitragen werde, Neues zu beweisen. Eine Ober-Theorie wäre ausdrucksstärker, aber schwächer für Beweise.

Zweitens bleibt die Suche nach neuen Axiomen, um die Beweisstärke der Mengenlehre zu erhöhen. Wie schon die Existenz von Antiaxiomen zeigt, ist es durchaus sinnvoll von *wahren* und *falschen* Axiomen zu sprechen. Ist ein Satz durch ein Antiaxiom beweisbar und in der Sprache ausdrückbar, kann er als Axiom gelten (siehe Seite 36). Die Wahl der Axiome wird aber keine alleinige Frage von Wahrheit sein, da das System nach Satz 27 immer syntaktisch unvollständig sein wird: Nicht-mathematische Faktoren werden eine Rolle spielen müssen.

Diese Folgerungen zur Mathematikphilosophie, die übrigens Gödel selbst in [Göd44] und [Göd47] zog,[15] können durch keinen der hier dargelegten Sätze von Skolem, Tarski oder Turing untermauert werden, sie begründen sich einzig durch den Gödelschen Unvollständigkeitssatz.

Der zweite Unvollständigkeitssatz

Aus der semantischen Variante des ersten Unvollständigkeitssatzes folgt nicht der zweite Unvollständigkeitssatz, auch keine semantische Variante. Die Folgerungen, die wir im Folgenden ziehen, sind also ebenfalls spezifisch für die Gödelschen Unvollständigkeitssätze.

Wären $\vdash$ und $\models$ in der Peano-Arithmetik gleichbedeutend, wären beide Sätze Gödels falsch. Ausschlaggebend ist also wie bereits bei Skolem und Tarski, dass die mathematisch wahre Folgerung $\models$ etwas mehr umfasst als die axiomatische Herleitung $\vdash$. Wir erinnern an Pascal (S. 13), der zwei Beweisformen unterschied: *Manipulation von Symbolen nach festen Spielregeln* und Widerspruchsbeweise. Das Paar entspricht genau dem Paar $\vdash$ und $\models$. Um das einzusehen, rekapitulieren wir Tarskis Begriff der mathematisch wahren Folgerung: $x \models y$ genau dann, wenn in jeder denkbaren Welt, in der x wahr ist, auch y wahr ist. In der Praxis ist es nicht so, dass man alle denkbaren Welten in Gedanken durchgeht. Formulieren wir es um: $x \models y$ genau dann, wenn es undenkbar ist, dass x wahr und y nicht wahr ist. Wobei *undenkbar* für *widersprüchlich* steht. Das ist die praktisch relevante Formulierung. So lässt sich jeder mathematische Beweis als Widerspruchsbeweis formulieren: Angenommen, y sei falsch, falls dann ein Widerspruch zur Wahrheit von x entsteht, ist y eine Folgerung aus x. Das ist das Schema, in das sich jeder mathematische Beweis bringen lässt. Das löst Tarskis Problem (S. 15) und war historisch auch Tarskis Motivation für seine Definition. Diese Formulierung setzt aber voraus, dass die Annahmen ohne die Folgerung widerspruchsfrei sind. Wenn jemand behauptet, einen mathematischen Satz bewiesen zu haben, so behauptet er, dass die Annahmen widerspruchsfrei sind.

Nun gehen wir zum zweiten Unvollständigkeitssatz Gödels: In den Beweisen zu Theorem 19, Satz 22 und Satz 27 war die Selbstreferenz entscheidend („ich" ist im System sagbar). Im Beweis zu Satz 28 zeigte Gödel, wie das System durch „ich" auf sich selbst Bezug nehmen kann: „Ich bin widerspruchsfrei". Satz 28 besagt, dass dieser Satz im System ausdrückbar, aber nicht beweisbar ist. Die Aussage ist aber genau dann wahr, wenn das System widerspruchsfrei ist. Die Frage ist nur, was bedeutet das? Nehmen wir an, die Aussage wäre falsch. Dann ist das System widersprüchlich, das Axiomensystem wäre somit undenkbar, da ich[i] es aber denke, wäre mein Denken falsch. Umgekehrt heißt das: Falls die Aussage wahr ist, ist das System widerspruchsfrei und ich denke richtig. Ergo: Wenn ich behaupte, einen Satz mathematisch bewiesen zu haben, so behaupte ich, dass die Annahmen widerspruchsfrei sind. Dies bedeutet nach dem Satz von Gödel, dass mein Denken auch jenseits der Logik vollkommen zuverlässig Widersprüche erkennt.

Aufgrund dieser Überlegung wäre es falsch, Mathematik als eine Art (inhaltsleere) Logik oder eine Art Sprache aufzufassen, mit der bestimmte Inhalte ausdrückbar sind (vgl. [Göd53]). Mathematik ist eher eine Theorie, da sie mindestens eine Voraussetzung macht: *Meine Vernunft ist ein vollkommenes Maß für Falschheit.* Man beachte den Unterschied zwischen *meine Vernunft ist fehleranfällig* und *meine Vernunft ist unfähig*, hier geht es allein um die prinzipielle Fähigkeit. Man braucht nur zu überlegen, wie anders das Argumentieren in der Mathematik wäre, wenn Widerspruchsbeweise nicht zulässig wären, weil die Vernunft nicht fähig wäre, Widersprüche zuverlässig zu erkennen. Jeder Beweis müsste als Manipulation von Symbolen dargestellt werden.

Eine einfache Konsequenz: Wenn ich an eine Realität jenseits des Denkens[ii] glaube, sodass meine Wahrheit durch die Realität bedingt wird, muss meine Vernunft die Realität (im Prinzip) vollständig denken können, da sonst die Realität undenkbar und also widersprüchlich wäre; die Realität muss außerdem so sein, wie ich sie denke. Jeder, der behauptet, etwas mathematisch bewiesen zu haben, und dies auf die Realität anwendet, behauptet zugleich diese Vollkommenheit der eigenen Vernunft. Für Mathematiker ist das eine Trivialität, die nicht erst aus Satz 28 folgt, aber Vorsicht: Hier wird ein Selbstbild der Mathematik begründet.

Wie schon der erste Unvollständigkeitssatz hat also auch der zweite nur Implikationen für das Selbstverständnis der Mathematik. Sie können daher schwerlich Grenzen für Wissenschaftstheorien nach sich ziehen. Doch das untersuchen wir jetzt genauer.

---

[i] Die Philosophie der Mathematik ist notgedrungen eine Ich-Philosophie: Niemand soll gezwungen werden, mathematisch zu denken, niemand soll belehrt werden. Nur das Ich überlegt, was es heißt, richtig zu denken, um (subjektive) Wahrheiten zu finden. Die Vorschriften gelten einzig und allein für das Ich.

[ii] Wir definieren nicht den Begriff des Denkens. Historisch (z. Bp. Rousseau, Hume) hielten sich hartnäckig zwei Missverständnisse gegen die „analytische Methode“ (gemeint war die moderne Mathematik), daher der Hinweis: fühlen und denken sind eine untrennbare Einheit, ebenso beobachten und denken. Wenn ich etwas beweise, dann ist es ein „Bauchgefühl“, das mir die Richtigkeit des Beweises suggeriert. Ich habe nichts, womit ich Denken vom Fühlen oder Beobachten unterscheiden könnte.

2. Wissenschaftsphilosophie

In diesem Kapitel wollen wir einige Voraussetzungen für die Anwendbarkeit metamathematischer Sätze auf eine Wissenschaftstheorie explizit benennen. Es muss Folgendes gelten:

(1) Wissenschaftliche Begründungen müssen der Methode DC folgen (wie gezeigt, ist das keine Methode, aber das Deduzieren und damit die Überprüfung der Begründung ist methodisch). Andernfalls würden die stärkeren Waffen bestimmen, was wissenschaftlich ist.
(2) Wissenschaftliche Begriffe müssen nach den Axiomen der Mengenlehre definiert sein. Andernfalls wären die Begriffe nicht unterscheidbar (*distinct*), wissenschaftliche Aussage wären mehrdeutig. Wissenschaftliche Wahrheit wäre relativ zur Macht.
(3) Wissenschaftliche Theorien müssen widerspruchsfrei sein, insbesondere dürfen Begriffe nicht (inhalts-)leer sein (sie müssen *clair* sein).[i] Andernfalls wären mathematische Folgerungen nicht anwendbar.
(4) Jede wissenschaftliche Erkenntnis soll eine mathematisch korrekte Folgerung sein. Andernfalls stünde die Wissenschaft in Widerspruch zur Mathematik.
(5) Mathematik darf keine Wissenschaft sein. Das wäre ein circulus vitiosus.

Der Punkt 1 sorgt dafür, dass wissenschaftliche Theorien logisch aufgebaut werden müssen. Nur dann sind die hier genannten Ergebnisse relevant (Skolem, Turing). Punkt 2 gewährleistet, dass die wissenschaftliche Sprache eine wissenschaftliche Semantik hat (somit werden die Ergebnisse Tarskis anwendbar). Gemäß Punkt 2 und Punkt 3 entwickelt die Wissenschaft mathematische Modelle, die nicht nur durch eine mathematisch korrekte Begrifflichkeit (2) gegeben sind, sondern auch mathematische Begründungen (3) erlauben. So entstehen axiomatisierbare mathematische Theorien. Die Modelle der Theorie sind aber nicht die Realität (was auch immer das sein mag). Metamathematische Aussagen beziehen sich immer nur auf die Theorien und ihre Modelle, nicht auf die Realität. Daher muss Punkt 4 sicherstellen, dass keine

[i] Ist eine Definition widersprüchlich (*undenkbar*), dann hat sie kein Modell, sie definiert semantisch die leere Menge. Ist umgekehrt eine Menge leer, erfüllen ihre Elemente jede Eigenschaft. Das hängt mit der Argumentation durch Widerspruch zusammen: Kein Element der Menge führt zum Widerspruch. Formal stehen nichtleere Mengen für Begriffe, die *denkbar* sind: Es gibt eine denkbare Welt, in der sie nicht leer sind. Früher sagte man, die Objekte, die unter dem Begriff fallen, seien *klar vorstellbar*, daher die traditionelle Bezeichnung *clair*.

weitere Erkenntnisquelle vorliegt. Gäbe es sie, wäre die Widerspruchsfreiheit kein Kriterium mehr für wissenschaftliche Theorien. Nebenbei impliziert dies, dass jede Beobachtung eine theoretische Folgerung ist.[16] Während also wissenschaftliche Erkenntnisse nicht durch „theoriefreie“ Beobachtungen gewonnen werden können, können sie durch die Analyse mathematischer Theorien und Modelle entstehen.[17] Dies setzt zwar die Entwicklung solcher Modelle als Teil der wissenschaftlichen Tätigkeit voraus. Aber die *Wahl des Modells* ist keine rein mathematische Frage, sie könnte von Traditionen, technischen Möglichkeiten und möglichen Anwendungen abhängen.[18] Aus diesem Grund dürfen sich metamathematische Grenzen für die Wissenschaft nur auf die Grenzen bei der Analyse der Modelle beziehen.

Definition. Eine semantisch widerspruchsfreie Theorie mit den genannten 4 Bedingungen nennen wir *wissenschaftlich*: Sie ist eine Menge widerspruchsfreier Sätze, die in einer Sprache mit einer wissenschaftlichen Semantik formuliert sind sowie unter den logischen Operationen einer mathematisch korrekten Logik abgeschlossen sind.

Der Punkt 5 ist notwendig, weil die Gültigkeit der hier präsentierten Ergebnisse naturgemäß davon abhängt, wie Mathematik definiert wird. Die Sätze sind im Allgemeinen weder im Logizismus noch im Intuitionismus und auch nicht im Formalismus gültig. Sie setzen eine bestimmte Mathematikphilosophie voraus, die nicht durch etwaige Wissenschaftstheorien annulliert werden darf. Es wäre ja analog unsinnig, wissenschaftstheoretische Einsichten etwa durch biologische Erkenntnisse widerlegen zu wollen. Implikationen können nur in der anderen Richtung existieren.

2.1. Gödel und die Wissenschaftsphilosophie

Punkt 5 macht es möglich, dass einige metamathematische Ergebnisse, die allein die Mathematikphilosophie betreffen, nichts für Wissenschaftstheorien implizieren. Wie im Abschnitt 1.4.2 ab Seite 29 gezeigt, wirken sich die Gödelschen Unvollständigkeitssätze direkt nur auf die Mathematikphilosophie aus. Ihre Bedeutung für Wissenschaftstheorien veranschaulichen wir mit einem Paradoxon.

Das TAKA-Paradoxon

Thoralf Skolem gab eine Methode an, die heute *skolemisieren* genannt wird; zusammen mit Alfred Tarskis Definition der mathematischen Wahrheit führt sie zu einem scheinbaren Widerspruch zu Kurt Gödels Satz 27 und Alan Turings Satz 23. Den Vornamen in historischer Reihenfolge entsprechend wollen wir es TAKA-Paradoxon nennen.

Definition. Seien T und K Theorien mit Modellen $\mathcal{T}$ und $\mathcal{K}$. Eine *Modell-Einbettung* von T in K ist eine Funktion, die zu jedem Satz a in der Sprache aus T einen Satz α in der Sprache aus K liefert, sodass gilt: a ist wahr in $\mathcal{T}$, falls α wahr in $\mathcal{K}$ ist.

Man erhält eine Art Orakel: Will man wissen, ob a wahr in $\mathcal{T}$ ist, fragt man die Einbettung. Natürlich kann jedes T in sich selbst eingebettet werden, aber das wäre langweilig. Interessant wäre, wenn man in $\mathcal{K}$ die Wahrheit von Sätzen durch Antiaxiome beweisen könnte:

Satz 29. *Zu jeder semantisch widerspruchsfreien $\mathcal{L}_1$-Theorie T kann systematisch eine Theorie K und eine Modell-Einbettung von T in K gefunden werden, sodass für jeden Satz α in der Sprache von K gilt: Ist α in K logisch unbeweisbar, dann ist α mathematisch wahr in $\mathcal{K}$. Dabei ist $\mathcal{K}$ das Modell von K.*

TAKA-Paradoxon. Du untersuchst ein Modell $\mathcal{T}$ und hast die Theorie T dazu. Findest du einen in T unentscheidbaren Satz a, gehe zu K über: Ist die eingebettete Aussage α in K unentscheidbar, dann ist a wahr in $\mathcal{T}$. Ist α eine logische Folgerung aus K, dann ist a wahr in $\mathcal{T}$, ist die Negation von α eine logische Folgerung, dann ist die Negation von a wahr in $\mathcal{T}$. Mit anderen Worten: Man kann jede Theorie so vervollständigen, dass die Wahrheit jeder Aussage entscheidbar wird.

Beweis. Die Idee ist, die Theorie so zu verändern, dass jeder Satz genau die Form erhält, auf die Antiaxiome anwendbar sind. Sei eine Aussage a gegeben, bringe a in Pränex-Form. Dies ist eine zu a äquivalente Aussage, bei der alle ontologischen Bedingungen vorne stehen („für alle“, „es gibt ein“). Diese Umformung kann systematisch erfolgen. Bringe dann die Aussage in Skolem-Form. Dies erfolgt durch ein systematisches Ersetzen von jedem „es gibt“ durch Funktionen („Skolem-Funktionen“). Die Idee ist so trivial wie genial:

$$\text{Für alle } x \text{ gibt es ein } y \text{, sodass } f(x, y)$$

wird ersetzt durch

$$\text{Es gibt eine Funktion } g \text{, sodass für alle } x \text{ gilt } f(x, g(x))$$

Wobei f eine beliebige Relation ist. Dann erweitert man die Sprache, indem man das Symbol g hinzufügt und man verändert die Theorie, indem man zu den Axiomen die Existenz der Skolem-Funktion g hinzufügt. Dies liefert eine Aussage α und eine Theorie K mit den gewünschten Eigenschaften. □

Das heißt *skolemisieren* und wird heute in fast jedem Logik-Kurs behandelt, weil es für KI häufig eingesetzt wird, zum Beispiel bei automatischen Theorem-Beweisern. Eine Aussage in Skolem-Form hat die Form „Für alle x_0, für alle x_1, für alle $x_2, \ldots$ gilt R“, wobei R eine Relation zwischen Individuen ist. In $\mathcal{L}_1$ ist das dasselbe wie „Es gibt kein x_0, kein x_1, kein $x_2, \ldots$, die nicht in Relation R stehen“. Damit ist eine Aussage in Skolem-Form durch Angabe konkreter Individuen widerlegbar.

Dazu kann irgendein $\mathcal{L}_1$-Algorithmus laufen (z. Bp. der Resolutionskalkül). Die Frage nach der Widerlegbarkeit könnte zwar unentscheidbar sein, aber da $\mathcal{L}_1$ semantisch vollständig ist, impliziert der Beweis der Unentscheidbarkeit erneut die Wahrheit der Aussage: Kein Gegenbeispiel kann in diesem Universum angegeben werden. Wann immer du in der skolemisierten Theorie auf Unentscheidbarkeit triffst, kannst du eine Entscheidung fällen.

Falls du aufmerksam gelesen hast, ist dir die Lösung des Paradoxon sicher klar: In einem Modell ist jede Aussage entweder wahr oder falsch.

Fazit

Gödels erster Unvollständigkeitssatz brachte alte Vorstellungen zu Fall, man glaubte an eine Allmacht der Logik und Axiomatik, und Gödel zeigte ihre Ohnmacht. Aber das bedeutet nicht, dass die Mathematik ohnmächtig wäre. Die Logik ist nur ein Hilfsmittel, und es gibt Wege, logisch-axiomatische Unentscheidbarkeiten zu umgehen.

Blicken wir auf die Wissenschaft: Unter der Annahme, dass die Wissenschaft mit mathematischen Modellen arbeitet, folgt unmittelbar, dass syntaktische Unvollständigkeit irrelevant ist, es geht um Modell-Vollständigkeit, und die erhält man durch Skolemisieren und Antiaxiome, oder andere Methoden. Sollte die syntaktische Unvollständigkeit Auswirkungen auf die Analyse eines mathematischen Modells in einer Wissenschaft haben, wäre die Mathematik gefragt, neue Axiome zu suchen, um die syntaktische Unentscheidbarkeit zu beheben. Wie bereits angemerkt, ist die Wahl des Modells keine mathematische Entscheidung. Nach Schließung der Lücke würde der Vorfall folgenlos bleiben. Es wäre sicherlich eine äußerst denkwürdige Angelegenheit für die Mathematik, hätte aber für die Wissenschaftsphilosophie keine Bedeutung.

Gödels zweiter Unvollständigkeitssatz hinterlässt dagegen schon eine Spur in der Wissenschaftsphilosophie. Wie wir sahen, impliziert er, dass bei der Anwendung mathematischer Theorien die Vollkommenheit der Vernunft beim Erkennen von Widersprüchen vorausgesetzt werden muss. Daraus folgten auch die Voraussetzungen, dass die Welt so ist, wie ich sie denke und sie mathematisch auch vollständig denkbar ist. Natürlich könnte ein böser Dämon existieren, der mich immerfort täuscht und mich im Irrglauben leben lässt, ich verstünde die Welt. Die Wissenschaft muss das ablehnen. Man könnte von metaphysischen Voraussetzungen sprechen; aber falls man keine Denkverbote mag, kann man sie auch als heuristische Prinzipien verstehen: Angenommen, meine Vernunft ist vollkommen, die Welt ist so, wie ich sie denke, und ich kann die Welt im Denken vollständig erfassen, wie weit komme ich damit?

2.2. **Die Sätze Turings und unser Weltbild**

Turing-Maschinen definieren unmittelbar einen Zeit-Begriff. Jede physikalische Theorie, die einen dazu kompatiblen Zeitbegriff hat und die physikalische Konstruktion von Turing-Maschinen ermöglicht (oder andere dazu äquivalente physikalische Systeme wie Zellautomaten), sieht sich den Ergebnissen Turings konfrontiert. Doch Vorsicht ist geboten.

Es wäre ein Missverständnis zu folgern, die physikalischen Zustände eines Computers seien als Funktion der Zeit nicht berechenbar. Nein, man kann für jeden Zeitpunkt vorausberechnen, ob ein Computer arbeiten wird oder sich in Ruhezustand befinden wird; was unmöglich ist, ist die Voraussage, ob er jemals den Ruhezustand erreichen wird.

Es wäre auch falsch, hieraus zu folgern, dass (physikalische Konstruierbarkeit der Turing-Maschine vorausgesetzt) das Eintreten gewisser Zustände nicht voraussehbar ist. Turings Beweis funktioniert nur, weil eine Turing-Maschine das Verhalten einer anderen Turing-Maschine voraussehen soll. Mathematisch lässt sich oft beweisen, ob ein Computer anhalten wird, obwohl kein Computer das vorausberechnen könnte. Eine Turing-Maschine ist ein Axiomensystem mit einer Logik – also ist sie unfähig, alle Wahrheiten zu erfassen.

Damit hängt auch das größte Missverständnis zusammen: der Begriff der Berechenbarkeit. Aus gutem Grund hat sich für *Berechnung* oder *Methode* die Berechenbarkeit durch eine Turing-Maschine durchgesetzt: Sprache, Axiomatik, Logik und Berechenbarkeit schmelzen zu einer Einheit. Das ist für uns Kinder Platons und Aristoteles' hübsch, aber es gibt andere Rechenmodelle, die physikalisch umsetzbar sind (am bekanntesten [Sie12]) oder einfach nur mathematisch sinnvoller sind wie die Modelle der algebraischen Komplexitätstheorie, zum Beispiel das BSS-Modell aus [Blu+98]. Mit ihnen kann man Dinge berechnen, die mit Computern unberechenbar sind. Die These von Church, falls du sie kennst, gehört besser in das Fach „falsch“ abgelagert.

Turings Beweis zu Satz 22 verdient daher mehr Beachtung. Er ist auf fast jeden Berechenbarkeitsbegriff anwendbar, die berechnende Maschine muss nur andere Maschinen ansprechen („du“) und nach Vorhersage ihres Verhaltens fragen können (dann ist Turings Dämon in Abhängigkeit von Laplace' Dämon definierbar). Also gilt sein Beweis auch für eine mathematisch denkende Vernunft, die jede Sprache transzendiert. David Wolpert folgert in der hervorragenden Arbeit [Wol17]: Im Universum kann höchstens ein berechnender Dämon existieren, der das Verhalten aller anderen berechnenden Dämonen voraussagt. Wolpert nennt es etwas ironisch *Monotheismus-Theorem*, obwohl der Beweis nur für Dämonen innerhalb des Universums gedacht ist.

3. Reduktionismus

3.1. Reduzierbarkeit auf logische Systeme

Die Sätze Skolems zeigen, dass gewisse Begriffe wie die Identität oder die Endlichkeit nicht logisch definierbar sind. Daher ist die alte Vorstellung, dass wissenschaftliche Theorien einige metaphysische Axiome, einige empirische Axiome und sonst nur logische Deduktionen aus diesen Axiomen enthalten, nicht haltbar. Sie enthalten wie jede mathematisch formulierte Theorie darüber hinaus noch einige unsagbare Voraussetzungen. Man kann jede Theorie der Astronomie in jede beliebige Unendlichkeit erweitern, es wird niemals genug sein, um logisch beweisen zu können, ob das Universum endlich ist, oder wenn unendlich, welcher Mächtigkeit: Wenn es unendlich sein soll, kann man Theorem 10 anwenden, wenn es endlich sein soll, wäre gemäß Folgerung 7 logisch möglich, *endlich* als *unendlich* zu lesen.

3.2. Reduzierbarkeit auf sprachliche Aussagen

Unabhängig von den Ergebnissen Skolems impliziert jede mathematische Theorie Wahrheiten, die jenseits der logischen Beweisbarkeit liegen, weil sie gemäß den Ergebnissen Tarskis jede wissenschaftliche Sprache transzendieren. Wenn ich glaubte, die Grenzen meiner Sprache wären die Grenzen meines Wissens, würde ich gemäß Theorem 14 an unerklärliche Wunder glauben müssen, die meiner Vernunft unzugänglich sind. Dadurch wäre ich gezwungen, Widerspruchsbeweise abzulehnen; die heutige Wissenschaft, in der jede Entdeckung und jede Messung ein Widerspruchsbeweis ist, müsste ich als Wahnsinn ablehnen – oder als unerklärliches Wunder hinnehmen.

3.3. Emergenzphänomene

Die Sätze Tarskis und Skolems zeigen, dass die Wissenschaft (sofern sie mit mathematischen Modellen arbeitet) keine Universalsprache sein kann, in der alle Phänomene ausdrückbar sind. Zwei Szenarien sind denkbar: Jedes mathematische Modell (jede Wissenschaft) steht für sich allein und ist unabhängig von den anderen. Dann könnte es durchaus sein, dass die Aussagen der Metaebene und somit die Emergenzphänomene wenig Relevanz haben. Das ist der Weg, den man anfangs für die KI vorsah (aus Not, weil hinter KI Turing-Maschinen stehen). Oder die Modelle hängen voneinander ab. Dann müssten sie als Hierarchie von Theorien entlang der Beth-Reihe denkbar sein. Die verschiedenen Theorien müssten nur untereinander widerspruchsfrei sein. Das ist der Weg der modernen Wissenschaft.

Das zweite Szenario lädt zum Reduktionismus ein: Einheitswissenschaft. Sofern damit der *Baum der Wissenschaften* gemeint ist, bei dem die Mathematik die Wurzel bildet, gefolgt von Physik und Chemie im Stamm usw., ist dagegen nichts einzuwenden. Wer aber die Ergebnisse Skolems, Henkins (S. 14) und Tarskis (S. 20) verkennt, könnte glauben, die Einheitswissenschaft sei das ersehnte Ziel der Wissenschaft. Genährt wird diese Ansicht durch die Wissenschaftsgeschichte: Während Aristoteles scharfe Trennungen zwischen den Wissenschaften voraussetzte (kein arithmetischer Satz könne geometrisch bewiesen werden und umgekehrt), waren gerade die Grenzübertretungen fruchtbar. Die Gegner der Einheitswissenschaft waren stets machtvoll, in Philosophie und Politik, aber es waren die Befürworter, die die Wissenschaft vorantrieben, oft von ihrer Zeit verschmäht oder ignoriert. Und trotzdem: Nein, die Einheitswissenschaft würde in immer höhere Ordnungen von Mengen von Mengen aufsteigen und damit hochgradig unvollständig sein. Es ist richtig, dass man die verschiedenen Wissenschaften widerspruchsfrei entlang der Beth-Reihe anordnen können muss, es ist richtig, dass ihre Grenzbereiche mit Logiken höherer Ordnung studiert werden können, aber das ist nicht das mathematische Modell, das wissenschaftlich im Fokus stehen darf, weil Logiken höherer Stufe semantisch unvollständig sind und die Emergenzphänomene vorherrschend bleiben – bereits beim ersten Übergang entstehen überabzählbar viele, aber laut Lemma 11 gibt es nur abzählbarviele Sätze, die sich mit irgendeiner Logik beweisen lassen. Man muss vielmehr jede Stufe als $\mathcal{L}_1$-Theorie neu formulieren und Emergenzen anders angehen. So ist der Reduktionismus auch als Heuristik zum Scheitern verurteilt, trotz seiner erfolgreichen Vergangenheit.

Es könnte eine historische Parallele zur Mathematik entstehen: Die Verknüpfung von Arithmetik und Geometrie in der analytischen Geometrie erwies sich als fruchtbar. Die Verknüpfung dieser mit der Algebra in der linearen Algebra ebenfalls. Alle mathematischen Zweige sind heute in der Mengenlehre vereint, aber die Mengenlehre selbst ist kaum Gegenstand der Forschung, mittlerweile wird sie an kaum einer Universität gelehrt, denn sie ist hochgradig unvollständig. Es ist zu erwarten, dass die Einheitswissenschaft ein ähnliches Schicksal erfährt.

3.4. **Laplacescher Dämon**

Nur informell: Die Kolmogorov-Komplexität von a ist gegeben durch die Größe des kleinsten Programms, dass a ausgibt (a wird also maximal komprimiert). Dabei heißt a zufällig, falls das Programm nicht kleiner als a ist (die komprimierte Datei ist so groß wie die Datei). Die Informationen, die das Komprimierungsprogramm benötigt, um a wiederherzustellen, heißen Kolmogorov-Informationen.

Aus Sicht der Informationstheorie Kolmogorovs ist die Existenz des Laplaceschen Dämons außerhalb des Universums evident, sofern irgendwo im Universum ein kausaler Zusammenhang existiert, wenn also das Universums nicht vollständig zufällig im Sinne Kolmogorovs ist. Davon ausgehend sind *algorithmic theories of everything* trotz Turing (und Chaitin) denkbar, siehe [Sch02]. Doch diese Sicht setzt voraus, dass Kolmogorovs Komplexitätsbegriff anwendbar ist, der ist aber nur anwendbar, wenn alle Informationen in einer Sprache codierbar sind. Die Kolmogorov-Komplexität, zweifellos eine Sternstunde der Philosophie, ist also nicht auf die Beth-Reihe anwendbar, da auf jeder Stufe eine neue Sprache benötigt wird und die Sprachen nicht ineinander übersetzbar sein können. Somit ist sie auch nicht auf den *Baum der Wissenschaft* anwendbar.

Auch David Wolpert verwendet in [Wol17] Kolmogorovs Informationstheorie, um den Laplaceschen Dämon trotz Zufall im Universum und Unentscheidbarkeiten in den Theorien sinnvoll definieren zu können. Aber er macht es vorsichtiger, indem er Kolmogorovs Informationstheorie nur auf vereinzelte Theorien anwendet (statt auf Universaltheorien) und somit innerhalb einer Sprache bleibt. So kann ein Dämon ein Geflecht widerspruchsfreier Theorien entwickeln und je nach Phänomen, das er voraussagen möchte, eine Theorie wählen. Wolperts Ergebnis ist, dass ein solcher Laplacescher Dämon innerhalb des Universums nur dann existieren kann, wenn kein weiterer Dämon existiert, der ihn fragen kann (*Monotheismus-Theorem*). Der Beweis ist von der physikalischen Theorie unabhängig, wird also auch dann gelten, wenn alle heutigen physikalischen Theorien als falsch gelten werden.

Falls aber die Welt mathematisch modellierbar ist, braucht der Dämon $\mathcal{L}_1$. In jeder Theorie kann also der Dämon auf Turings Berechenbarkeitsbegriff beschränkt bleiben (alles, was berechenbar ist, ist auch logisch-axiomatisch in einer formalen Sprache ableitbar).

Wieder aufgrund der Emergenzphänomene für axiomatische Theorien entlang der $\beth$-Reihe kann der Laplacesche Dämon weder innerhalb noch außerhalb des Universums existieren. Unabhängig von der physikalischen Konstruierbarkeit von Turing-Maschinen.

Anmerkungen

1. Auch die für die Philosophie noch wichtigere Ergebnisse von Kolmogorov und Chaitin zur Informationstheorie sind Diagonalargumente. Auf die inhaltlichen Zusammenhänge des Diagonalarguments gehen wir nicht ein. Sie sind ausgehend von [Law69] untersucht worden, aber sie sind, so scheint es uns, noch nicht genügend verstanden.

2. *Anm. für Historiker.* Aristoteles formuliert die Trennung nicht ausdrücklich und es darf bezweifelt werden, ob er seine logische Konzeption tatsächlich so meinte. Aber die Interpretation setzte sich auch nicht völlig zu unrecht durch: Analytica Priora I.1 24a18-19, I.4 25b39-40 und I.4 26a27 lassen keinen Zweifel daran, dass für ihn die logischen Operationen unabhängig von der Semantik gelten.

3. *Anm. für Historiker.* In 28a26 beim *Felapton* setzt er voraus, dass Begriffe nicht leer sein dürfen. Dies war für ihn und seine Zeit selbstverständlich, ist aber eine ontologische Annahme. Es ergeben sich zahlreiche „Folgefehler". Das Ziel, einen Begriff der logischen Folgerung zu finden, der von ontologischen Annahmen unabhängig ist, wurde somit verfehlt.

4. Wir interpretieren hier die Logik des Aristoteles als Modallogik wie Michael Wolff, nicht als Monadenlogik wie Stanisław Jaśkowski.

5. *Anm. für Historiker.* Als Persönlichkeit ist Skolem schwer zu durchschauen, er hielt nicht viel von Hilbert, noch weniger vom Hilbertprogramm und noch viel weniger vom Ruhm, den er geradezu zu verabscheuen schien. Lange weigerte er sich, einen Doktortitel zu erlangen und mied stets die Zentren der Forschung. Doch Tatsache ist, dass die moderne Mathematik ihm vieles verdankt: Die Mengenlehre, so wie sie heute gelehrt wird, wird zwar nach Zermelo und Fraenkel benannt, ist aber in dieser Form allein von Skolem. Er leistete wichtige Beiträge im Kontext des Hilbertprogramms und löste auch ein Problem aus Hilberts Liste der Jahrhundertprobleme. Skolem verschwieg immer diesen Kontext und tat so, als ob er davon nichts wüsste. Das ist unglaubwürdig, Hilbert war das Zentrum der Mathematikgemeinde, er war eine Art Idol der jungen Generation, und Skolem hatte selbst ein Jahr bei Hilbert gelernt. Mit den folgenden Sätzen ist es ähnlich: Man findet sie alle bei Skolem ([Sko20],[Sko22],[Sko29],[Sko34]), da sie aber das Hilbertprogramm betreffen, sagt Skolem nicht, was er da beweist, oder wenn er es sagt, fügt er hinzu, den Beweis habe er von anderen gehört oder bei anderen gelesen. Wir möchten hier diese Sätze trotzdem Skolem zuschreiben.

6. Von Giwi Margwelaschwili entliehen, der zu literarischen Texten und anderen Kunstwerken systematisch Nicht-Standard-Interpretationen entwickelt.

7. *Anm. für Historiker.* Das ist nicht Skolems Beweis. Die Existenz abzählbarer Nicht-Standard-Modelle der Arithmetik bewies Skolem in [Sko22], aber da dies für ihn ein „a priori sehr plausibles Ergebnis" war, machte er sich nicht die Mühe, ein konkretes Modell anzugeben. Er gab erst in [Sko34] ein konkretes Modell an. Das hier angegebene Modell stammt von Abraham Robinson. Dieselbe Beweisidee auf $\mathbb{R}$ angewendet ergibt die Existenz unendlichgroßer Zahlen, die wie jede reelle Zahl ungleich 0 auch multiplikative Inverse haben müssen, also unendlichkeine Zahlen. So konnte Robinson die Korrektheit der Analysis in ihrer Ursprünglichen Form (*Infinitesimalrechnung*) bei Leibniz und Euler nachweisen.

8. Levi ben Gerson und andere nutzten Induktion schon im Mittelalter. Aber es dauerte über Pascal hinaus weiter bis Abraham Kästner, bis sie als Beweisprinzip anerkannt wurde.

9. *Anm. für Historiker.* Viele Sätze scheinen heute trivial. Die Leistung Tarskis bestand vor allem darin, allgemein Sprachen zu betrachten, die Begriffe *Wahrheit*,

Metaebene und *Objektebene* genau zu definieren und davon ausgehend die Semantik formaler Sprachen zu bestimmen. Schließlich zu zeigen, dass Wahrheit überhaupt für Axiomensysteme definierbar ist (in [Tar35] bezeichnet Tarski dies als „Hauptergebnis A“, wir beziehen uns stets auf die Korrekturen ab S. 192). Erst mit dieser Begrifflichkeit sind die Sätze trivial.

10. Das war *keine* Revolution: Die Mengenlehre wurde von Cantor entwickelt und er meinte, mit dem Begriff „Menge“ das platonische *eidos* zu erfassen. Für das griechische Wort hatte Christian Wolf eine Neuschöpfung geprägt: *Begriff*. Cantor kam recht spät zur Bezeichnung „Menge“, er gebrauchte zuerst die Bezeichnung „Inbegriff“, dann „Mannigfaltigkeit“ und sprach auch von „Archetypen“. Auch bei Bernhard Bolzano heißen Mengen *Inbegriffe* (das ist die Verdeutschung von *Archetyp*). Die Warschauer Schule und die Schule um Luzin in der Sowjetunion pflegten diese philosophische Interpretation. Bekanntlich hat sie Bourbaki flächendeckend durchgesetzt. Aber nur in der Tarskischen Modelltheorie lässt sich rein formal beweisen, dass es so ist: Die Axiome der Mengenlehre und ihre Lehrsätze liefern eine Theorie korrekter Begriffsbildungen. Deswegen ist die Mengenlehre so wichtig.

11. *Anm. für Philosophen*. Die Interpretation der Ergebnisse Tarskis sorgen bis heute für Kontroversen. Das ist nicht nur inhaltlich bedingt, sondern zum Teil auch historisch: Undefinierte Begriffe und Missinterpretationen der eigenen Resultate erschweren die Lektüre der philosophischen Texte Tarskis. Er verstand sich nicht als Revolutionärer und deutete seine Ergebnisse oft (fälschlicherweise) als Bestätigung traditioneller Ansichten. Man merkt, wie seine Ansichten mit den Jahren kontinuierlich reifen, manchmal innerhalb einer Schrift wie in [Tar35], wo er im Nachwort seine Deutungen aus dem Haupttext umkrempelt. Wir präsentieren hier die moderne Interpretation seiner Sätze innerhalb der Modelltheorie. Jene Theorie, die Tarski selbst durch [Tar36] begründete.

12. *Anm. für Historiker*. Vgl. „Hauptergebnis A“ in [Tar35, S. 192]. Man beachte, dass Tarski oft von einer Metasprache, die ihre Objektsprache als Teilsprache enthält, ausgeht, während wir von einer Objektsprache sprechen, die Teile ihrer Metasprache formalisieren kann. Das ist formal äquivalent, erklärt aber, warum seine Formulierung von der unsrigen abweicht. In späteren Schriften wird er klarer und in [Tar69, S. 412-420] ist die Einsicht, dass Wahrheit sich nur auf Metaebene definieren lässt, das zentrale Argument.

13. Bei der ersten historisch belegten Formulierung im *Discours* wird als zusätzlicher Punkt die metasprachliche Annahme hinzugefügt, dass die so ermittelte Ordnung der Realität entspreche. Sie ist aber keine methodische Anweisung und wir übergehen sie stillschweigend. Man beachte, dass das Prinzip DC sich selbst aufruft und also „ich“ sagen kann, genauer: „ich kann beweisen, dass ...“.

14. Siehe [Fef06] für einen Überblick der heute weitgehend anerkannten Folgerungen aus den Unvollständigkeitssätzen.

15. *Anm. für Historiker*. Mit einem kleinen Unterschied im Kontext. Gödel schreibt: „Es gibt a priori drei Möglichkeiten für Cantors Kontinuumshypothese: Sie könnte beweisbar, widerlegbar oder unentscheidbar sein“ und sagt über die dritte Möglichkeit, sie sei eine „präzise Formulierung“ seiner These, dass die Frage nicht rein

mathematisch beantwortet werden kann [Göd47, S. 181, 259]. Gödel glaubte zeitweise an absolut unentscheidbare Sätze. Wir argumentieren hier dagegen über rein mathematisch wahre Aussagen, weil in einem Modell jede Aussage entweder wahr ist oder nicht: Ob das Auswahlaxiom wahr ist oder nicht, hängt nicht von den Axiomen ab, sondern von jener Mengenlehre in unserem Denken, die wir mit den Axiomen formalisieren wollen.

16. *Anm. für Philosophen*: Wenn ich an einem regnerischen Tag aus dem Fenster schaue und Menschen auf der Straße beobachte, dann sehe ich Hütte, Mäntel, Regenschirme und *folgere*, es seien Menschen da. Aber das gilt für jede Beobachtung: Die Beobachtung, dass ich mir den Finger verbrenne, setzt unter anderem eine Theorie des Ichs voraus und dass das Ich die von seinem Körper gesendeten Informationen lesen, richtig dechiffrieren und richtig zuordnen kann.

17. *Anm. für Philosophen*: Das klingt paradox, ist aber klar: Die Evolutionstheorie entstand nicht, als Darwin um die Welt reiste, sondern viele Jahre später, als er am Schreibtisch über seine Erlebnisse nachdachte und andere Forscher seine Theorie korrigierten, indem sie die theoretischen Folgerungen überprüften und sich zu Revisionen gezwungen sahen. Die treibende Kraft war stets die Widerspruchsfreiheit.

18. Die Wirbeltheorie erwies sich im 17. Jh. dem Modell der instantanen Fernwirkung mit Hilfe Gottes als unterlegen, wurde aber mit der Feldtheorie im 19. Jh. wieder aufgenommen und durch die Relativitätstheorie im 20. Jh. vollständig rehabilitiert. Die Wirbeltheorie implizierte nicht nur die Relativität des Raumes, sondern auch die Äquivalenz von Masse und Energie. Dies damals zu postulieren, wäre angesichts der technischen und begrifflichen Möglichkeiten praktisch sinnlos. Newtons Theorie eines jederzeit eingreifenden Gottes war damals gewiss sinnvoller und eroberte bald die gesamte wissenschaftliche Gemeinde.

Literatur

[Blu+98] Lenore Blum u. a. „Complexity and real computation“. Springer Science & Business Media, 1998.

[Fef06] Solomon Feferman. „The nature and significance of Gödel's incompleteness theorems“. In: *Princeton, Institute for Advanced Study, Gödel Centenary Program* 17 (2006).

[Göd44] Kurt Gödel. „Russell's mathematical logic“. In: *Collected Works* 2 (1944), S. 119–141.

[Göd47] Kurt Gödel. „What is Cantor's continuum problem?“ In: *Collected Works* 2 (1947, 1964), S. 176–187, 254–270.

[Göd53] Kurt Gödel. „Is mathematics syntax of language?“ In: *Collected Works* 3 (1953-9), S. 334–362.

[Göd84] Kurt Gödel. „Collected Works“. Hrsg. von Solomon Feferman. Oxford University Press, 1984-2003.

[Hof09] Dirk W. Hoffmann. „Theoretische Informatik“. Hanser-Verlag, 2009.

[Hof18] Dirk W. Hoffmann. „Grenzen der Mathematik“. Springer-Verlag, 2018.

[Law69] William Lawvere. „Diagonal arguments and cartesian closed categories“. In: *Category theory, homology theory and their applications II*. Springer, 1969, S. 134–145.

[Sch02] Jürgen Schmidhuber. „Hierarchies of generalized Kolmogorov complexities and nonenumerable universal measures computable in the limit“. In: *International Journal of Foundations of Computer Science* 13.04 (2002), S. 587–612.

[Sie12] Hava Siegelmann. „Neural networks and analog computation: Beyond the Turing limit“. Springer Science, 2012.

[Sko20] Thoralf Skolem. „Logisch-Kombinatorische Untersuchungen über die Erfüllbakeit und Beweisbarkeit mathematischen Sätze nebst einem Theoreme über dichte Mengen“. In: *Selected Works in Logic* (1920), S. 103–136.

[Sko22] Thoralf Skolem. „Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre“. In: *Selected Works in Logic* (1922), S. 137–152.

[Sko29] Thoralf Skolem. „Über einige Grundlagenfragen der Mathematik“. In: *Selected Works in Logic* (1929), S. 227–273.

[Sko34] Thoralf Skolem. „Über die Nichtcharakterisierbarkeit der Zahlenreihe mittels endlicher oder abzählbar unendlich vieler Aussagen mit ausschließlich Zahlenvariablen“. In: *Selected Works in Logic* (1934), S. 354–366.

[Sko62] Thoralf Skolem. „Set Theory“. Notre Dame, 1962.

[Sko70] Thoralf Skolem. „Selected Works in Logic“. Hrsg. von Jens Erik Fenstad. Universitetsforlaget, 1970.

[Sma98] Steve Smale. „Mathematical problems for the next century“. In: *The Mathematical Intelligencer* 20.2 (1998), S. 7–15.

[Tar35] Alfred Tarski. „Der Wahrheitsbegriff in den formalisierten Sprachen“. In: *Collected Papers* 2, 1935-1944 (1935), S. 186–198.

[Tar36] Alfred Tarski. „Über den Begriff der logischen Folgerung“. In: *Collected Papers* 2, 1935-1944 (1936), S. 271–281.

[Tar69] Alfred Tarski. „Truth and proof“. In: *Collected Papers* 4, 1928-1979 (1969), S. 401–423.

[Tar85] Alfred Tarski. „Collected Papers“. Hrsg. von Ralph McKenzie Steven Givant. Birkhäuser, 1985-1986.

[Wol17] David Wolpert. „Constraints on physical reality arising from a formalization of knowledge“. In: *arXiv:1711.03499* (2017).