Studienbücherei

P. Schreiber

Grundlagen der Mathematik

VEB Deutscher Verlag der Wissenschaften

Mathematik für Lehrer
Band 12

Herausgegeben von:
W. Engel, S. Brehmer, M. Schneider, H. Wussing

Unter Mitarbeit von:
G. Asser, J. Böhm, J. Flachsmeyer, G. Geise, T. Glocke, K. Härtig, G. Kasdorf, O. Krötenheerdt, H. Lugowski, P. H. Müller, G. Porath

Studienbücherei

Grundlagen der Mathematik

P. Schreiber

Mit 4 Abbildungen

Zweite Auflage

VEB Deutscher Verlag
der Wissenschaften
Berlin 1984

Verlagslektor: Brigitte Mai
Umschlaggestaltung: Rudolf Wendt

Lizenz-Nr. 206 · 435/86/84
Printed in the German Democratic Republic
Satz: VEB Druckhaus „Maxim Gorki", 7400 Altenburg
Fotomechanischer Nachdruck: Grafische Werke Zwickau, 9541 Zwickau
LSV 1014
Bestellnummer: 570 417 5
01980

Vorwort

Dem Anliegen der Reihe „Mathematik für Lehrer“ entsprechend ist das vorliegende Lehrbuch vor allem für die obligatorische Ausbildung der angehenden Mathematiklehrer im Fach Grundlagen der Mathematik bestimmt und lehnt sich daher inhaltlich eng an das präzisierte Lehrprogramm dieser Veranstaltung an. Ein Lehrbuch folgt jedoch stets anderen Gesetzen als eine Vorlesung, und so ist auch hier ein Buch entstanden, das bei aller prinzipiellen Beschränkung auf den Stoff der genannten Vorlesung und bei aller Wahrung seines einführenden Charakters als erste Quelle für weiterführende Studien dienen kann und in einigen Richtungen mehr und anderes bringt als alle mir bekannten Lehrbücher aus dem Bereich der Grundlagen der Mathematik.

Grundlagen der Mathematik im Sinne dieses Buches sind weder mit mathematischer Logik oder gar einem bestimmten Logikkalkül noch mit gewissen traditionellen Inhalten des Begriffs Grundlagen (etwa Mengenlehre, Aufbau der Zahlbereiche, Grundlagen der Geometrie und anderer spezieller mathematischer Disziplinen) zu identifizieren. Vielmehr wird eine möglichst elementare und durchsichtige Einführung in die semantischen (inhaltlichen), syntaktischen (formalen) und algorithmischen Aspekte der Metatheorie beliebiger deduktiver Theorien sowie in die gegenseitigen Beziehungen zwischen diesen Aspekten angestrebt. Im Gegensatz zu vielen Lehrbüchern ähnlicher Thematik wurden viele (für den Anfänger meist langweilige bzw. abschreckende) technische Einzelheiten von Beweisen zugunsten von ausführlichen Begriffsmotivationen, Diskussion von Beweismethoden, Beispielen und Anwendungen unterdrückt. Die letzten ergeben insbesondere insgesamt einen gewissen Aufschluß über die Grundlagen dreier spezieller mathematischer Theorien: der Mengenlehre, der Arithmetik der natürlichen Zahlen und der ebenen euklidischen Geometrie.

Die Abschnitte 4.7., 6.4., 7.4., 7.5., 8.3. und das Kapitel 5 können überschlagen werden, ohne daß der rote Faden des Buches verlorengeht. Sie tragen zum Teil (etwa 4.7. und Kapitel 5) reinen Beispielcharakter, zum Teil vermitteln sie Stoff, der über ein Minimalprogramm hinausgeht (etwa 6.4.), oder sie tragen gewissen im Lehrprogramm der Vorlesung Grundlagen der Mathematik vorgesehenen Alternativen Rechnung (zum Beispiel verschiedene Varianten der Einführung des präzisierten Berechenbarkeitsbegriffes). Ausdrücklich zu empfehlen ist ein solches abgekürztes Studium des Buches jedoch nicht.

Die wesentlichsten Unterschiede gegenüber anderen Einführungen in die Grundlagen der Mathematik sind

1. der Bezug auf im allgemeinen mehrsortige Strukturen und dementsprechend mehrsortige Sprachen von Anfang an,
2. eine sehr allgemeine Auffassung des Begriffs nichtelementare Sprache,
3. eine sehr allgemeine Auffassung des Begriffs definitorische Spracherweiterung,
4. die Zulassung partieller Operationen zur Interpretation von Operationssymbolen formalisierter Sprachen,
5. eine sehr allgemeine Auffassung der Begriffe Schlußregel und Beweiskalkül.

Die ersten drei dieser „Neuerungen" stehen miteinander in engem Zusammenhang, genauer: 3. erfordert 2., 2. erfordert 1. Die genannten fünf Schritte vom Gewohnten weg in Richtung auf größere Abstraktheit bzw. Allgemeinheit fordern, wie es dem Verfasser auf Grund langjähriger Erfahrungen erscheint, den „Fachmann", der bereits alles auf andere Weise kennt, mehr heraus als den Anfänger, der hier sowieso mit einer Fülle von gänzlich Neuem konfrontiert wird. Sie bringen letzten Endes sämtlich die in der Metamathematik untersuchten Strukturen und Sprachen näher an die in der mathematischen Praxis verwendeten heran und fördern dadurch die Bereitschaft des Unvoreingenommenen zum „Mitmachen".

Den Herren Professor H. Wussing und Dr. G. Kasdorf danke ich für einige berichtigende bzw. vertiefende Hinweise zum historischen Abschnitt 9. Ferner danke ich allen an der Herstellung beteiligten Kollegen des VEB Deutscher Verlag der Wissenschaften und des VEB Druckhaus „Maxim Gorki" für ihre sorgfältige Arbeit. Vor allem aber gilt mein Dank meinem verehrten Lehrer, Herrn Professor G. Asser, dessen prinzipielle Positionen zu den Grundlagen der Mathematik wesentlich in dieses Buch eingeflossen sind. Ich widmete es ihm in Dankbarkeit zu seinem 50. Geburtstag am 26. Februar 1976.

Stralsund, im Dezember 1976 Peter Schreiber

Inhalt

1. Zeichenreihen

In diesem Buch wird die Mathematik selbst zum Gegenstand mathematischer Untersuchungen gemacht. Das ist aber nur dadurch möglich, daß die in der Mathematik benutzten „Sprachen“ und die in diesen Sprachen formulierten „Theorien“ in recht formaler Weise als Mengen von geschriebenen (oder zumindest prinzipiell schreibbaren) Zeichenfolgen aufgefaßt werden. Demgemäß wird viel von solchen Begriffen wie Alphabet, Zeichenreihe, Vorkommen eines Zeichens in einer Zeichenreihe usw. die Rede sein. Dieses Kapitel führt in die wichtigsten Begriffe und Sachverhalte aus diesem Bereich ein, wobei wir uns bemühen werden, den Leser nicht schon im ersten Kapitel durch die unnötig abstrakte Behandlung anschaulicher Dinge abzuschrecken.

Unter einem *Alphabet* verstehen wir im folgenden stets eine beliebige nichtleere (nicht notwendig endliche) Menge. Die Elemente eines Alphabets A werden als *Buchstaben* oder auch als *Grundzeichen* bezeichnet. Eine *Zeichenreihe im Alphabet A* (im folgenden meist kürzer als *Wort im Alphabet A* bezeichnet) ist eine beliebige endliche Folge von Elementen von A. Ein Wort wird im allgemeinen durch einfaches Hintereinanderschreiben seiner Folgeglieder angegeben. Zum Beispiel sind abbab, ababc, cc, cab, b, c und ba Wörter im Alphabet $\{a, b, c\}$. Die Menge aller Wörter im Alphabet A bezeichnen wir mit $W(A)$. Ist A eine höchstens abzählbare (d. h. endliche oder abzählbare) Menge, so ist $W(A)$ stets abzählbar. Dies werden wir später beweisen.

Aus verschiedenen Gründen ist es zweckmäßig, in jede Menge $W(A)$ außer den oben betrachteten nichtleeren Wörtern ein *leeres Wort* aufzunehmen, das dadurch bestimmt ist, daß es keinen Buchstaben enthält bzw. eine Folge der Länge 0 ist. Das leere Wort wird (da man es offenbar nicht durch Hintereinanderschreiben seiner

Buchstaben notieren kann) mit dem Symbol Λ bezeichnet. Die Nützlichkeit des leeren Wortes sei an folgendem Beispiel erläutert: Es sei $\boldsymbol{A}$ ein Alphabet und $\mathrm{a} \in \boldsymbol{A}$ einer seiner Buchstaben. Dann kann man jedem Wort $\mathrm{W} \in W(\boldsymbol{A})$ ein eindeutig bestimmtes Wort $\mathrm{W}' \in W(\boldsymbol{A})$ zuordnen, das dadurch aus W entsteht, daß man alle in W vorkommenden Buchstaben a herausstreicht. Zum Beispiel ist $\mathrm{W}' = \mathrm{bb}$ für $\mathrm{W} = \mathrm{abab}$ und auch für $\mathrm{W} = \mathrm{bbaaaa}$, und allgemein ist $\mathrm{W}' = \mathrm{W}$ genau dann, wenn W den Buchstaben a nicht enthält. Ohne die Vereinbarung über das leere Wort wäre nun die Operation des Streichens von a an solchen Wörtern, die keinen von a verschiedenen Buchstaben enthalten (d. h. für $\mathrm{W} \in W(\{\mathrm{a}\}) \subseteq W(\boldsymbol{A})$), nicht ausführbar. Diesen Sachverhalt können wir jedoch unter Benutzung von Λ einfacher und kürzer formulieren: $\mathrm{W}' = \Lambda$ *genau dann, wenn* $\mathrm{W} \in W(\{\mathrm{a}\})$. Übrigens kann man das Vorhandensein von Λ in jeder Wortmenge $W(\boldsymbol{A})$ auch formaler dadurch rechtfertigen, daß 0 eine natürliche Zahl und daher jede Folge der Länge 0 eine endliche Folge ist. Somit kommt unter den anfangs als Elemente von $W(\boldsymbol{A})$ definierten endlichen Folgen von Buchstaben aus $\boldsymbol{A}$ automatisch eine (und nur eine) Folge der Länge 0 vor.

In jeder Menge $W(\boldsymbol{A})$ ist eine *Verkettung* genannte und mit dem Symbol $\circ$ bezeichnete zweistellige Operation definiert, die beliebigen Wörtern $\mathrm{U}, \mathrm{V} \in W(\boldsymbol{A})$ dasjenige Wort $\mathrm{U} \circ \mathrm{V}$ zuordnet, das man durch Hintereinanderschreiben der Wörter U, V (in dieser Reihenfolge) erhält. Es ist also zum Beispiel $\mathrm{abc} \circ \mathrm{cac} = \mathrm{abccac}$ und insbesondere $\mathrm{U} \circ \Lambda = \Lambda \circ \mathrm{U} = \mathrm{U}$ für jedes Wort $\mathrm{U} \in W(\boldsymbol{A})$. Es ist anschaulich klar, daß die Verkettungsoperation *assoziativ* ist, d. h.

$$(\mathrm{W} \circ \mathrm{U}) \circ \mathrm{V} = \mathrm{W} \circ (\mathrm{U} \circ \mathrm{V})$$

für $\mathrm{U}, \mathrm{V}, \mathrm{W} \in W(\boldsymbol{A})$ gilt. Aus diesem speziellen Assoziativgesetz für die Operation $\circ$ folgt bekanntlich[1]) das allgemeine Assoziativgesetz: Die Verkettung endlich vieler Wörter $\mathrm{W}_1, \mathrm{W}_2, \ldots, \mathrm{W}_n$ ist unabhängig von der Klammerung und kann daher ohne Mißverständnisse in der Form $\mathrm{W}_1 \circ \mathrm{W}_2 \circ \cdots \circ \mathrm{W}_n$ geschrieben werden. Insbesondere läßt sich jedes nichtleere Wort $a_{i_1} a_{i_2} \ldots a_{i_n}$ in dieser Form auf genau eine Weise als Verkettung einbuchstabiger Wörter $a_{i_1} \circ a_{i_2} \circ \cdots \circ a_{i_n}$ schreiben, und die einbuchstabigen Wörter sind unter allen nichtleeren Wörtern gerade dadurch ausgezeichnet, daß sie sich nicht in zwei echte Verkettungsfaktoren zerlegen lassen, d. h., daß für sie aus der Gleichung $\mathrm{W} = \mathrm{U} \circ \mathrm{V}$ stets $\mathrm{W} = \mathrm{U}$ (und damit $\mathrm{V} = \Lambda$) oder $\mathrm{W} = \mathrm{V}$ (und damit $\mathrm{U} = \Lambda$) folgt.

Bisher haben wir uns unter Buchstaben immer geschriebene bzw. schreibbare Zeichen vorgestellt, und die Tatsache, daß man Erörterungen über Buchstaben von möglicherweise anderer Natur nicht niederschreiben kann, ohne für sie wenigstens Bezeichnungen in Form „echter“ Buchstaben zu wählen, unterstützt diese Vorstellung. Die allgemeine Definition der Begriffe Alphabet und Wort (beliebige nicht-

[1]) Vgl. MfL, Band 1, S. 125. Der dort für eine bestimmte assoziative Operation geführte Beweis kann wörtlich auf jede andere assoziative Operation übertragen werden.

leere Menge bzw. endliche Folge von Elementen dieser Menge) schließt jedoch nicht aus, daß im konkreten Fall die Buchstaben z. B. auch Laute oder Lichtsignale (verschiedener Farbe oder Länge) sein können und die Verkettung der so gebildeten Wörter als zeitliche Aufeinanderfolge von Buchstaben realisiert ist. Wie es auch in anderen Gebieten der Mathematik üblich ist, lösen wir unsere Untersuchungen dadurch von den Problemen der verschiedenartigen Deutungsmöglichkeiten unserer Begriffe, daß wir im folgenden von einer axiomatischen Beschreibung der Begriffe Buchstabe, Wort, Verkettung ausgehen, wobei die Axiome durch Abstraktion aus den Eigenschaften der ursprünglich intendierten Objekte entstehen.

Definition 1. Es sei F eine nichtleere Menge, E eine nichtleere Teilmenge von F, $e \in F$ ein beliebiges Element von F, das Symbol $\circ$ bezeichne eine binäre Operation in F, d. h. eine Abbildung von $F \times F$ in F. Das System $\mathfrak{F} = (F, \circ, e, E)$ heißt eine *freie Halbgruppe* mit dem *neutralen Element* e und dem *Erzeugendensystem* E (kurz *freie Halbgruppe*), wenn folgende Axiome erfüllt sind:

A 1 *$\circ$ ist assoziativ, d. h., es gilt*

$$(x \circ y) \circ z = x \circ (y \circ z) \quad \textit{für alle } x, y, z \in F.$$

A 2 *e ist neutrales Element, d. h., es gilt*

$$x \circ e = e \circ x = x \quad \textit{für alle } x \in F.$$

A 3 *E ist Erzeugendensystem, d. h., für alle Teilmengen M von F gilt: Wenn $e \in M$ und $E \subseteq M$ und für $x, y \in M$ stets $x \circ y \in M$ folgt, so ist $M = F$* (d. h., F ist die kleinste bezüglich $\circ$ abgeschlossene Teilmenge von F, die e und alle Elemente von E enthält).

A 4 *F ist frei, d. h.,*

a) *für $x, y \in F$ folgt aus $x \circ y = e$ stets $x = y = e$,*

b) *für $x, y \in F$ und $a, b \in E$ folgt aus $x \circ a = y \circ b$ stets $x = y$ und $a = b$.*

A 5 *$e \notin E$.*

Zunächst ist anschaulich klar, daß das System $(W(A), \circ, \Lambda, A)$, wobei $\circ$ hier die Verkettung bezeichnet, bei beliebigem Alphabet A die Axiome A 1 bis A 5 erfüllt und daher eine freie Halbgruppe ist. Solche freien Halbgruppen werden wir *Worthalbgruppen* nennen. Der folgende Satz, dem wir aber noch eine Definition voranstellen, präzisiert, in welcher Weise der Begriff der freien Halbgruppe eine axiomatische Charakterisierung der anschaulichen Begriffe Alphabet, Buchstabe, Wort, Verkettung usw. liefert.

Definition 2. Ein *Isomorphismus* einer freien Halbgruppe $(F_1, \circ_1, e_1, E_1)$ auf eine freie Halbgruppe $(F_2, \circ_2, e_2, E_2)$ ist eine eineindeutige Abbildung f von F_1 auf F_2, für die gilt:

$$f(x \circ_1 y) = f(x) \circ_2 f(y) \quad \textit{für alle } x, y \in F_1. \tag{1}$$

Zwei freie Halbgruppen heißen *isomorph*, wenn ein Isomorphismus zwischen ihnen existiert. (Die Isomorphie von freien Halbgruppen ist offenbar eine Äquivalenzrelation.)

Satz 1. *Sind* $\mathfrak{F}_1 = (F_1, \circ_1, e_1, E_1)$ *und* $\mathfrak{F}_2 = (F_2, \circ_2, e_2, E_2)$ *freie Halbgruppen und ist* f' *eine eineindeutige Abbildung von* E_1 *auf* E_2, *so gibt es genau einen Isomorphismus* f *von* F_1 *auf* F_2, *für den*

$$f(a) = f'(a) \qquad \text{für } a \in E_1 \tag{2}$$

gilt, d. h., jede eineindeutige Abbildung von E_1 *auf* E_2 *läßt sich auf genau eine Weise zu einem Isomorphismus von* F_1 *auf* F_2 *fortsetzen. Ist umgekehrt* f *ein Isomorphismus von* F_1 *auf* F_2, *so ist* $f(E_1) = E_2$, *d. h., freie Halbgruppen sind genau dann isomorph, wenn ihre Erzeugendensysteme gleichmächtig sind. Durch Vorgabe einer (endlichen oder unendlichen) Mächtigkeit ist jeweils eine und bis auf Isomorphie genau eine freie Halbgruppe bestimmt.*

Beweis. Zunächst ziehen wir einige Folgerungen aus den Axiomen A1 bis A5.

1. *Es gibt außer* e *kein weiteres Element* e' *von* F *mit*

$$x \circ e' = x \qquad \text{für alle } x \in F \tag{3}$$

oder

$$e' \circ x = x \qquad \text{für alle } x \in F. \tag{4}$$

Wäre z. B. (3) erfüllt, so wäre insbesondere $e \circ e' = e$, aber nach A2 auch $e \circ e' = e'$, folglich $e = e'$. Analog ist (4) zu behandeln.

2. $F = E = \{e\}$ *mit* $e \circ e = e$ *erfüllt die Axiome* A1 *bis* A4, *jedoch nicht* A5. Existiert wenigstens ein $x \neq e$ in F, so muß es auch ein $a \neq e$ in E geben, denn $E = \{e\}$ hätte nach A2 und A3 auch $F = \{e\}$ zur Folge. Wäre nun $e \in E$, so wäre wegen $(a \circ a) \circ e = a \circ a$ auf Grund von A4b) auf $a \circ a = a$ und $e = a$ zu schließen, im Widerspruch zu $a \neq e$. Folglich dient das Axiom A5 lediglich dazu, den Fall $E = F = \{e\}$ auszuschließen. Es kann gleichwertig durch das Axiom „*Es gibt ein* $x \neq e$“ ersetzt werden.

3. Auf Grund von A1 läßt sich jedes Halbgruppenelement x, das überhaupt als Verkettung von endlich vielen *Erzeugenden* (d. h. Elementen von E) darstellbar ist, z. B. $(a \circ b) \circ \big((c \circ d) \circ a\big)$, wobei $a, b, c, d \in E$, darstellen, in der *kanonischen Form* $\Big(\big(\cdots \big((a_{i_1} \circ a_{i_2}) \circ a_{i_3}\big) \circ \cdots\big) \circ a_{i_n}\Big)$ mit $a_{i_\nu} \in E$ für $\nu = 1, \ldots, n$. Dabei sei der Fall $n = 1$ nicht ausgeschlossen. Die kanonische Darstellung kürzen wir im folgenden mit $a_{i_1} \circ a_{i_2} \circ \cdots \circ a_{i_n}$ ab und bezeichnen sie auch als ein „E-Wort“. Diejenige Menge M, die aus e und allen durch ein E-Wort darstellbaren Halbgruppenelementen besteht, erfüllt die Voraussetzungen von A3. Hierbei ist nur der Fall nichttrivial, daß $x, y \in M$, $x = a_{i_1} \circ \cdots \circ a_{i_n}$, $y = a_{j_1} \circ \cdots \circ a_{j_m}$ gilt. Auf Grund des Assoziativgesetzes A1 ist aber dann $x \circ y = a_{i_1} \circ \cdots \circ a_{i_n} \circ a_{j_1} \circ \cdots a_{j_m}$, d. h. durch ein E-Wort darstellbar und folglich

Element von M. Nach A3 folgt nun $M = F$, d. h., F besteht genau aus allen durch ein E-Wort darstellbaren Elementen und dem (etwa durch das leere E-Wort dargestellten) neutralen Element e von F. Es ist jetzt noch zu zeigen:

4. Zwei durch (eventuell leere) E-Worte dargestellte Halbgruppenelemente sind genau dann gleich, wenn ihre Darstellungen „buchstabenweise" übereinstimmen, d. h., für $n > 0$ ist niemals $a_{i_1} \circ \cdots \circ a_{i_n} = e$, und aus $a_{i_1} \circ \cdots \circ a_{i_n} = a_{j_1} \circ \cdots \circ a_{j_m}$ folgt $n = m$ und $a_{i_\nu} = a_{j_\nu}$ für $\nu = 1, \ldots, n$ (in beiden Fällen $a_\varkappa \in E$ vorausgesetzt). Die erste dieser Behauptungen folgt für $n = 1$ aus A5 und für $n \geqq 2$ aus A4a. Die zweite Behauptung zeigt man leicht unter der unwesentlichen Annahme $n \geqq m$ mit Hilfe von A4b durch vollständige Induktion bezüglich n.

Wir kommen nun zum eigentlichen Beweis von Satz 1. Auf Grund des bisher Gezeigten ist schon klar, daß jede freie Halbgruppe $(F, \circ, e, E)$ im wesentlichen als Worthalbgruppe $\bigl(W(E), \circ, \Lambda, E\bigr)$ aufgefaßt werden kann, wenn man die ursprünglichen Halbgruppenelemente durch ihre ihnen eineindeutig zugeordneten Darstellungen mittels E-Wörtern ersetzt. Sind nun $(F_1, \circ_1, e_1, E_1)$ und $(F_2, \circ_2, e_2, E_2)$ freie Halbgruppen und ist f' eine eineindeutige Abbildung von E_1 auf E_2 und soll so zu einer Abbildung f von F_1 auf F_2 fortgesetzt werden, daß (1) gilt, so muß insbesondere

$$f(x) = f(x \circ_1 e_1) = f(x) \circ_2 f(e_1) \qquad \textit{für alle } x \in F_1,$$

d. h. für alle $f(x) \in F_2$ gelten. Wegen der Eindeutigkeit des neutralen Elements ist daher $f(e_1) = e_2$. Ein beliebiges Element x von F_1, das von e_1 verschieden ist, wird durch ein ihm eineindeutig entsprechendes E-Wort $a_{i_1} \circ_1 a_{i_2} \circ_1 \cdots \circ_1 a_{i_n}$ dargestellt. Aus (1) folgt daher sofort (genauer: durch Induktion bezüglich n)

$$f(x) = f(a_{i_1}) \circ_2 f(a_{i_2}) \circ_2 \cdots \circ_2 f(a_{i_n}) = f'(a_{i_1}) \circ_2 \cdots \circ_2 f'(a_{i_n}).$$

Das bedeutet, daß ein Isomorphismus von F_1 auf F_2, falls er existiert, durch seine Werte für die Erzeugenden bereits eindeutig bestimmt ist. Umgekehrt erfüllt aber die durch

$$\begin{aligned} &f(a) := f'(a) \qquad \textit{für } a \in E_1,\\ &f(e_1) := e_2,\\ &f(a_{i_1} \circ_1 \cdots \circ_1 a_{i_n}) := f(a_{i_1}) \circ_2 \cdots \circ_2 f(a_{i_n}) \end{aligned}$$

definierte Abbildung f von F_1 auf F_2 die Isomorphiebedingung (1). Es bleibt noch zu zeigen, daß isomorphe freie Halbgruppen gleichmächtige Erzeugendensysteme haben. Dazu genügt es zu zeigen, daß für einen Isomorphismus f für $a \in E_1$ stets $f(a) \in E_2$ gilt. (Anwendung auf den Isomorphismus f^{-1} von F_2 auf F_1 ergibt die Umkehrung.) Es sei $a \in E_1$ und f ein Isomorphismus. Dann ist $f(a) \in F_2$ genau dann kein Element von E_2, wenn $f(a)$ sich als Verkettung $b \circ_2 c$ zweier Elemente $b, c \neq e_2$ von F_2 darstellen läßt. Nehmen wir dies an, so ist $a = f^{-1}\bigl(f(a)\bigr) = f^{-1}(b \circ_2 c)$ $= f^{-1}(b) \circ_1 f^{-1}(c)$, da mit f auch f^{-1} ein Isomorphismus ist. Daher folgt auch wegen $b, c \neq e_2$, daß $f^{-1}(b), f^{-1}(c) \neq e_1$ gilt. Mithin ist a als Verkettung zweier von e_1 verschiedener Elemente dargestellt, folglich kein Element von E_1. Damit ist Satz 1 vollständig bewiesen.

Der durch Definition 1 erfaßte Begriff der freien Halbgruppe ist in gewissem Sinn etwas allgemeiner als der anschauliche Begriff der Worthalbgruppe, da in den Axiomen A 1 bis A 5 die Mächtigkeit des Erzeugendensystems überhaupt keine Rolle spielt, während es kaum möglich sein dürfte, ein System von überabzählbar vielen schreibbaren und paarweise unterscheidbaren Zeichen als Alphabet im anschaulichen Sinn zu definieren. Als ein abzählbares Alphabet im anschaulichen Sinn könnten z. B. die Zeichen

I, ⊔, ⊔⊔, ⊔⊔⊔, ⊔⊔⊔⊔, ⊔⊔⊔⊔⊔, ...

dienen. In diesem Buch wird nirgends ein überabzählbares Alphabet vorausgesetzt oder benutzt werden. Im folgenden werden wir wieder anstelle von Erzeugendensystem, Erzeugendes (Element), Halbgruppenelement und neutrales Element die anschaulicheren Bezeichnungen Alphabet, Buchstabe, Wort bzw. leeres Wort verwenden, jedoch in den folgenden Definitionen, Sätzen und Beweisen keine darüber hinausgehenden Grundbegriffe und keine Voraussetzungen außer den in den Axiomen A 1 bis A 5 formulierten benutzen. Wir beginnen mit der Definition einiger Begriffe. („$:\leftrightarrow$" bedeutet: Das Linksstehende ist definitionsgemäß gleichbedeutend mit dem Rechtsstehenden. „$:=$" bedeutet: Das links bezeichnete Objekt ist definitionsgemäß gleich dem rechts bezeichneten Objekt.) Großbuchstaben, vorzugsweise P, Q, R, S, T, U, V, W, dienen als Variablen für beliebige Wörter, Kleinbuchstaben, vorzugsweise a, b, c, x, y, als Variablen für Buchstaben, d. h. Elemente des Erzeugendensystems. Das Symbol für die Verkettungsoperation wird unterdrückt, d. h., UV steht für U $\circ$ V, RaS für R $\circ$ a $\circ$ S, usw.

W *beginnt mit* U $:\leftrightarrow$ *Es gibt ein Wort* R, *so daß* W $=$ UR.

Da R $= \Lambda$ sein kann, beginnt W stets mit W. Da R $=$ W sein kann, beginnt W stets mit Λ.

W *endet mit* U $:\leftrightarrow$ *Es gibt ein Wort* S, *so daß* W $=$ SU.

Da S $= \Lambda$ sein kann, endet W stets mit W. Da S $=$ W sein kann, endet W stets mit Λ.

U *kommt in* W *vor* $:\leftrightarrow$ *Es gibt Wörter* S, R, *so daß* W $=$ SUR. (5)

Statt U *kommt in* W *vor* sagen wir auch: U *ist Teilwort von* W. Durch die Definition (5) ist die zukünftige Verwendung dieser Redeweisen klar von anderen denkbaren Bedeutungen abgegrenzt. (In gewissem Sinne kommt z. B. auch aa in bababb vor, jedoch nicht im Sinne von Definition (5).) Da die Wörter S, R gleich dem leeren Wort sein können, folgt sofort: Wenn W mit U beginnt, so kommt U in W vor. Wenn W mit U endet, so kommt U in W vor. Stets kommt W in W vor. Kommt U in W vor und ist ungleich W, so sagen wir auch: U ist *echtes* Teilwort von W.

Wir definieren nun eine dreistellige Operation *Sub* (*Substitution*), und zwar soll *Sub* (W, a, U) für Wörter W, U und Buchstaben a dasjenige Wort bezeichnen, das aus W entsteht, wenn man alle in W vorkommenden Buchstaben a durch das Wort U ersetzt. Anders ausgedrückt: Hat W die Form

$$S_1 a S_2 a S_3 \ldots S_{n-1} a S_n, \tag{6}$$

wobei a in den Wörtern $S_1, \ldots, S_n$ nicht vorkommt, so soll

$$\mathit{Sub}(W, a, U) = S_1 U S_2 U S_3 \ldots S_{n-1} U S_n$$

sein. (Da S_1 bzw. S_n das leere Wort sein kann, sind auch die Fälle erfaßt, in denen W mit a beginnt bzw. endet.) Insbesondere soll $\mathit{Sub}(W, a, U) = W$ sein, falls a nicht in W vorkommt. Man könnte also z. B. definieren:

$$\mathit{Sub}(W, a, U) := \begin{cases} W, \textit{ falls a nicht in W vorkommt}, \\ S_1 U S_2 U S_3 \ldots S_{n-1} U S_n, \textit{ falls } W = S_1 a S_2 a S_3 \ldots S_{n-1} a S_n \textit{ ist} \\ \quad \textit{und a in } S_1, \ldots, S_n \textit{ nicht vorkommt.} \end{cases}$$

Die in dieser Definition unvermeidlichen „...“ kann man durch eine Art induktiver Definition vermeiden, die sich auf die Tatsache stützt, daß jedes Wort W auf genau eine Weise durch fortlaufendes Anhängen von Buchstaben aus dem leeren Wort erhalten werden kann:

$$\mathit{Sub}(\Lambda, a, U) = \Lambda,$$

ist schon $\mathit{Sub}(W, a, U)$ *definiert, so sei*

$$\mathit{Sub}(Wx, a, U) = \begin{cases} \mathit{Sub}(W, a, U) \circ x, \textit{ falls } x \neq a, \\ \mathit{Sub}(W, a, U) \circ U, \textit{ falls } x = a. \end{cases}$$

Als Beispiel für die Anwendung der Operation *Sub* betrachten wir die Elimination einer Variablen x aus einer Gleichung der Form $f(x, y) = 0$ mittels einer zweiten, bereits nach x aufgelösten Gleichung $x = g(y)$. Hierbei sei vorausgesetzt, daß $f(x, y) = 0$ und $x = g(y)$ (und eine Schar weiterer Gleichungen bestimmten Typs) wirklich Zeichenreihen über einem bestimmten Alphabet sind. Ist z. B. das Alphabet $A = \{x, y, a, b, c, =, 0, (,), +, -, {}^2\}$ gegeben, wobei ^{2} wie üblich zur Bezeichnung der zweiten Potenz dient, so könnte das Wort g(y) die Gestalt $ay^2 + by + c$ und das Wort f(x, y) die Gestalt $(x + y)^2 + a(y - x) + cxy$ haben. Nun ergibt aber $\mathit{Sub}\big(f(x, y) = 0, x, g(y)\big)$ die auf Grund der üblichen Klammerregeln inhaltlich falsche Schlußfolgerung

$$(ay^2 + by + c + y)^2 + a(y - ay^2 + by + c) + cay^2 + by + cy = 0.$$

Dieser Fehler kann jedoch leicht dadurch behoben werden, daß man statt g(y) das inhaltlich gleichwertige Wort $(ay^2 + by + c)$ für x einsetzt. Der Leser bilde selbst

$Sub(\mathrm{f}(\mathrm{x}, \mathrm{y}) = 0, \mathrm{x}, (\mathrm{ay}^2 + \mathrm{by} + \mathrm{c}))$ und denke dabei ein wenig über die Rolle der Klammern in arithmetischen Formeln nach, deren üblicher abgekürzter Gebrauch zwar das „intelligente" Rechnen erleichtert, jedoch einem formalen Operieren mit Zeichenreihen nach bestimmten Regeln hinderlich ist.

Für einige spätere Anwendungen benötigen wir die *mehrfache* oder *simultane* Substitution: Sind $\mathrm{a}_1, \ldots, \mathrm{a}_k$ paarweise verschiedene Buchstaben, so sei

$$Sub(\mathrm{W}; \mathrm{a}_1, \mathrm{U}_1; \mathrm{a}_2, \mathrm{U}_2; \ldots; \mathrm{a}_k, \mathrm{U}_k) :=$$

$$:= \begin{cases} \mathrm{W}, \textit{ falls } \mathrm{a}_1, \ldots, \mathrm{a}_k \textit{ nicht in } \mathrm{W} \textit{ vorkommen,} \\ \mathrm{S}_0\mathrm{U}_{i_1}\mathrm{S}_1\mathrm{U}_{i_2}\mathrm{S}_2 \ldots \mathrm{S}_{n-1}\mathrm{U}_{i_n}\mathrm{S}_n, \textit{ falls} \\ \qquad \mathrm{W} = \mathrm{S}_0\mathrm{a}_{i_1}\mathrm{S}_1\mathrm{a}_{i_2}\mathrm{S}_2 \ldots \mathrm{S}_{n-1}\mathrm{a}_{i_n}\mathrm{S}_n \textit{ ist und } \mathrm{a}_1, \ldots, \mathrm{a}_k \textit{ nicht} \\ \qquad \textit{in } \mathrm{S}_0, \mathrm{S}_1, \ldots, \mathrm{S}_n \textit{ vorkommen.} \end{cases}$$

Es ist also z. B. $Sub(\mathrm{abcbb}; \mathrm{a}, \mathrm{bb}; \mathrm{b}, \mathrm{ca}) = \mathrm{bbcaccaca}$.

Wir wollen nun die einfache Substitution auf den Fall ausdehnen, daß nicht notwendig ein Buchstabe a, sondern ein beliebiges nichtleeres Wort V an allen Stellen seines Vorkommens in W durch ein anderes Wort U ersetzt wird. Dies spielt z. B. überall dort eine Rolle, wo Texte durch Anwendung einer abkürzenden Definition für gewisse mehrfach vorkommende Teilstücke verkürzt werden sollen. Eine allgemeine Definition stößt hier jedoch auf die Schwierigkeit, daß sich verschiedene Vorkommen des Teilwortes V im Wort W überschneiden können. Zum Beispiel kommt aba in babababab an den drei unterstrichenen Stellen vor. Was soll nun etwa $Sub(\mathrm{babababab}, \mathrm{aba}, \mathrm{b})$ sein? Der Bestimmtheit halber legt man fest: Die Stellen des Vorkommens von V in W sind von links beginnend so zu markieren, daß jeweils hinter dem zuletzt markierten Teilwort mit der Suche nach dem nächsten Vorkommen begonnen wird. Damit dieser Prozeß stets ein Resultat liefert, muß insbesondere $\mathrm{V} \neq \Lambda$ sein, da sonst das zu untersuchende Restwort immer wieder gleich W wäre. Danach sind alle markierten Vorkommen von V durch U zu ersetzen. Zur bequemeren Formulierung einer exakten Definition führen wir zunächst einen Hilfsbegriff ein:

$$\mathrm{V} \textit{ in } \mathrm{X} \textit{ höchstens hinten} :\leftrightarrow \textit{Für alle Wörter } \mathrm{PQ} \textit{ folgt aus } \mathrm{X} = \mathrm{PVQ} \textit{ stets } \mathrm{Q} = \Lambda. \tag{7}$$

$$Sub(\mathrm{W}, \mathrm{V}, \mathrm{U}) := \begin{cases} \textit{nicht definiert für } \mathrm{V} = \Lambda, \\ \mathrm{W}, \textit{ falls } \mathrm{V} \textit{ nicht in } \mathrm{W} \textit{ vorkommt,} \\ \mathrm{S}_1\mathrm{US}_2\mathrm{US}_3 \ldots \mathrm{S}_{n-1}\mathrm{US}_n, \textit{ falls } \mathrm{W} = \mathrm{S}_1\mathrm{VS}_2\mathrm{VS}_3 \ldots \mathrm{S}_{n-1}\mathrm{VS}_n \\ \qquad \textit{ist und } \mathrm{V} \textit{ in } \mathrm{S}_i\mathrm{V} \textit{ höchstens hinten } (i = 1, \ldots, n-1) \\ \qquad \textit{und } \mathrm{V} \textit{ nicht in } \mathrm{S}_n \textit{ vorkommt.} \end{cases}$$

Offenbar stimmt diese Definition[1]) für den Fall, daß V ein Buchstabe a ist, mit der ursprünglichen Definition überein, denn für Buchstaben a gilt:

a *in* Sa *höchstens hinten genau dann, wenn* a *nicht in* S *vorkommt.*

Dem Leser sei zur Übung empfohlen, diesen Satz unter alleiniger Benutzung der beiden Definitionen (5) und (7) und der Axiome A1 bis A5 zu beweisen.

Es sei A ein beliebiges Alphabet. Jedem Wort W $\in W(A)$ ist in anschaulich unmißverständlicher Weise eine natürliche Zahl l(W) als seine *Länge* zugeordnet, so daß l eine Abbildung von $W(A)$ auf die Menge N der natürlichen Zahlen ist. Unter Benutzung des bereits zur Definition von *Sub*(W, a, U) herangezogenen Induktionsprinzips könnte man auch induktiv definieren:

$l(\Lambda) = 0,$

ist schon l(W) *definiert, so sei*

$l(\mathrm{Wx}) = l(\mathrm{W}) + 1$ *für beliebige Buchstaben* $\mathrm{x} \in A$.

Für natürliche Zahlen $n = 0, 1, 2, 3, \ldots$ sei $W^n(A)$ die Menge aller Wörter der Länge n im Alphabet A. Wie aus der Kombinatorik bekannt ist, besteht $W^n(A)$ aus k^n Wörtern, falls das Alphabet A aus k Buchstaben besteht. Dies gilt auch noch für $n = 0$.

Die Abbildung l ist offenbar genau dann eineindeutig, wenn das Alphabet A genau einen Buchstaben a enthält. In diesem Fall können wir die Wörter aus $W(\{\mathrm{a}\})$ aber selbst mit einiger Berechtigung als die natürlichen Zahlen ansehen. (Kinder lernen die natürlichen Zahlen zuerst in dieser wortartigen Form kennen: n auf einen Draht gefädelte Kugeln gleicher Art sind nichts anderes als ein Wort der Länge n in einem einbuchstabigen Alphabet!) Das leere Wort entspricht hier der Null, der einzige Buchstabe der Eins, die Verkettung der Addition. Da, wie wir gezeigt haben, die Axiome A1 bis A5 allgemein eine axiomatische Charakterisierung der freien Halbgruppen leisten, ergeben A1 bis A5 zusammen mit dem Zusatzaxiom

A6 *Es gibt genau ein Element a von E*

eine neue axiomatische Charakterisierung der natürlichen Zahlen in bezug auf die Addition (statt wie üblich Nachfolgerbildung) als Grundverknüpfung. Umgekehrt kann man beliebige freie Halbgruppen in einer der Peanoschen Axiomatisierung der natürlichen Zahlen[2]) ähnlicheren Weise statt durch Verkettung durch ein System E von Nachfolgerfunktionen, d. h. einstelligen Abbildungen von der Menge aller Wörter in sich, charakterisieren, wobei jedes vom leeren Wort Λ verschiedene Wort auf genau eine Weise in der Form

$$f_{i_n}(f_{i_{n-1}}(\cdots f_{i_1}(\Lambda)\cdots))$$

[1]) Bezüglich anderer Möglichkeiten, die Substitution zu definieren, umgekehrt die Substitution statt Verkettung als Grundbegriff zu wählen, und vieler weiterer interessanter Details der Theorie der Zeichenreihen sei der Leser verwiesen auf K. Härtig, Explizite Definitionen einiger Eigenschaften von Zeichenreihen, Zeitschr. f. Math. Logik u. Grdlg. d. Math. 2 (1956), 177–203; siehe auch 3 (1957), 151–156.

[2]) Vgl. MfL, Band 1, Abschnitt 3.2.

dargestellt werden kann, die dem Wort $f_{i_1} f_{i_2} \ldots f_{i_n}$ im Alphabet E entspricht. Es ist dann axiomatisch zu fordern:

A'1 $f(W) \neq \Lambda$ für $f \in E$ und $\mathrm{W} \in W(E)$.

A'2 *Ist* $f_1(\mathrm{V}) = f_2(\mathrm{W})$, *so ist* $f_1 = f_2$ *und* $\mathrm{V} = \mathrm{W}$.

A'3 *Für alle Teilmengen* M *von* $W(E)$ *gilt: Ist* $\Lambda \in M$ *und für* $\mathrm{W} \in M$ *und* $f \in E$ *stets* $f(\mathrm{W}) \in M$, *so ist* $M = W(E)$ (d. h., $W(E)$ ist die kleinste bezüglich aller Operationen $f \in E$ abgeschlossene und Λ enthaltende Teilmenge von $W(E)$).

Wie man sieht, ergeben sich hier für den Spezialfall $E = \{f\}$ unmittelbar die Axiome von PEANO. Geht man allgemein von der Axiomatisierung durch A1 bis A5 bezüglich des Grundbegriffs Verkettung aus, so lassen sich die Nachfolgerfunktionen sofort durch

$$f_{\mathrm{a}}(\mathrm{W}) := \mathrm{W} \circ \mathrm{a} \quad \textit{für } \mathrm{W} \in W(A) \textit{ und } \mathrm{a} \in A$$

definieren und für die so definierten Funktionen A'1 bis A'3 als Sätze beweisen. Geht man umgekehrt vom Grundbegriff „Nachfolgerfunktionen" und den Axiomen A'1 bis A'3 aus, so läßt sich die Verkettung nur induktiv definieren:

$$\mathrm{V} \circ \Lambda = \mathrm{V},$$
$$\textit{ist schon } \mathrm{V} \circ \mathrm{W} \textit{ definiert, so sei } \mathrm{V} \circ f(\mathrm{W}) = f(\mathrm{V} \circ \mathrm{W}) \textit{ für } f \in E.$$

Für den Spezialfall eines nur aus einer Nachfolgerfunktion f bestehenden Systems E kommt hier natürlich die bekannte induktive Definition der Addition heraus.

Ein *endliches geordnetes Alphabet* A ist eine endliche Menge, die zugleich mit einer durch die Reihenfolge der Aufzählung ihrer Elemente angegebenen Ordnungsrelation $<$ versehen ist. Ein geordnetes Alphabet werden wir in der Form $A = (\mathrm{a}_1, \ldots, \mathrm{a}_n)$ (statt $\{\mathrm{a}_1, \ldots, \mathrm{a}_n\}$) angeben. Mit Hilfe der vorgegebenen Ordnung eines Alphabets A kann man zunächst für jede feste Zahl n die Menge $W^n(A)$ in einer aus der Kombinatorik wohlbekannten Weise *lexikographisch ordnen*. Diese Ordnung in $W^n(A)$ läßt sich mittels Verkettung und der Ordnung $<$ des Alphabets wie folgt exakt definieren:

$$\mathrm{V} \underset{lex^*}{<} \mathrm{W} :\leftrightarrow \textit{Es existieren Wörter } \mathrm{U, R, S} \textit{ und Buchstaben } \mathrm{a, b}, \textit{ so daß } \mathrm{a} < \mathrm{b}$$
$$\textit{und } \mathrm{V} = \mathrm{UaR} \textit{ und } \mathrm{W} = \mathrm{UbS}.$$

Es gibt nun zwei wesentlich verschiedene Arten, diese lexikographische Ordnung zu einer Ordnung der gesamten Menge $W(A)$ fortzusetzen. Die in Wörterbüchern benutzte Ordnung (der die lexikographische Ordnung ihren Namen verdankt) läßt sich wie folgt definieren:

$$\mathrm{V} \underset{''lex''}{\leqq} \mathrm{W} :\leftrightarrow \mathrm{W} \textit{ beginnt mit } \mathrm{V}, \textit{ oder es existieren Buchstaben } \mathrm{a, b} \textit{ mit } \mathrm{a} < \mathrm{b}$$
$$\textit{und ein Wort } \mathrm{U}, \textit{ so daß } \mathrm{V} \textit{ mit } \mathrm{Ua} \textit{ und } \mathrm{W} \textit{ mit } \mathrm{Ub} \textit{ beginnt.}$$

Diese Relation bewährt sich aber in den Wörterbüchern nur deshalb, weil diese immer nur eine endliche Teilmenge einer Menge $W(A)$ enthalten. Wollte man die gesamte Menge $W(A)$ nach diesem Prinzip ordnen ($\underset{''lex''}{\leqq}$ ist immerhin reflexiv, transitiv und antisymmetrisch!), so käme man zu einer sehr unübersichtlichen Reihenfolge. Zum Beispiel folgen auf jedes Wort W erst alle Wörter der Form Waa ... a (a kleinster Buchstabe von A), bevor wieder kürzere Wörter wie Wb folgen.[1]) Daher bewährt sich für die Gesamtmenge $W(A)$ besser die folgende, mit $\underset{lex}{<}$ bezeichnete Variante, bei der die Wörter zunächst nach ihrer Länge und nur die Wörter gleicher Länge untereinander lexikographisch geordnet werden:

$$\mathrm{V} \underset{lex}{<} \mathrm{W} :\leftrightarrow l(\mathrm{V}) < l(\mathrm{W}) \textit{ oder}$$
$$l(\mathrm{V}) = l(\mathrm{W}) \textit{ und } \mathrm{V} \underset{lex^*}{<} \mathrm{W}$$

Die so definierte Relation $\underset{lex}{<}$ nennen wir die durch die vorgegebene Ordnung $<$ des Alphabets *induzierte lexikographische Ordnung* von $W(A)$. Wir zeigen im folgenden, daß diese Relation tatsächlich die Grundeigenschaften einer (irreflexiven) Ordnung und darüber hinaus die charakteristischen Ordnungseigenschaften der natürlichen Zahlen besitzt.

1. Aus der Definition folgt unmittelbar: *Für kein Wort* W *ist*

$$\mathrm{W} \underset{lex}{<} \mathrm{W}.$$

2. *Ist* $\mathrm{U} \underset{lex}{<} \mathrm{V}$ *und* $\mathrm{V} \underset{lex}{<} \mathrm{W}$, *so ist* $\mathrm{U} \underset{lex}{<} \mathrm{W}$.

Die Behauptung ist klar, falls $l(\mathrm{U}) < l(\mathrm{V})$ oder $l(\mathrm{V}) < l(\mathrm{W})$ ist, da dann $l(\mathrm{U}) < l(\mathrm{W})$ folgt. Ist $l(\mathrm{U}) = l(\mathrm{V}) = l(\mathrm{W})$, so gibt es Wörter S, T und Buchstaben a, b, c, d mit $\mathrm{a} < \mathrm{b}$ und $\mathrm{c} < \mathrm{d}$, so daß gilt:

U *beginnt mit* Sa,

V *beginnt mit* Sb,

V *beginnt mit* Tc,

W *beginnt mit* Td.

Falls hierbei $\mathrm{S} = \mathrm{T}$ ist, folgt $\mathrm{b} = \mathrm{c}$ und damit $\mathrm{a} < \mathrm{d}$, d. h. $\mathrm{U} \underset{lex}{<} \mathrm{W}$. Ist $l(\mathrm{S}) < l(\mathrm{T})$, so beginnt T mit Sb, folglich beginnt auch W mit Sb. Ist $l(\mathrm{T}) < l(\mathrm{S})$, so beginnt S mit Tc, folglich auch U mit Tc. In beiden Fällen folgt $\mathrm{U} \underset{lex}{<} \mathrm{W}$.

3. *Für* $\mathrm{V} \neq \mathrm{W}$ *ist* $\mathrm{V} \underset{lex}{<} \mathrm{W}$ *oder* $\mathrm{W} \underset{lex}{<} \mathrm{V}$.

Dies ist klar, falls $l(\mathrm{V}) \neq l(\mathrm{W})$ ist. Andernfalls gibt es unter den gemeinsamen Anfangsstücken U von V und W eines von maximaler Länge, die aber kleiner als $l(\mathrm{V})$ sein muß, da sonst $\mathrm{V} = \mathrm{W}$ wäre. Folglich beginnt V mit Ua und W mit Ub, wobei $\mathrm{a} \neq \mathrm{b}$, d. h. $\mathrm{a} < \mathrm{b}$ oder $\mathrm{b} < \mathrm{a}$ ist.

4. *Jede nichtleere Teilmenge M von $W(A)$ besitzt ein bezüglich $\underset{lex}{<}$ kleinstes Element.*

Zunächst gibt es in der nichtleeren Menge $\{l(\mathrm{W}) : \mathrm{W} \in M\}$ von natürlichen Zahlen eine kleinste natürliche Zahl n_0 und wegen der bekannten Anzahl aller Wörter der Länge n_0 höchstens k^{n_0}

[1]) Vgl. hierzu P. Schreiber, Lexikographische Ordnung als Grundbegriff der Semiotik, Math. Nachr. 38 (1968), 53–60.

Wörter der Länge n_0 in M (k Anzahl der Elemente von A). Unter diesen endlich vielen gibt es aber wegen 3. ein kleinstes.

5. *Insbesondere ist* $\Lambda \underset{lex}{<} \mathrm{W}$ *für* $\mathrm{W} \neq \Lambda$, da dann stets $l(\Lambda) < l(\mathrm{W})$ ist.

6. Aus 4. folgt, daß *jedes Wort* $\mathrm{W} \in W(A)$ *einen unmittelbaren Nachfolger* W' *bezüglich* $\underset{lex}{<}$ *besitzt*, nämlich das kleinste Element der Menge $\{\mathrm{V} : \mathrm{V} \in W(A)$ *und* $\mathrm{W} \underset{lex}{<} \mathrm{V}\}$, die nichtleer ist, da sie insbesondere alle Wörter V mit $l(\mathrm{W}) < l(\mathrm{V})$ enthält.

7. *Jedes Wort* $\mathrm{W} \neq \Lambda$ *besitzt einen unmittelbaren Vorgänger* V, *so daß* $\mathrm{V}' = \mathrm{W}$ *ist*. Dazu zeigt man: Die Menge M aller Wörter, die gleich Λ sind oder einen unmittelbaren Vorgänger besitzen, erfüllt die Voraussetzungen von Axiom A3. Hierbei ist nur der Fall interessant, daß aus $\mathrm{U}, \mathrm{W} \in M$, $\mathrm{U}, \mathrm{W} \neq \Lambda$ auf $\mathrm{U} \circ \mathrm{W} \in M$ zu schließen ist. Ist V Vorgänger von W, so ist $\mathrm{U} \circ \mathrm{V}$ Vorgänger von $\mathrm{U} \circ \mathrm{W}$.

Aus 7. ergibt sich rein ordnungstheoretisch, daß jede Menge $M \subseteq W(A)$, die Λ und mit W auch W′ enthält, gleich $W(A)$ ist: Träfe dies auf eine Menge M nicht zu, so wäre $K = W(A) \setminus M$ nichtleer. Es sei dann U das kleinste Element von K; folglich ist $\mathrm{U} \neq \Lambda$ (wegen $\Lambda \in M$) und V der daher nach 7. existierende Vorgänger von U. Wegen $\mathrm{V} \in M$ ist auch $\mathrm{U} = \mathrm{V}' \in M$ im Widerspruch zu $\mathrm{U} \in K$.

Aus unseren Betrachtungen 1. bis 7. über die lexikographische Ordnung folgt insbesondere:

Satz 2. *Ist* A *ein endliches Alphabet, so ist* $W(A)$ *abzählbar.*

Beweis. A sei irgendwie geordnet, und $\underset{lex}{<}$ sei die durch diese Ordnung induzierte lexikographische Ordnung von $W(A)$. Dann ist die durch

$$f(\Lambda) = 0,$$
$$f(\mathrm{W}') = f(\mathrm{W}) + 1$$

induktiv definierte Abbildung eineindeutig von $W(A)$ auf die Menge N der natürlichen Zahlen.

Satz 3. *Ist* A *ein abzählbares Alphabet, so ist* $W(A)$ *abzählbar.*

Beweis. Es sei $A = \{\mathrm{a}_n : n = 0, 1, 2, \ldots\}$, $B = \{\mathrm{a}, \mathrm{b}\}$ ein zweibuchstabiges Alphabet. Nach Satz 2 ist $W(B)$ abzählbar. Die induktiv durch

$$g(\Lambda) = \Lambda,$$
$$g(\mathrm{W}\mathrm{a}_k) = g(\mathrm{W}) \circ \mathrm{b}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{k}\mathrm{b}$$

definierte Abbildung von $W(A)$ in $W(B)$ ist offenbar eineindeutig. Demnach ist $W(A)$ einer unendlichen Teilmenge einer abzählbaren Menge gleichmächtig, folglich selbst abzählbar.

2. Aussagenlogik

2.1. Aussagen und Wahrheitsfunktionen

Eine *Aussage* ist ein zu Mitteilungszwecken dienendes Objekt, das einen bestimmten Sachverhalt darstellt.[1]) In der Praxis treten Aussagen in den verschiedensten Formen auf: als gesprochene oder geschriebene Sätze (in deutscher oder einer anderen „natürlichen" Sprache), als Formeln, als Licht- oder Flaggensignale usw. Es genügt jedoch, die folgenden Betrachtungen auf solche Aussagen zu beziehen, die den Charakter von Zeichenreihen über einem bestimmten Alphabet haben.

Wenn der durch eine Aussage dargestellte Sachverhalt in der Realität besteht, bezeichnen wir diese Aussage als *wahr*, andernfalls als *falsch*. In der Umgangssprache trifft man auf zahlreiche „Aussagen", die nicht in eindeutiger Weise als wahr bzw. falsch bewertet werden können. Zum Beispiel bleibt die Wahrheit des Satzes „Heute ist schönes Wetter" wegen der Verschwommenheit des Begriffs „schönes Wetter" auch dann noch unbestimmt, wenn man weiß, auf welchen Tag und Ort sich „heute" bezieht. Wir beschränken uns im folgenden auf die Betrachtung von Aussagen, denen auf Grund hinreichend exakter Vereinbarungen über die benutzte „Sprache" und in einem bestimmten Zusammenhang genau einer der beiden *Wahrheitswerte* wahr bzw. falsch (kurz ***W***, ***F***) zugeordnet ist.

Jede natürliche Sprache erlaubt gewisse grammatische Konstruktionen, mit deren Hilfe man aus einer oder mehreren Aussagen beliebigen Inhalts eine neue Aussage so bilden kann, daß der Wahrheitswert dieser neuen Aussage nicht vom konkreten Inhalt der verwendeten Teilaussagen, sondern nur in einer bestimmten Weise von deren Wahrheitswerten abhängt. Zum Beispiel kann man in der deutschen

---

[1]) Auf eine genauere Definition und Analyse des Begriffs Aussage vom philosophischen Standpunkt muß hier verzichtet werden. Der interessierte Leser sei auf [21], insbesondere S. 31ff. verwiesen.

Sprache aus einer beliebigen Aussage A die Aussage *A ist falsch* (kurz *nicht A*) bilden, die genau dann wahr ist, wenn A falsch ist. Analog kann man aus beliebigen Aussagen A, B die Aussagen *A und B*, *A oder B*, *entweder A oder B*, *weder A noch B*, *A gilt genau dann, wenn B gilt* usw. bilden. Der Wahrheitswert der Gesamtaussage läßt sich dabei jeweils durch eine Wertetabelle darstellen, in deren Eingangsspalte alle möglichen Kombinationen von Wahrheitswerten für die Aussagen A, B erscheinen:

A	B	A *und* B	A *oder* B	*entweder* A *oder* B	*weder* A *noch* B
W	**W**	**W**	**W**	**F**	**F**
W	**F**	**F**	**W**	**W**	**F**
F	**W**	**F**	**W**	**W**	**F**
F	**F**	**F**	**F**	**F**	**W**

Wir bemerken, daß durch diese Tabelle gleichzeitig im Fall der Wörter *oder* bzw. *entweder oder*, deren Gebrauch in der Umgangssprache nicht immer einheitlich ist, eine gewisse Festlegung getroffen wurde, die innerhalb der Mathematik zweckmäßig und üblich ist. Sprachkonstruktionen nach Art der behandelten Beispiele, die auf eine feste Anzahl (nicht notwendig zwei) beliebiger Aussagen anwendbar sind und eine Gesamtaussage liefern, deren Wahrheitswert in einer durch eine Wertetabelle angebbaren Weise von den Wahrheitswerten der verwendeten Teilaussagen abhängt, bezeichnet man als *extensionale Aussagenverbindungen*. Um die Bedeutung des Beiwortes *extensional* in diesem Zusammenhang zu erläutern, geben wir ein Beispiel einer nichtextensionalen Aussagenverbindung an. Dazu denken wir uns eine bestimmte Person X fixiert, die über alles eine — manchmal richtige und manchmal falsche — Meinung hat. Nun kann man aus einer beliebigen Aussage A die neue Aussage *X hält A für wahr* bilden. Der Wahrheitswert dieser Aussage hängt offenbar nicht nur von der Wahrheit bzw. Falschheit von A ab.

Die Wertetabelle einer extensionalen Aussagenverbindung spiegelt das Wesentliche viel besser wider als die Aussagenverbindung selbst. Es ist z. B. ganz gleichgültig, ob man *A und B* oder *sowohl A als auch B* sagt. Beide Aussagenverbindungen haben die gleiche Wertetabelle, stellen den gleichen logischen Zusammenhang zwischen den Teilaussagen A, B her. Die durch die Wertetabelle einer extensionalen Aussagenverbindung gegebene Funktion, und allgemeiner jede Funktion (beliebiger endlicher Stellenzahl) im Bereich der Wahrheitswerte **W**, **F** nennt man *Wahrheitsfunktion* (häufig auch *Boolesche Funktion* nach dem englischen Mathematiker George Boole (1815—1864)). Als mögliche Argumente einer k-stelligen Wahrheitsfunktion treten alle Kombinationen der Länge k aus den beiden Wahrheitswerten **W**, **F** auf (d. h. im wesentlichen alle Wörter der Länge k im Alphabet $\{\mathbf{W}, \mathbf{F}\}$). Der Definitionsbereich einer k-stelligen Booleschen Funktion besteht demnach aus 2^k Elementen. Da jedem dieser Elemente wiederum ein beliebiger der beiden Wahrheitswerte als

Funktionswert zugeordnet sein kann, gibt es 2^{2^k} k-stellige Wahrheitsfunktionen. Wir geben im folgenden die Tabellen aller ein- und zweistelligen Wahrheitsfunktionen an, wobei wir der Übersichtlichkeit halber die Argumente in senkrechter Richtung und die Spalten der zugehörigen Werte in waagerechter Richtung lexikographisch ordnen (in der durch $W < F$ induzierten lexikographischen Ordnung). Allgemein bezeichnen wir die Wahrheitsfunktionen in diesem Buch mit Φ_i^k, wobei der obere Index die Stellenzahl angibt und der untere Index in der angegebenen Weise zur Numerierung gleichstelliger Funktionen dient.

	Φ_1^1	Φ_2^1	Φ_3^1	Φ_4^1
W	W	W	F	F
F	W	F	W	F

	Φ_1^2	Φ_2^2	Φ_3^2	Φ_4^2	Φ_5^2	Φ_6^2	Φ_7^2	Φ_8^2	Φ_9^2	Φ_{10}^2	Φ_{11}^2	Φ_{12}^2	Φ_{13}^2	Φ_{14}^2	Φ_{15}^2	Φ_{16}^2
W W	W	W	W	W	W	W	W	W	F	F	F	F	F	F	F	F
W F	W	W	W	W	F	F	F	F	W	W	W	W	F	F	F	F
F W	W	W	F	F	W	W	F	F	W	W	F	F	W	W	F	F
F F	W	F	W	F	W	F	W	F	W	F	W	F	W	F	W	F

Unter den so systematisch aufgezählten Wahrheitsfunktionen kommen natürlich auch diejenigen vor, die den bereits bekannten Aussagenverbindungen *und*, *oder* usw. entsprechen. Für die übrigen kann man nachträglich nach darstellenden Aussagenverbindungen suchen. Zum Beispiel läßt sich Φ_7^2 durch *A genau dann, wenn B* und Φ_5^2 durch *wenn A, so B*, aber auch durch *nicht A oder B* ausdrücken. Die letztere Aussagenverbindung müßte man eigentlich, um Mißverständnisse auszuschließen, in der Form (*nicht A*) *oder* B schreiben, da sie andernfalls auch als *nicht* (A *oder* B) verstanden werden kann, was der Funktion Φ_{15}^2 entspricht. Derartige Probleme verschärfen sich bei komplizierteren Ineinanderschachtelungen von einfachen Aussagenverbindungen. In der täglichen Umgangssprache treten solche kompliziert zusammengesetzten Aussagenverbindungen kaum auf, und man kann dann meist aus dem Zusammenhang erkennen, was gemeint ist. In der Mathematik spielen jedoch komplizierte Aussagenverbindungen eine große Rolle. Dabei treten u. a. folgende Fragen auf:

— Welche Wahrheitsfunktionen kann man aus einem gegebenen Grundsystem von Wahrheitsfunktionen überhaupt durch fortgesetzte Superposition (Ineinanderschachtelung) erhalten?

— Gibt es endliche Systeme von einfachen (niedrigstelligen) Wahrheitsfunktionen, aus denen sich jede Wahrheitsfunktion durch Superposition gewinnen läßt?

— Wie kann man den Werteverlauf einer als Superposition dargestellten Wahrheitsfunktion rationell ermitteln?
— Wie kann man feststellen, ob zwei verschachtelte Aussagenverbindungen logisch gleichwertig sind, d. h. die gleiche Wahrheitsfunktion darstellen?

Der Beantwortung dieser Fragen sind die beiden folgenden Abschnitte gewidmet.

2.2. Aussagenlogische Ausdrücke, Allgemeingültigkeit und semantische Äquivalenz

Wir wählen in Anlehnung an den üblichen mathematischen Sprachgebrauch fünf Wahrheitsfunktionen als Grundfunktionen aus und bauen einen Funktionskalkül zur exakten Beschreibung aller aus diesen Grundfunktionen erhältlichen Wahrheitsfunktionen auf.

Funktion	traditionelle Bezeichnung	darstellende Aussagenverbindung	Symbol[1])
Φ_3^1	Negation	nicht	$\neg$
Φ_8^2	Konjunktion	und	$\wedge$
Φ_2^2	Alternative	oder	$\vee$
Φ_5^2	Implikation	wenn, so	$\rightarrow$
Φ_7^2	Äquivalenz	genau dann, wenn	$\leftrightarrow$

[1]) In der Literatur sind verschiedene andere Symbole verbreitet, u. a. $\sim$ statt $\neg$, & statt $\wedge$, $\cdot$ statt $\vee$, $\supset$ statt $\rightarrow$, $\equiv$ statt $\leftrightarrow$.

Das Alphabet C bestehe aus den fünf Funktionssymbolen $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$, den beiden Zeichen (,) und den abzählbar vielen Buchstaben $p_0, p_1, p_2, \ldots$, die als Variable für Wahrheitswerte benutzt werden. Aus der Wortmenge $W(C)$ sondern wir durch eine induktive Definition die Teilmenge *ausd* der (aussagenlogischen) Ausdrücke aus:

Definition 1.

(a) $p_i \in$ *ausd* $(i = 0, 1, 2, \ldots)$;

(b) *Wenn* $H \in$ *ausd, so sei auch* $\neg H \in$ *ausd*;

(Ausführlich bedeutet (b): Wenn die Zeichenreihe H zur Menge *ausd* gehört, so gehört auch diejenige Zeichenreihe $\neg$ H zu *ausd*, die aus H entsteht, indem man den Buchstaben $\neg$ davorschreibt. Wegen (a) ist also beispielsweise

$$\neg p_3 \in \textit{ausd}, \neg p_{238} \in \textit{ausd},$$

folglich auch

$$\neg\neg\, p_3 \in \mathit{ausd}, \neg\neg\neg\, p_3 \in \mathit{ausd}, \neg\neg\neg\neg\neg\neg\neg\, p_{238} \in \mathit{ausd} \text{ usw.}$$

Der Buchstabe H, eventuell mit Indizes, wird hier und im folgenden als Variable zur Bezeichnung beliebiger Ausdrücke benutzt.)

(c) *Wenn* $H_1 \in \mathit{ausd}$ *und* $H_2 \in \mathit{ausd}$, *so sei auch* $(H_1 \wedge H_2) \in \mathit{ausd}$, $(H_1 \vee H_2) \in \mathit{ausd}$, $(H_1 \to H_2) \in \mathit{ausd}$ *und* $(H_1 \leftrightarrow H_2) \in \mathit{ausd}$;

(d) *Eine Zeichenreihe* $Z \in W(C)$ *sei nur dann Element der Menge ausd, wenn sich dies durch endlich viele Anwendungen der Regeln* (a), (b), (c) *ergibt.*

Definition 1 bedeutet anders formuliert: Die Menge *ausd* ist die kleinste Teilmenge von $W(C)$, die alle einbuchstabigen Wörter der Form p_i enthält, mit jedem Element H auch die Zeichenreihe $\neg$ H und mit je zwei (nicht notwendig verschiedenen) Elementen H_1, H_2 auch die Zeichenreihen $(H_1 \wedge H_2)$, $(H_1 \vee H_2)$, $(H_1 \to H_2)$ und $(H_1 \leftrightarrow H_2)$ enthält.

Beispiel 1. Aufbau eines Ausdrucks durch Anwendung von (a), (b), (c):

$p_1, p_2 \in \mathit{ausd}$ $\qquad$ $p_3, p_4 \in \mathit{ausd}$ $\qquad$ nach (a).

Nach (b) folgt: $\qquad$ Nach (c) folgt:

$\neg\, p_1, \neg\, p_2 \in \mathit{ausd}$ $\qquad$ $(p_3 \to p_4) \in \mathit{ausd}$

Nach (c) folgt: $\qquad$ Nach (b) folgt:

$(\neg\, p_1 \wedge \neg\, p_2) \in \mathit{ausd}$ $\qquad$ $\neg\, (p_3 \to p_4) \in \mathit{ausd}$

Nach (c) folgt:

$$\big((\neg\, p_1 \wedge \neg\, p_2) \vee \neg\, (p_3 \to p_4)\big) \in \mathit{ausd}$$

Nach (b) folgt:

$$\neg\big((\neg\, p_1 \wedge \neg\, p_2) \vee \neg\, (p_3 \to p_4)\big) \in \mathit{ausd} \tag{1}$$

Jeder Ausdruck beschreibt offenbar eine bestimmte, durch Superposition aus den gewählten Grundfunktionen erzeugbare Wahrheitsfunktion, und umgekehrt wird jede derartige Wahrheitsfunktion durch wenigstens einen Ausdruck dargestellt. Diese anschaulich klare Beziehung zwischen Ausdrücken und Wahrheitsfunktionen wird durch die folgende Definition präzisiert.

Definition 2. Ist $H \in \mathit{ausd}$, so heißt eine Abbildung f, die mindestens den in H vorkommenden Variablen je einen der beiden Wahrheitswerte zuordnet, eine *Belegung von* H. Ist f eine Belegung von H, so sei $\mathit{Wert}(H, f)$ der Wahrheitswert, den die durch H dargestellte Wahrheitsfunktion an der durch f gegebenen Stelle ihres Definitionsbereiches annimmt. Ausführlich wird dies durch folgende induktive

Regeln zur Definition (und Berechnung) von $\mathit{Wert}(\mathrm{H}, f)$ ausgedrückt:

(a) $\mathit{Wert}\,(\mathrm{p}_i, f) \overset{\prime}{=} f(\mathrm{p}_i)$;

(b) $\mathit{Wert}\,(\neg\,\mathrm{H}, f) = \Phi_3^1\left(\mathit{Wert}\,(\mathrm{H}, f)\right)$;

(Das heißt, $\mathit{Wert}(\neg\,\mathrm{H}, f) = \boldsymbol{W}$ genau dann, wenn $\mathit{Wert}(\mathrm{H}, f) = \boldsymbol{F}$. Wir bemerken: Ist f eine Belegung von $\neg\,\mathrm{H}$, so ist f erst recht eine Belegung von H. Ist also schon $\mathit{Wert}(\mathrm{H}, f)$ für alle Belegungen von H definiert, so ist durch (b) $\mathit{Wert}(\neg\,\mathrm{H}, f)$ für alle Belegungen von $\neg\,\mathrm{H}$ definiert. Analoge Erklärungen zu den folgenden Schritten möge der Leser selbst formulieren.)

(c) $\mathit{Wert}\left((\mathrm{H}_1 \wedge \mathrm{H}_2), f\right) = \Phi_8^2\left(\mathit{Wert}(\mathrm{H}_1, f), \mathit{Wert}(\mathrm{H}_2, f)\right)$;

(d) $\mathit{Wert}\left((\mathrm{H}_1 \vee \mathrm{H}_2), f\right) = \Phi_2^2\left(\mathit{Wert}(\mathrm{H}_1, f), \mathit{Wert}(\mathrm{H}_2, f)\right)$;

(e) $\mathit{Wert}\left((\mathrm{H}_1 \rightarrow \mathrm{H}_2), f\right) = \Phi_5^2\left(\mathit{Wert}(\mathrm{H}_1, f), \mathit{Wert}(\mathrm{H}_2, f)\right)$;

(f) $\mathit{Wert}\left((\mathrm{H}_1 \leftrightarrow \mathrm{H}_2), f\right) = \Phi_7^2\left(\mathit{Wert}(\mathrm{H}_1, f), \mathit{Wert}(\mathrm{H}_2, f)\right)$.

Beispiel 2. Berechnung der Wertetabelle des Ausdrucks (1).

1	2	3	4	5	6	7	8	9	10	11
p_1	p_2	p_3	p_4	$\neg\,\mathrm{p}_1$	$\neg\,\mathrm{p}_2$	$(\neg\,\mathrm{p}_1 \wedge \neg\,\mathrm{p}_2)$	$(\mathrm{p}_3 \rightarrow \mathrm{p}_4)$	$\neg 8$	$(7 \vee 9)$	$\neg 10$
W	*W*	*W*	*W*	*F*	*F*	*F*	*W*	*F*	*F*	*W*
W	*W*	*W*	*F*	*F*	*F*	*F*	*F*	*W*	*W*	*F*
W	*W*	*F*	*W*	*F*	*F*	*F*	*W*	*F*	*F*	*W*
W	*W*	*F*	*F*	*F*	*F*	*F*	*W*	*F*	*F*	*W*
W	*F*	*W*	*W*	*F*	*W*	*F*	*W*	*F*	*F*	*W*
W	*F*	*W*	*F*	*F*	*W*	*F*	*F*	*W*	*W*	*F*
W	*F*	*F*	*W*	*F*	*W*	*F*	*W*	*F*	*F*	*W*
W	*F*	*F*	*F*	*F*	*W*	*F*	*W*	*F*	*F*	*W*
F	*W*	*W*	*W*	*W*	*F*	*F*	*W*	*F*	*F*	*W*
F	*W*	*W*	*F*	*W*	*F*	*F*	*F*	*W*	*W*	*F*
F	*W*	*F*	*W*	*W*	*F*	*F*	*W*	*F*	*F*	*W*
F	*W*	*F*	*F*	*W*	*F*	*F*	*W*	*F*	*F*	*W*
F	*F*	*W*	*W*	*W*	*W*	*W*	*W*	*F*	*W*	*F*
F	*F*	*W*	*F*	*W*	*W*	*W*	*F*	*W*	*W*	*F*
F	*F*	*F*	*W*	*W*	*W*	*W*	*W*	*F*	*W*	*F*
F	*F*	*F*	*F*	*W*	*W*	*W*	*W*	*F*	*W*	*F*

In Spalte 1 bis 4 sind in lexikographischer Reihenfolge alle Belegungen der Variablen $\mathrm{p}_1, \ldots, \mathrm{p}_4$ aufgeführt. Spalte 5 wird aus Spalte 1 durch Anwendung der Negation gewonnen, ebenso Spalte 6 aus Spalte 2. Spalte 7 entsteht durch Anwendung der Konjunktion auf Spalte 5 und Spalte 6 usw. Die letzte Spalte enthält das Endresultat, d. h. die Wertetabelle der durch den Ausdruck (1) dargestellten Wahrheitsfunktion. Es ist klar, daß man prinzipiell die Wertetabelle jedes Ausdrucks nach dem gleichen Verfahren berechnen kann, jedoch ist dies für hinreichend komplizierte (lange) Ausdrücke ohne Hilfe der EDV kaum durchführbar.

Definition 3. Ein Ausdruck H heißt *allgemeingültig*, wenn $Wert(\mathrm{H}, f) = \boldsymbol{W}$ für alle Belegungen f von H gilt. Die Menge der allgemeingültigen Ausdrücke wird mit *ag* bezeichnet.

Die Allgemeingültigkeit eines Ausdrucks kann man offenbar grundsätzlich durch Berechnung seiner Wertetabelle nachweisen bzw. widerlegen. Der Leser prüfe dies für die folgenden Ausdrücke nach.

Beispiel 3. Allgemeingültige Ausdrücke sind unter anderen

$$(p_1 \to p_1),$$

$$(p_1 \vee \neg p_1),$$

$$(p_1 \to (p_2 \to p_1)),$$

$$((p_1 \leftrightarrow p_2) \to ((p_1 \wedge p_2) \vee (\neg p_1 \wedge \neg p_2))).$$

Jeder allgemeingültige Ausdruck kann als ein Schema zur Bildung von zusammengesetzten Aussagen aufgefaßt werden, die bei beliebigem Inhalt der verwendeten Teilaussagen insgesamt stets wahr sind, anders formuliert: Jeder allgemeingültige Ausdruck ist das Schema einer Schar von Aussagen, die allein auf Grund ihrer logisch sprachlichen Struktur, also unabhängig vom konkreten Inhalt, wahr sind. Das systematische Aufsuchen derartiger Schemata, später die Erfassung und Beschreibung aller derartigen Schemata, gehört zu den ältesten Aufgaben der traditionellen Logik.

Definition 4. Ausdrücke H_1, H_2 heißen *semantisch äquivalent*, wenn der Ausdruck $(\mathrm{H}_1 \leftrightarrow \mathrm{H}_2)$ allgemeingültig ist.

Auf Grund des Wertverlaufs der Äquivalenzfunktion Φ_7^2 sind Ausdrücke genau dann semantisch äquivalent, wenn für jede Belegung f aller in H_1 und H_2 insgesamt vorkommenden Variablen $Wert(\mathrm{H}_1, f) = Wert(\mathrm{H}_2, f)$ gilt, d. h., wenn H_1 und H_2, abgesehen von gewissen eventuell vorkommenden fiktiven Variablen, die den Wertverlauf nicht beeinflussen, die gleiche Wahrheitsfunktion darstellen. Die semantische Äquivalenz von Ausdrücken ist von großer Bedeutung für das richtige logische Schließen, da jedes Paar semantisch äquivalenter Ausdrücke die exakte Begründung für die Gleichwertigkeit zweier extensionaler Aussagenverbindungen liefert. Zum Beispiel ist die Gleichwertigkeit der Aussagenverbindungen *wenn A, so B* bzw. *nicht A oder B* eine Folge der semantischen Äquivalenz der Ausdrücke $(p_1 \to p_2)$ und $(\neg p_1 \vee p_2)$ bzw. der Allgemeingültigkeit des Ausdrucks $((p_1 \to p_2) \leftrightarrow (\neg p_1 \vee p_2))$. Allgemein ist die intuitiv richtige Anwendung einfacher semantischer Äquivalenzen ein gewisses Kriterium für den Entwicklungsgrad des „logischen Denkens". In komplizierteren Fällen kann jedoch die semantische Äquivalenz auch von erfahrenen Mathematikern nicht ohne Rechnung erkannt werden.

Beispiel 4. Paare semantisch äquivalenter Ausdrücke:

$\neg\neg p_1$ *semantisch äquivalent* p_1,

$(p_1 \wedge p_2)$ *semantisch äquivalent* $(p_2 \wedge p_1)$,

$(p_1 \rightarrow p_2)$ *semantisch äquivalent* $(\neg p_2 \rightarrow \neg p_1)$,

$\big(p_1 \wedge (p_2 \vee \neg p_2)\big)$ *semantisch äquivalent* p_1.

Bevor wir weitere Beispiele formulieren, sollen einige allgemein übliche *Klammersparungsregeln* zur Vereinfachung der Schreibweise konkreter Ausdrücke vereinbart werden.

(a) Die Außenklammern fertiger Ausdrücke werden weggelassen.

(b) Die zweistelligen Funktionssymbole $\leftrightarrow$, $\rightarrow$, $\vee$, $\wedge$ haben in der angegebenen Reihenfolge zunehmende Bindekraft.

$p_1 \leftrightarrow p_2 \rightarrow p_3$ bedeutet demnach $\big(p_1 \leftrightarrow (p_2 \rightarrow p_3)\big)$ und nicht $\big((p_1 \leftrightarrow p_2) \rightarrow p_3\big)$.

$p_7 \wedge p_6 \rightarrow p_3 \vee p_4$ bedeutet demnach $\big((p_7 \wedge p_6) \rightarrow (p_3 \vee p_4)\big)$ und nicht $\Big(p_7 \wedge \big(p_6 \rightarrow (p_3 \vee p_4)\big)\Big)$.

(c) Konjunktion und Alternative sind assoziativ, d. h.,

$\big((p_1 \wedge p_2) \wedge p_3\big)$ *semantisch äquivalent* $\big(p_1 \wedge (p_2 \wedge p_3)\big)$,

$\big((p_1 \vee p_2) \vee p_3\big)$ *semantisch äquivalent* $\big(p_1 \vee (p_2 \vee p_3)\big)$.

Unter Berufung auf die hieraus folgenden allgemeinen Assoziativgesetze schreiben wir im folgenden $p_1 \wedge p_2 \wedge \ldots \wedge p_n$ statt $\Big(p_1 \wedge \Big(p_2 \wedge \big(\ldots \wedge (p_{n-1} \wedge p_n) \ldots\big)\Big)\Big)$ oder einer semantisch äquivalenten Klammerung, analog $p_1 \vee p_2 \vee \ldots \vee p_n$.

(d) Statt der Variablen p_i ($i = 0, 1, 2, 3, \ldots$) verwenden wir gelegentlich p, q, r, s, t.

Es sei nachdrücklich betont, daß Zeichenreihen, die unter Verwendung der Regeln (a) bis (d) geschrieben sind, selbst keine Ausdrücke im Sinne von Definition 1, sondern lediglich bequemer les- und schreibbare Abkürzungen solcher Ausdrücke sind. Insbesondere beziehen sich auch im folgenden allgemeine Betrachtungen (Definitionen, Sätze usw.) stets auf Ausdrücke im Sinne der Definition 1. Der Leser „rückübersetze" zur Übung einige der folgenden „Ausdrücke" in die unabgekürzte Form.

Wichtige semantische Äquivalenzen, als allgemeingültige Ausdrücke geschrieben, sind:

$p_1 \wedge p_2 \leftrightarrow p_2 \wedge p_1$ (*Kommutativgesetz der Konjunktion*)

$p_1 \wedge p_1 \leftrightarrow p_1$ (*Verschmelzungsgesetz der Konjunktion*)

$p_1 \vee p_2 \leftrightarrow p_2 \vee p_1$ (*Kommutativgesetz der Alternative*)

$$\begin{array}{ll}
p_1 \vee p_1 \leftrightarrow p_1 & (\textit{Verschmelzungsgesetz der Alternative}) \\
\left.\begin{array}{l} p_1 \vee p_2 \wedge p_3 \leftrightarrow (p_1 \vee p_2) \wedge (p_1 \vee p_3) \\ p_1 \wedge (p_2 \vee p_3) \leftrightarrow p_1 \wedge p_2 \vee p_1 \wedge p_3 \end{array}\right\} & (\textit{Distributivgesetze}) \\
\left.\begin{array}{l} \neg (p_1 \wedge p_2) \leftrightarrow \neg p_1 \vee \neg p_2 \\ \neg (p_1 \vee p_2) \leftrightarrow \neg p_1 \wedge \neg p_2 \end{array}\right\} & (\textit{de Morgansche Regeln}) \\
p_1 \rightarrow p_2 \leftrightarrow \neg p_2 \rightarrow \neg p_1 & (\textit{Regel der Kontraposition}) \\
(p_1 \leftrightarrow p_2) \leftrightarrow (p_1 \rightarrow p_2) \wedge (p_2 \rightarrow p_1) & (\textit{Elimination der Äquivalenz}) \\
p_1 \rightarrow p_2 \leftrightarrow \neg p_1 \vee p_2 & (\textit{Elimination der Implikation}) \\
p_1 \rightarrow (p_1 \rightarrow p_2) \leftrightarrow p_1 \rightarrow p_2 & (\textit{Prämissenverschmelzung}) \\
p_1 \rightarrow (p_2 \rightarrow p_3) \leftrightarrow p_1 \wedge p_2 \rightarrow p_3 & (\textit{Prämissenverbindung}) \\
p_1 \rightarrow (p_2 \rightarrow p_3) \leftrightarrow p_2 \rightarrow (p_1 \rightarrow p_3) & (\textit{Prämissenvertauschung})
\end{array}$$

Abschließend sei vermerkt, daß auch die Äquivalenz assoziativ ist:

$$\big((p_1 \leftrightarrow p_2) \leftrightarrow p_3\big) \leftrightarrow \big(p_1 \leftrightarrow (p_2 \leftrightarrow p_3)\big).$$

Es sei aber die Warnung angebracht, daß die Schreibweise $p_1 \leftrightarrow p_2 \leftrightarrow p_3$ im üblichen mathematischen Sprachgebrauch nicht etwa $p_1 \leftrightarrow (p_2 \leftrightarrow p_3)$ bedeutet wie im Fall der Konjunktion und Alternative, sondern gleich

$$p_1 \wedge p_2 \wedge p_3 \vee \neg p_1 \wedge \neg p_2 \wedge \neg p_3$$

zu setzen ist.

Es wurde schon darauf hingewiesen, daß ein Ausdruck im wesentlichen nichts anderes ist als das Schema einer ineinandergeschachtelten Aussagenverbindung. Insbesondere kann man sich für die in einem Ausdruck vorkommenden Variablen solche Aussagen eingesetzt denken, die ihrerseits nach einem durch einen Ausdruck beschreibbaren Schema aussagenlogisch zusammengesetzt sind. Demnach muß gelten:

Satz 1. *Sind* H_1 *und* H_2 *beliebige Ausdrücke, so ist die Zeichenreihe* $Sub(H_1, p_i, H_2)$, *die aus* H_1 *entsteht, indem man die Variable* p_i *an allen Stellen ihres Vorkommens* in H_1 *durch* H_2 *ersetzt, ebenfalls ein Ausdruck* H_3. *Ist* f *eine Belegung von* H_3, *so sei* f' *die durch*

$$f'(p_j) = \begin{cases} Wert(H_2, f) & \text{für } i = j, \\ f(p_j) & \text{sonst} \end{cases}$$

definierte Belegung. Dann gilt

$$Wert\big(Sub(H_1, p_i, H_2), f\big) = Wert(H_1, f'). \tag{2}$$

Einen exakten Beweis dieses anschaulich recht klaren Satzes kann man unter Benutzung von Definition 1 durch Induktion über die Kompliziertheit von H_1 führen: Es sei zunächst H_1 eine einzelne Variable p_j. Falls p_j gleich der zu ersetzenden

Variablen p_i ist, ist $Sub(H_1, p_i, H_2) = H_2$ und $Wert(H_2, f) = f'(p_i) = Wert(H_1, f')$ für jede Belegung f von H_2. Im Fall $j \neq i$ ist $Sub(H_1, p_i, H_2) = H_1$ und $Wert(H_1, f) = f(p_j) = f'(p_j) = Wert(H_1, f')$ für jede Belegung f von H_1. Es sei nun $H_0 = \neg H_1$ und der Satz für H_1 schon (bei beliebigem p_i und H_2) bewiesen. Zunächst ist

$$Sub(H_0, p_i, H_2) = \neg Sub(H_1, p_i, H_2).$$

Da nach Voraussetzung $Sub(H_1, p_i, H_2)$ ein Ausdruck ist, trifft dies daher auch auf $Sub(H_0, p_i, H_2)$ zu, und f ist genau dann eine Belegung des einen dieser Ausdrücke, wenn f auch Belegung des anderen Ausdrucks ist. Es sei nun f eine solche Belegung. Da (2) nach Voraussetzung für H_1 gilt, ist

$$\begin{aligned} & Wert\big(Sub(H_0, p_i, H_2), f\big) = Wert\big(\neg Sub(H_1, p_i, H_2), f\big) \\ & = \Phi_3^1\Big(Wert\big(Sub(H_1, p_i, H_2), f\big)\Big) \\ & = \Phi_3^1\big(Wert(H_1, f')\big) = Wert(\neg H_1, f') \\ & = Wert(H_0, f'). \end{aligned}$$

Folglich ist der Satz für H_0 bewiesen.

Es sei nun $H_0 = (H_1 \wedge H_1')$ und der Satz für H_1 und H_1' bei beliebigem f und H_2 bewiesen. Offenbar ist $Sub(H_0, p_i, H_2)$ gleich $\big(Sub(H_1, p_i, H_2) \wedge Sub(H_1', p_i, H_2)\big)$, und folglich ein Ausdruck. (Hier erweist es sich als zweckmäßig, daß wir bei der Formulierung des Satzes das Vorkommen von p_i in H_1 nicht verlangt haben. Eine Variable p_i, die in H_0 vorkommt, muß nicht in H_1 und in H_1' vorkommen.) Es sei nun H_2 ein beliebiger Ausdruck, f eine Belegung von $Sub(H_0, p_i, H_2)$. Dann ist f auch Belegung von $Sub(H_1, p_i, H_2)$ und von $Sub(H_1', p_i, H_2)$, und es gilt

$$\begin{aligned} & Wert\big(Sub(H_0, p_i, H_2), f\big) \\ & = \Phi_8^2\Big(Wert\big(Sub(H_1, p_i, H_2), f\big), Wert\big(Sub(H_1', p_i, H_2), f\big)\Big) \\ & = \Phi_8^2\big(Wert(H_1, f'), Wert(H_1', f')\big) = Wert\big((H_1 \wedge H_1'), f'\big) \\ & = Wert(H_0, f'). \end{aligned}$$

Für die restlichen Fälle verläuft der Beweis völlig analog.

Aus Satz 1 folgt unmittelbar

Satz 2. *Ist* H_1 *ein allgemeingültiger Ausdruck, so ist für beliebige Variablen* p_i *und Ausdrücke* H_2 *der Ausdruck* $Sub(H_1, p_i, H_2)$ *ebenfalls allgemeingültig.*

Als Spezialfall von Satz 2 ergibt sich weiter

Satz 3. *Sind* H_1 *und* H_2 *semantisch äquivalent, so sind für beliebige Variablen* p_i *und Ausdrücke* H_0 *die Ausdrücke* $Sub(H_1, p_i, H_0)$ *und* $Sub(H_2, p_i, H_0)$ *ebenfalls semantisch äquivalent.*

Zum Beweis von Satz 3 hat man außer Satz 2 und Definition 4 nur die offenbar richtige Gleichung

$$Sub\big((H_1 \leftrightarrow H_2), p_i, H_0\big) = \big(Sub(H_1, p_i, H_0) \leftrightarrow Sub(H_2, p_i, H_0)\big)$$

für Zeichenreihen zu benutzen.

Satz 4. *Sind* H_1 *und* H_2 *semantisch äquivalent und entsteht die Zeichenreihe* H_4 *dadurch aus dem Ausdruck* H_3, *daß man ein Vorkommen von* H_1 *in* H_3 *durch* H_2 *ersetzt* (*d. h., es gibt Wörter* P, Q, *so daß* $H_3 = PH_1Q$ *und* $H_4 = PH_2Q$), *so ist* H_4 *ein zu* H_3 *semantisch äquivalenter Ausdruck.*

Beweis. Wir betrachten zunächst den Spezialfall $H_1 = H_3$. Dann ist $H_4 = H_2$ und folglich H_4 ein zu H_3 semantisch äquivalenter Ausdruck. Der allgemeine Fall wird nun durch Induktion über die Kompliziertheit von H_3 bewiesen. Ist H_3 eine einzelne Variable, so tritt der Spezialfall ein. Der Satz gelte schon für einen Ausdruck H_3, und H_1 sei ein Teilausdruck von $\neg H_3$. Dann ist H_1 Teilausdruck von H_3, und nach Induktionsannahme entsteht durch Ersetzen von H_1 durch H_2 in H_3 der zu H_3 semantisch äquivalente Ausdruck H_4. Demnach entsteht durch Ersetzung von H_1 in $\neg H_3$ der zu $\neg H_3$ semantisch äquivalente Ausdruck $\neg H_4$. Gilt der Satz schon für Ausdrücke H_3', H_3'', so ist ein Teilausdruck H_1 von $(H_3' \circ H_3'')$ (wobei $\circ$ eines der Zeichen $\wedge, \vee, \rightarrow, \leftrightarrow$ bedeutet) entweder gleich dem Gesamtausdruck (Spezialfall) oder einem Teilausdruck von H_3' oder einem Teilausdruck von H_3''. In den letzten beiden Fällen schließt man analog zum Fall $\neg H_3$ weiter.

Satz 4 bedeutet, daß man jedes Paar semantisch äquivalenter Ausdrücke als Regel zur schrittweisen semantisch äquivalenten Umformung beliebiger anderer Ausdrücke auffassen kann. Insbesondere gestatten z. B. die bereits als Eliminationsregeln bezeichneten semantischen Äquivalenzen

$$(p_1 \leftrightarrow p_2) \leftrightarrow (p_1 \rightarrow p_2) \wedge (p_2 \rightarrow p_1),$$

$$(p_1 \rightarrow p_2) \leftrightarrow \neg p_1 \vee p_2,$$

die auf Grund von Satz 2 auch als Schemata

$$(H_1 \leftrightarrow H_2) \leftrightarrow (H_1 \rightarrow H_2) \wedge (H_2 \rightarrow H_1),$$

$$(H_1 \rightarrow H_2) \leftrightarrow \neg H_1 \vee H_2$$

geschrieben werden können, die schrittweise Umformung eines beliebigen Ausdrucks H in einen zu H semantisch äquivalenten Ausdruck H', in dem die beiden Zeichen $\leftrightarrow$ und $\rightarrow$ nicht mehr vorkommen. Ferner folgt aus der ersten de Morganschen Regel

$$\neg (p_1 \wedge p_2) \leftrightarrow \neg p_1 \vee \neg p_2 \tag{3}$$

die Eliminationsregel

$$p_1 \wedge p_2 \leftrightarrow \neg (\neg p_1 \vee \neg p_2), \tag{4}$$

analog folgt aus der zweiten de Morganschen Regel

$$p_1 \vee p_2 \leftrightarrow \neg(\neg p_1 \wedge \neg p_2).$$

(Die Allgemeingültigkeit von (4) kann man natürlich direkt mittels der Wertetabelle erschließen. Wenn wir hier formuliert haben, daß (4) aus (3) folgt, so bedeutet dies ausführlich: Da $p \leftrightarrow \neg\neg p$ allgemeingültig ist, ist nach Satz 2 auch

$$p_1 \wedge p_2 \leftrightarrow \neg\neg(p_1 \wedge p_2) \tag{5}$$

allgemeingültig. Ersetzt man hierin nach Satz 4 den Teilausdruck $\neg(p_1 \wedge p_2)$ durch den laut (3) semantisch äquivalenten Ausdruck $\neg p_1 \vee \neg p_2$, so erhält man einen zu (5) semantisch äquivalenten, mithin wieder allgemeingültigen Ausdruck (4).)

Folgerung. *Jede durch Superposition aus den fünf hier betrachteten Grundfunktionen darstellbare Wahrheitsfunktion läßt sich allein mittels Negation und Konjunktion und allein mittels Negation und Alternative erzeugen.*

2.3. Normalformen

Es bleibt zu zeigen, daß die fünf betrachteten Grundfunktionen ausreichen, um jede Wahrheitsfunktion beliebiger Stellenzahl durch Superposition zu erzeugen. Zu diesem Zweck betrachten wir Normalformen.

Ein Ausdruck der Form $H_1 \wedge H_2 \wedge \ldots \wedge H_n$, wobei für $i = 1, 2, \ldots, n$ jeder der Ausdrücke H_i (unabhängig von den anderen) entweder gleich p_i oder gleich $\neg p_i$ ist, heißt eine *Elementarkonjunktion in den Variablen* $p_1, \ldots, p_n$. Es sind zum Beispiel $p_1 \wedge \neg p_2 \wedge p_3$, $p_1 \wedge \neg p_2 \wedge \neg p_3$ Elementarkonjunktionen in den Variablen p_1, p_2, p_3. Offenbar gibt es für $n \geq 1$ genau 2^n verschiedene Elementarkonjunktionen in den Variablen $p_1, \ldots, p_n$, und wie man durch Induktion über n leicht bestätigt, nimmt jede Elementarkonjunktion für genau diejenige Belegung ihrer Variablen den Wert $\boldsymbol{W}$ an, die jeder unnegiert vorkommenden Variablen den Wert $\boldsymbol{W}$ und jeder negiert vorkommenden Variablen den Wert $\boldsymbol{F}$ zuordnet. Man kann also einer Elementarkonjunktion sofort ihren vollständigen Werteverlauf ansehen und umgekehrt zu jeder Wahrheitsfunktion, die genau einmal den Wert $\boldsymbol{W}$ annimmt, sofort eine sie darstellende Elementarkonjunktion aufschreiben. Zum Beispiel entsprechen sich auf diese Weise die durch

p_1	p_2	p_3	p_4	p_5	Φ
W	W	F	W	F	W
		sonst			F

gegebene Wahrheitsfunktion Φ und die Elementarkonjunktion

$$p_1 \wedge p_2 \wedge \neg p_3 \wedge p_4 \wedge \neg p_5.$$

Sind E_1 und E_2 verschiedene Elementarkonjunktionen in den Variablen $p_1, \ldots, p_n$, so sei $E_1 < E_2$, falls an der ersten Stelle von links, an der sich E_1 und E_2 unterscheiden, die negierte Variable in E_2 und die unnegierte Variable in E_1 steht. (Das entspricht einer „lexikographischen“ Ordnung nach der Vorschrift „unnegiert vor negiert“.)

Ein Ausdruck heißt eine (*eigentliche*) *kanonische alternative Normalform* (kurz *aNf*), wenn er die Form $E_1 \vee E_2 \vee \ldots \vee E_k$ $(k \geqq 1)$ hat, wobei $E_1, \ldots, E_k$ paarweise verschiedene Elementarkonjunktionen in den gleichen Variablen sind und $E_i < E_j$ für $1 \leqq i < j \leqq k$ gilt. Neben den eigentlichen *aNf* bezeichnen wir auch den Ausdruck $p_1 \wedge \neg p_1 \vee p_2 \wedge \neg p_2 \vee \ldots \vee p_n \wedge \neg p_n$ als (*uneigentliche*) *aNf* in den Variablen $p_1, \ldots, p_n$. Dann gibt es offenbar für beliebiges $n \geqq 1$ insgesamt genau 2^{2^n} *aNf* in den Variablen $p_1, \ldots, p_n$.

Sind $H_1, \ldots, H_k$ beliebige Ausdrücke, so gilt, wie man leicht durch Induktion über k bestätigt, für beliebige Belegungen f des Ausdrucks $H_1 \vee H_2 \vee \ldots \vee H_k$ genau dann

$$Wert(H_1 \vee H_2 \vee \ldots \vee H_k, f) = \boldsymbol{W},$$

wenn $Wert(H_i, f) = \boldsymbol{W}$ für wenigstens einen der Ausdrücke $H_1, \ldots, H_k$ gilt.

Unter Berücksichtigung des bereits diskutierten Wertverlaufs der Elementarkonjunktionen folgt hieraus sofort:

Jede eigentliche aNf $E_1 \vee E_2 \vee \ldots \vee E_k$ *wird für genau diejenigen k Belegungen wahr, bei denen jeweils eine der Elementarkonjunktionen* $E_1, \ldots, E_k$ *wahr wird. Die uneigentliche aNf* (*und nur diese*) *nimmt bei jeder Belegung ihrer Variablen den Wert* $\boldsymbol{F}$ *an.*

Folglich kann man zu jeder Wahrheitsfunktion beliebiger Stellenzahl sofort eine sie darstellende *aNf* aufschreiben und umgekehrt von jeder *aNf* ohne besondere Rechnung sofort ihren vollständigen Werteverlauf ablesen. Die Zuordnung zwischen den n-stelligen Wahrheitsfunktionen und den sie darstellenden *aNf* ist eineindeutig. Zum Beispiel entsprechen sich die *aNf*

$$p_1 \wedge p_2 \wedge p_3 \vee p_1 \wedge \neg p_2 \wedge \neg p_3 \vee \neg p_1 \wedge p_2 \wedge \neg p_3$$

und die Funktion

p_1	p_2	p_3	Φ_{106}^3
$\boldsymbol{W}$	$\boldsymbol{W}$	$\boldsymbol{W}$	$\boldsymbol{W}$
$\boldsymbol{W}$	$\boldsymbol{F}$	$\boldsymbol{F}$	$\boldsymbol{W}$
$\boldsymbol{F}$	$\boldsymbol{W}$	$\boldsymbol{F}$	$\boldsymbol{W}$
	sonst		$\boldsymbol{F}$

Folgerung. *Jede Wahrheitsfunktion ist als Superposition von Negation und Konjunktion bzw. von Negation und Alternative darstellbar.* (Man kann ja wahlweise noch die Konjunktion oder die Alternative aus den *aNf* eliminieren.)

Wir zeigen abschließend, daß eine geeignet gewählte Wahrheitsfunktion allein ausreicht, um die Negation und Konjunktion und damit alle Wahrheitsfunktionen auszudrücken. Um diesen Sachverhalt möglichst kurz und übersichtlich formulieren zu können, denken wir uns das Zeichen | als Operationssymbol für die Funktion Φ_{15}^2 (weder noch, Nicodsche Funktion) genommen und die Definition der Ausdrücke entsprechend erweitert. Dann ist offenbar

$$\neg p \textit{ semantisch äquivalent } p \mid p,$$

$$p \wedge q \textit{ semantisch äquivalent } \neg p \mid \neg q;$$

folglich gilt

$$p \wedge q \textit{ semantisch äquivalent } (p \mid p) \mid (q \mid q).$$

Dem entspricht die logische Gleichwertigkeit der Aussagenverbindungen *A und B* und *weder* (*weder A noch A*) *noch* (*weder B noch B*).

Aufgaben. 1. Man zeige, daß auch die Funktion Φ_9^2 (Sheffersche Funktion) allein zur Erzeugung von Negation und Konjunktion genügt.

2. Unter Benutzung der Assoziativgesetze für Konjunktion, Alternative und Äquivalenz charakterisiere man jeweils die Gesamtheit aller durch eine dieser Funktionen allein erzeugbaren Wahrheitsfunktionen und folgere daraus, daß keine dieser Funktionen allein zur Erzeugung aller Wahrheitsfunktionen ausreicht.

3. Durch Vertauschen der Operationssymbole $\wedge$ und $\vee$ entstehen aus den Begriffen Elementarkonjunktion bzw. (eigentliche und uneigentliche) kanonische alternative Normalform die Begriffe *Elementaralternative* bzw. (eigentliche und uneigentliche) *kanonische konjunktive Normalform*. Man zeige, daß jede Wahrheitsfunktion auch auf genau eine Weise durch eine kanonische konjunktive Normalform dargestellt werden kann.

4. Durch systematische Anwendung von semantisch äquivalenten Umformungen kann man jeden Ausdruck schrittweise in die ihm äquivalente *aNf* überführen. Man eliminiere zunächst alle Symbole $\leftrightarrow$ und $\rightarrow$, treibe die Negationszeichen mittels der de Morganschen Regeln so weit wie möglich nach innen, kürze doppelte Negationen, verwandle den erhaltenen Ausdruck durch Anwendung des entsprechenden Distributivgesetzes in eine Alternative von Konjunktionen, ordne die erhaltenen Konjunktions- und Alternativglieder mittels Kommutativ- und Assoziativgesetzen in der gewünschten Reihenfolge, verschmelze dabei doppelt auftretende Glieder und fülle fehlende Variablen mittels der Regel

$$p \leftrightarrow p \wedge q \vee p \wedge \neg q$$

auf. Der Leser baue diese Andeutungen zu einem allgemeinen Verfahren aus und wende es auf Beispiele an.

3. Strukturen und formalisierte Sprachen

3.1. Kartesische Produkte, Relationen und Operationen

Das aus beliebigen Dingen $x_1, \ldots, x_n$ gebildete *n-Tupel* $(x_1, \ldots, x_n)$ ist in ähnlicher Weise ein gedankliches Objekt wie z. B. die Menge $\{x_1, \ldots, x_n\}$. Für den praktischen Gebrauch der n-Tupel ist nur die Voraussetzung wesentlich, daß die n-Tupel folgende Grundeigenschaft besitzen:

$$(x_1, \ldots, x_n) = (y_1, \ldots, y_n) \textit{ genau dann, wenn } x_i = y_i \ (i = 1, \ldots, n). \qquad (1)$$

Die These, daß der Mengenbegriff allein ausreicht, alle in der Mathematik vorkommenden Begriffe zu definieren, erfordert es, im jeweils gewählten System der Mengenlehre für jede natürliche Zahl $n \geqq 2$ eine n-stellige Operation zu definieren, die je n Dingen eine Menge $(x_1, \ldots, x_n)$ so zuordnet, daß (1) erfüllt ist. Dies ist stets möglich, jedoch im allgemeinen auf sehr viele verschiedene Arten, so daß jeder mengentheoretischen n-Tupel-Definition eine gewisse Willkür anhaftet.

Sind $M_1, \ldots, M_n$ beliebige Mengen ($n \geqq 2$), so sei

$$M_1 \times \cdots \times M_n := \{(x_1, \ldots, x_n) : x_1 \in M_1 \textit{ und} \ldots \textit{und } x_n \in M_n\}.$$

Die so definierte Menge $M_1 \times \cdots \times M_n$ heißt *kartesisches Produkt* (auch *Kreuzprodukt*) der Mengen $M_1, \ldots, M_n$. Insbesondere sei

$$M^1 := M, \quad M^n := \underbrace{M \times \cdots \times M}_{n\text{-mal}} \quad \textit{für } n \geqq 2.$$

Sind $M_1, \ldots, M_n$ beliebige nichtleere und paarweise disjunkte Mengen ($n \geqq 1$), so verstehen wir unter einer *Relation im Bereich der Mengen* $M_1, \ldots, M_n$ eine beliebige Teilmenge eines kartesischen Produkts der Form $M_{i_1} \times M_{i_2} \times \cdots \times M_{i_m}$, wobei $m \geqq 1$, $i_\mu \in \{1, \ldots, n\}$ für $\mu = 1, \ldots, m$ ist. Dabei heißt das m-Tulpel $(i_1, \ldots, i_m)$ von natürlichen Zahlen der *Typ* der Relation und die Zahl m die *Stellenzahl* der Relation.

Relationen im Bereich einer Menge M und nur solche sind allein durch ihre Stellenzahl charakterisiert. Eine solche Relation R ist einfach eine Teilmenge von M^m und wird kurz als *m-stellige Relation in M* bezeichnet.

Beispiel 1. Es sei M_1 die Menge aller Punkte und M_2 die Menge aller Geraden einer beliebigen fest gewählten euklidischen Ebene. Ferner sei

$$R_{inz} := \{(x_1, x_2)\colon x_1 \in M_1 \text{ und } x_2 \in M_2 \text{ und } x_1 \text{ liegt auf } x_2\},$$

$$R_{zwi} := \{(x_1, x_2, x_3)\colon x_1, x_2, x_3 \in M_1 \text{ und } x_2 \text{ liegt zwischen } x_1 \text{ und } x_3 \text{ (auf der Strecke } x_1x_3)\},$$

$$R_{kong} := \{(x_1, x_2, x_3, x_4)\colon x_1, x_2, x_3, x_4 \in M_1 \text{ und die Strecke } x_1x_2 \text{ ist kongruent der Strecke } x_3x_4\},$$

$$R_{par} := \{(x_1, x_2)\colon x_1, x_2 \in M_2 \text{ und } x_1 \text{ ist parallel zu } x_2\},$$

$$R_{orth} := \{(x_1, x_2)\colon x_1, x_2 \in M_2 \text{ und } x_1 \text{ steht senkrecht auf } x_2\}.$$

Die angegebenen fünf Mengen sind Relationen im Bereich der Mengen M_1, M_2. (In konkreten Fällen wie dem vorliegenden sagen wir z. B. auch: Relationen im Bereich der Punkte und Geraden der betrachteten Ebene.) R_{inz} ist eine zweistellige Relation vom Typ (1, 2) und heißt in der Geometrie *Inzidenzrelation*. R_{zwi} ist eine dreistellige Relation vom Typ (1, 1, 1) (d. h. eine dreistellige Relation in M_1) und heißt in der Geometrie *Zwischenrelation*. R_{kong} ist eine vierstellige M_1-Relation und wird in der Geometrie als *(Strecken-) Kongruenz* bezeichnet. R_{par} und R_{orth} sind zweistellige Relationen in M_2 (d. h. vom Typ (2, 2)).

Sind $M_{i_1}, \ldots, M_{i_m}$ irgendwelche (nicht notwendig verschiedene) der nichtleeren und paarweise disjunkten Mengen $M_1, \ldots, M_n$, so heißt eine beliebige Abbildung (Zuordnung) α, die jedem Element der Menge $M_{i_1} \times \cdots \times M_{i_m}$ einen der beiden Wahrheitswerte W, F zuordnet, ein *m-stelliges Attribut* (in manchen Büchern auch *Prädikat*) *vom Typ* $(i_1, \ldots, i_m)$ im Bereich der Mengen $M_1, \ldots, M_n$. Ist α ein Attribut vom Typ $(i_1, \ldots, i_m)$ im Bereich der Mengen $M_1, \ldots, M_n$, so ist offenbar die durch

$$R_\alpha := \{(x_1, \ldots, x_m)\colon \alpha(x_1, \ldots, x_m) = W\}$$

definierte Menge eine Relation vom Typ $(i_1, \ldots, i_m)$ im Bereich der Mengen $M_1, \ldots, M_n$. Ist umgekehrt R eine beliebige solche Relation, so ist die durch

$$\alpha_R(x_1, \ldots, x_m) := \begin{cases} W, \text{ falls } (x_1, \ldots, x_m) \in R, \\ F, \text{ falls } (x_1, \ldots, x_m) \in (M_{i_1} \times \cdots \times M_{i_m}) \setminus R, \end{cases}$$

definierte Abbildung α_R ein Attribut vom Typ $(i_1, \ldots, i_m)$ im Bereich der Mengen $M_1, \ldots, M_n$. Dabei gilt $R_{\alpha_R} = R$ und $\alpha_{R_\alpha} = \alpha$.
Demnach entsprechen Relationen und Attribute einander umkehrbar eindeutig. In gewisser Weise ist ein Attribut nichts anderes als die charakteristische Funktion

einer Relation. Beide Begriffe stellen im wesentlichen die gleiche Art von Sachverhalten dar und werden häufig nebeneinander benutzt bzw. gar nicht streng unterschieden. Wir werden im folgenden stets den Relationsbegriff benutzen.

Ist F eine Relation vom Typ $(i_1, \ldots, i_m, j)$ im Bereich der Mengen $M_1, \ldots, M_n$ mit der Eigenschaft, daß es zu jedem $(x_1, \ldots, x_m) \in M_{i_1} \times \cdots \times M_{i_m}$ höchstens ein $y \in M_j$ mit $(x_1, \ldots, x_m, y) \in F$ gibt, so heißt F eine *m-stellige partielle Operation vom Typ* $(i_1, \ldots, i_m; j)$ *im Bereich der Mengen* $M_1, \ldots, M_n$. Statt $(x_1, \ldots, x_m, y) \in F$ schreiben wir in diesem Fall $y = F(x_1, \ldots, x_m)$.

Der so definierte Begriff der (partiellen) Operation ist inhaltlich identisch mit den Begriffen Funktion bzw. Abbildung. Welche dieser Bezeichnungen jeweils benutzt wird, hängt vor allem von gewissen Traditionen ab. Insbesondere spricht man in der Algebra vorwiegend von Operationen, in der Geometrie vorzugsweise von Abbildungen usw.

Für eine partielle m-stellige Operation F sei

$$D_F := \{(x_1, \ldots, x_m) \colon \textit{es gibt ein } y \textit{ mit } (x_1, \ldots, x_m, y) \in F\}$$

der *Definitionsbereich* von F,

$$W_F := \{y \colon \textit{es gibt } x_1, \ldots, x_m \textit{ mit } (x_1, \ldots, x_m, y) \in F\}$$

der *Wertebereich* von F. Ist $D_F = M_{i_1} \times \cdots \times M_{i_m}$ für eine Operation F vom Typ $(i_1, \ldots, i_m; j)$ im Bereich der Mengen $M_1, \ldots, M_n$, so bezeichnen wir F als eine *volle* Operation. Meist werden wir im folgenden die Beiworte voll bzw. partiell fortlassen und nur in besonderen Fällen ausdrücklich darauf hinweisen, daß eine Operation voll oder (echt) partiell ist.

Beispiel 2. Es seien M_1, M_2 die in Beispiel 1 betrachteten Mengen. Ferner sei

$$F_{lin} := \{(x_1, x_2, x_3) \colon x_1, x_2 \in M_1 \textit{ und } x_3 \in M_2 \textit{ und } x_1 \neq x_2 \textit{ und } x_1 \textit{ liegt auf } x_3 \textit{ und } x_2 \textit{ liegt auf } x_3\},$$

$$F_{lot} := \{(x_1, x_2, x_3) \colon x_1 \in M_1 \textit{ und } x_2, x_3 \in M_2 \textit{ und } x_1 \textit{ liegt auf } x_3 \textit{ und } x_2 \textit{ steht senkrecht auf } x_3\}.$$

F_{lin} und F_{lot} sind dreistellige Relationen im Bereich der Mengen M_1, M_2. Zugleich ist F_{lin} eine zweistellige echt partielle Operation vom Typ $(1, 1; 2)$ und F_{lot} eine zweistellige volle Operation vom Typ $(1, 2; 2)$. Beim Übergang von der Relations- zur Operationsschreibweise bedeutet $F_{lin}(x_1, x_2)$ diejenige Gerade, die durch die beiden (als verschieden vorauszusetzenden) Punkte x_1, x_2 geht und $F_{lot}(x_1, x_2)$ diejenige Gerade, die durch den Punkt x_1 geht und auf der Geraden x_2 senkrecht steht. Offenbar ist

$$D_{F_{lin}} = \{(x_1, x_2) \colon x_1, x_2 \in M_1 \textit{ und } x_1 \neq x_2\}, \qquad W_{F_{lin}} = M_2,$$
$$D_{F_{lot}} = M_1 \times M_2, \qquad W_{F_{lot}} = M_2.$$

3.2. Strukturen

Unter einer *n-sortigen Struktur* ($n \geqq 1$) verstehen wir ein $(n + m + l + p)$-Tupel der Form

$$(M_1, \ldots, M_n; \varrho_1, \ldots, \varrho_m; \varphi_1, \ldots, \varphi_l; \gamma_1, \ldots, \gamma_p), \tag{1}$$

wobei $M_1, \ldots, M_n$ nichtleere und paarweise disjunkte Mengen, $\varrho_1, \ldots, \varrho_m$ Relationen im Bereich der Mengen $M_1, \ldots, M_n$, $\varphi_1, \ldots, \varphi_l$ partielle Operationen im Bereich dieser Mengen und $\gamma_1, \ldots, \gamma_p$ gewisse Elemente der Menge $\bigcup_{\nu=1}^{n} M_\nu$ sind. Die Mengen $M_1, \ldots, M_n$ heißen die *Grundbereiche* (auch *Grundmengen*) der Struktur. Im konkreten Fall kann jede der Zahlen m, l, p gleich Null sein. Jedoch soll $m + l > 0$ sein. Strukturen, die durch $m = 0$ gekennzeichnet sind, werden häufig als (*universelle*) *Algebren* bezeichnet. In letzter Zeit wird diese Bezeichnung auch schon gleichbedeutend mit dem hier definierten Strukturbegriff gebraucht. Strukturen mit $l = p = 0$ wurden eine Zeitlang mit dem heute nicht mehr gebräuchlichen Namen *Relativ* bezeichnet. Da jede k-stellige partielle Operation eine spezielle $(k + 1)$-stellige Relation ist und jede der Konstanten $\gamma_\varkappa$ ($\varkappa = 1, \ldots, p$) einer Struktur auch durch die einstellige Relation $\{\gamma_\varkappa\}$ beschrieben werden kann, ist der Spezialfall „Relativ" des Strukturbegriffs eigentlich genauso allgemein wie der Strukturbegriff. Man kann sich sogar auf die Betrachtung einsortiger Strukturen beschränken, indem man aus den ursprünglichen Grundbereichen $M_1, \ldots, M_n$ den einheitlichen neuen Grundbereich $M = \bigcup_{\nu=1}^{n} M_\nu$ bildet und die Mengen $M_1, \ldots, M_n$ als in ihm gegebene einstellige Relationen (sogenannte *Sortenprädikate*) auffaßt. Für die meisten gegenwärtigen Anwendungen erweist sich jedoch der Strukturbegriff in der oben eingeführten Form als am zweckmäßigsten.

Ist $\gamma_\varkappa \in M_j$, so bezeichnen wir $\gamma_\varkappa$ als eine *Konstante vom Typ* (; j). (Die Bezeichnung (j) könnte mit dem Typ einer einstelligen Relation verwechselt werden.) Aus der Gesamtheit aller Typen der zur Struktur (1) gehörigen Relationen, Operationen und Konstanten und der Sortenzahl kann man nun den *Typ der Struktur* (1) wie folgt bilden:

$$\big(n; \mathit{Typ}(\varrho_1), \ldots, \mathit{Typ}(\varrho_m); \mathit{Typ}(\varphi_1), \ldots, \mathit{Typ}(\varphi_l); \mathit{Typ}(\gamma_1), \ldots, \mathit{Typ}(\gamma_p)\big).$$

Für einsortige Strukturen läßt sich diese Angabe wieder wesentlich vereinfachen, da der Typ einer solchen Struktur bereits völlig durch die Stellenzahl der Relationen und Operationen und die Anzahl der Konstanten gekennzeichnet ist.

Beispiele.

1. Die natürlichen Zahlen bilden unter anderem die Struktur

$$\mathfrak{N} := (\mathbf{N}; \varrho_1; \varphi_1, \varphi_2; \gamma_1, \gamma_2).$$

Dabei ist $\mathbf{N}$ die Menge der natürlichen Zahlen, ϱ_1 die übliche Ordnung der Menge $\mathbf{N}$, d. h. eine zweistellige Relation in $\mathbf{N}$; φ_1 bzw. φ_2 bezeichnen die Addition bzw. Multiplikation in $\mathbf{N}$, d. h. zweistellige Operationen in $\mathbf{N}$; γ_1 bezeichnet die Zahl Null, γ_2 die Zahl Eins. Vom gleichen Typ wie die Struktur $\mathbf{N}$ ist z. B. jeder geordnete Körper. Ein solcher wird bekanntlich durch eine Grundmenge K, eine in K definierte zweistellige Ordnungsrelation, zwei in K definierte zweistellige Operationen (Addition und Multiplikation) und die neutralen Elemente bezüglich Addition und Multiplikation gegeben.

2. Bei einem axiomatischen Aufbau der ebenen euklidischen Geometrie nach dem Vorbild HILBERTS[1]) kann man euklidische Ebenen als gewisse Strukturen der Form

$$(M_1, M_2; R_{inz}, R_{zwi}, R_{kong})$$

charakterisieren, wobei die einzelnen Bestandteile die in Abschnitt 3.1., Beispiel 1, eingeführte Bedeutung haben. Demnach sind euklidische Ebenen gewisse Strukturen vom Typ (2; (1, 2), (1, 1, 1), (1, 1, 1, 1)).

3. Ein Vektorraum $\mathfrak{V}$ über einem Körper K als Skalarbereich ist eine Struktur vom Typ

$$\big(2;\ (1, 1; 1), (2, 1; 1), (2, 2; 2), (2, 2; 2); (;1), (;2), (;2)\big). \tag{2}$$

Aus (2) kann man ablesen: Es gibt zwei Grundbereiche (nämlich die Menge M_1 der Vektoren und die Menge M_2 der Skalare), keine Relationen, eine zweistellige Operation in M_1 (die Vektoraddition), eine Operation, die je einem Skalar und einem Vektor einen Vektor zuordnet (die Skalarmultiplikation), zwei zweistellige Operationen in M_2 (die Addition und Multiplikation der Skalare untereinander), schließlich einen ausgezeichneten Vektor (den Nullvektor) und zwei ausgezeichnete Skalare (Null- und Einselement des Skalarkörpers).

Um Aussagen über Strukturen eines bestimmten Typs formulieren zu können, benötigt man eine geeignete Sprache. Wesentliche Bestandteile einer solchen Sprache sind:

(a) Für Dinge jeder Sorte benötigt man einen Vorrat von Variablen, der abzählbar sein soll, um endliche Aussagen beliebiger Länge bzw. Kompliziertheit formulieren zu können.

Zum Beispiel verabredet man in der Geometrie, Punkte mit großen Buchstaben A, B, C, ..., P_0, P_1, P_2, ... (eventuell mit Indizes) und Geraden mit kleinen Buchstaben g, h, a, b, ..., g_0, g_1, g_2, ... (eventuell mit Indizes) zu bezeichnen. In der linearen Algebra benutzt man kleine lateinische Buchstaben zur Bezeichnung von Vektoren und kleine griechische Buchstaben zur Bezeichnung von Skalaren oder

[1]) Vgl. [16].

(vor allem in älteren Büchern) kleine Frakturbuchstaben für Vektoren und kleine lateinische Buchstaben für Skalare.

(b) Zur Bezeichnung einer k-stelligen Relation benötigt man eine k-stellige Aussageform. Dabei verstehen wir unter einer *k-stelligen Aussageform* ein sprachliches Gebilde mit k Variablen, das jedesmal dann, wenn für jede dieser k Variablen der Name eines bestimmten Objekts aus dem zu dieser Variablen gehörigen Grundbereich eingesetzt wird, zu einer Aussage wird.

Zum Beispiel wählt man in der Geometrie zur Bezeichnung der vierstelligen Punktrelation R_{kong} das Zeichen $\cong$ mit der Verabredung, daß sich davor und dahinter je zwei Leerstellen für Punktvariablen befinden, d. h., man bildet die Aussageformen $\mathrm{AB} \cong \mathrm{CD}$, $\mathrm{AP} \cong \mathrm{PQ}$, $\mathrm{P_1P_2} \cong \mathrm{P_1P_3}$ usw. Die Zeichenreihe $\mathrm{AB} \cong \mathrm{CD}$ wird häufig fälschlich als Aussage angesehen. Sie ist jedoch weder wahr noch falsch, solange man A, B, C und D nicht als Bezeichnungen ganz bestimmter Punkte auffaßt. Soweit A, B, C, D als Punktvariablen benutzt werden, ist die Zeichenreihe $\mathrm{AB} \cong \mathrm{CD}$ eine Aussageform im oben definierten Sinn. Andere Beispiele für Aussageformen zur sprachlichen Bezeichnung von Relationen sind

$\mathrm{g} \parallel \mathrm{h}$ zur Bezeichnung der Parallelitätsrelation R_{par},

$\mathrm{m} \leq \mathrm{n}$ zur Bezeichnung der Ordnungsrelation für natürliche Zahlen,

A auf g zur Bezeichnung der Inzidenzrelation R_{inz},

A zwischen B und C zur Bezeichnung der Zwischenrelation R_{zwi}.

(c) Für jede Operation benötigt man ein Operationssymbol (das auch aus mehreren getrennten Zeichen bestehen kann) nebst der Verabredung über die Plätze, an denen die entsprechenden Variablen stehen sollen. Für zweistellige Operationen verwendet man traditionsgemäß Zeichen wie $+$, $-$, $\cdot$, $:$, $\times$, $\circ$ mit der Verabredung, daß die Variable für das erste Argument davor und die für das zweite Argument dahinter stehen soll. Eine bei beliebiger Stellenzahl anwendbare Art, eine Operation zu kennzeichnen, ist das vorangestellte Operationszeichen, etwa F, G, $\mathfrak{P}$, mit in Klammern eingeschlossenen und durch Komma getrennten nachgestellten Variablen, beispielsweise $F(\mathrm{x}, \mathrm{y}, \mathrm{z})$, $G(\mathrm{x})$, $\mathfrak{P}(\mathrm{M})$. Wir erwähnen noch das nachgestellte Operationszeichen $'$ zur Bezeichnung der einstelligen Nachfolgeroperation im Bereich der natürlichen Zahlen, wobei also n' den Nachfolger der durch n bezeichneten Zahl darstellt. Schließlich haben auch gewisse Wortkombinationen der Umgangssprache den Charakter von Operationssymbolen, z. B.

— das Lot von P auf g (wofür man, um den Operationscharakter deutlicher zu machen, Lot(P, g) schreiben könnte),

— die Parallele zu g durch P (wofür man, auch zwecks Abkürzung, Par(g, P) oder ähnlich schreiben kann),

— das Produkt von x und y (wofür sich bereits zu einem frühen Zeitpunkt symbolische Abkürzungen, die gleichzeitig das Wesen der Sache verdeutlichen, herausbildeten).

(d) Zur Bezeichnung einer Konstanten einer Struktur benötigt man eine spezielle Zeichenreihe oder einen einzelnen Buchstaben, die natürlich von den verwendeten Variablen verschieden sein müssen.

Von dieser Art sind z. B. allgemein bekannte Symbole wie 0, 1, e (für das neutrale Element einer Gruppe). Grundsätzlich sind auch alle Eigennamen der natürlichen Sprachen wie Euklid, Nero, Goethe von dieser Art.

Im nächsten Abschnitt werden wir so allgemein wie möglich diejenige Sprache beschreiben, die durch die Wahl der in (a) bis (d) genannten Bestandteile implizit bereits eindeutig bestimmt ist. Eine solche Sprache bezeichnet man als eine *formalisierte* (*prädikatenlogische*) *Sprache*. Um die Beschreibung einer beliebigen formalisierten Sprache, die ohnehin recht kompliziert ist, nicht noch weiter zu erschweren, wollen wir dabei annehmen, daß die Bestandteile (a) bis (d) in einer normierten Form gewählt sind. Wer die Definition der formalisierten Sprachen in dieser normierten Form verstanden hat, wird sie leicht auf solche Fälle abwandeln können, in denen die Bestandteile (a) bis (d) unter Berücksichtigung von Traditionen und gedächtnisstützenden Symbolen variabler gehandhabt werden. So erweist sich, daß — eventuell nach geringen Korrekturen — jede in einem genau abgegrenzten Gebiet der Mathematik benutzte Sprache im wesentlichen eine formalisierte Sprache ist.

3.3. Formalisierte Sprachen

Eine *Basis* $\boldsymbol{B}$ einer formalisierten Sprache ist ein System $(n; \boldsymbol{B}_r, \boldsymbol{B}_f, \boldsymbol{B}_c)$, bestehend aus einer natürlichen Zahl $n \geqq 1$ (der *Sortenzahl*), einer Menge $\boldsymbol{B}_r$ von Relationssymbolen, einer Menge $\boldsymbol{B}_f$ von Operationssymbolen und einer Menge $\boldsymbol{B}_c$ von Konstantensymbolen. Jedes Relationssymbol hat die Form $R_i^{i_1,\ldots,i_k}$, wobei die obere Indexkombination als *Typ* des Symbols bezeichnet wird und den Typ der bezeichneten Relation angibt, während der untere Index zur Unterscheidung bzw. Aufzählung aller Elemente von $\boldsymbol{B}_r$ dient. Jedes Operationssymbol hat die Form $F_i^{i_1,\ldots,i_k;j}$ und jedes Konstantensymbol die Form c_i^j, wobei die unteren und oberen Indizes jeweils die analoge Rolle spielen wie bei den Relationssymbolen. Ist n der erste Bestandteil der Basis $\boldsymbol{B}$, so müssen alle als obere Indizes auftretenden natürlichen Zahlen $i_\varkappa$ bzw. j kleiner oder gleich n sein.

Ist $\boldsymbol{B}$ eine Basis, so besteht das Grundalphabet $\boldsymbol{A}$ der durch $\boldsymbol{B}$ bestimmten Sprache aus den Elementen von $\boldsymbol{B}_r \cup \boldsymbol{B}_f \cup \boldsymbol{B}_c$ sowie aus den Zeichen x_i^j $(1 \leqq j \leqq n,\ i \geqq 0)$, $=$, $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\bigwedge$, $\bigvee$, ((Klammer auf),) (Klammer zu) und , (Komma). Dabei sind die Variablen x_i^j und die Elemente von $\boldsymbol{B}_r \cup \boldsymbol{B}_f \cup \boldsymbol{B}_c$ trotz der Indizes jeweils als ein Buchstabe anzusehen. Aus der Wortmenge $W(\boldsymbol{A})$ sondern wir durch induktive Definitionen zunächst die Mengen $\boldsymbol{T}^\nu$ $(1 \leqq \nu \leqq n)$ und danach die Menge $\boldsymbol{S}$ aus. Die Elemente von $\boldsymbol{T}^\nu$ heißen *Terme der Sorte* ν und werden sich als Formeln zur Bezeichnung von Dingen der Sorte ν in denjenigen n-sortigen Strukturen erweisen, auf die sich die zu beschreibende Sprache bezieht. Die Elemente von S heißen *Ausdrücke* und werden sich als Aussageformen bzw. Aussagen in bezug auf die genannten Strukturen erweisen. Die Menge S heißt die *durch die Basis* B *definierte formalisierte prädikatenlogische Sprache* (kurz Sprache).

Definition 1.

(a) $x_i^j \in T^j \qquad (i \geqq 0,\ 1 \leqq j \leqq n)$;

(b) $c_i^j \in T^j \qquad (c_i^j \in B_c)$;

(c) *Ist* $F_i^{j_1,\ldots,j_k;j} \in B_f$ *und* $t_\varkappa \in T^{j_\varkappa}$ *für* $\varkappa = 1, \ldots, k$, *so ist*

$$F_i^{j_1,\ldots,j_k;j}(t_1, \ldots, t_k) \in T^j.$$

Definition 2. Die Variable x_i^j kommt in der Zeichenreihe $Z \in W(A)$ *vollfrei* vor, wenn x_i^j in Z vorkommt, jedoch die Zeichenreihen $\bigwedge x_i^j$ und $\bigvee x_i^j$ nicht in Z vorkommen. $H(x_i^j)$ bezeichnet stets einen Ausdruck, in dem die Variable x_i^j vollfrei vorkommt.

Definition 3.

(a) *Ist* $R_i^{j_1,\ldots,j_k} \in B_r$ *und* $t_\varkappa \in T^{j_\varkappa}$ *für* $\varkappa = 1, \ldots, k$, *so ist*

$$R_i^{j_1,\ldots,j_k}(t_1, \ldots, t_k) \in S.$$

(b) *Ist* $t_1 \in T^j$ *und* $t_2 \in T^j$ *für ein* $j \leqq n$, *so ist* $t_1 = t_2 \in S$.

(c) *Ist* $H \in S$, *so ist* $\neg H \in S$; *sind* $H_1, H_2 \in S$, *so ist* $(H_1 \wedge H_2) \in S$, $(H_1 \vee H_2) \in S$, $(H_1 \rightarrow H_2) \in S$ *und* $(H_1 \leftrightarrow H_2) \in S$.

(d) *Ist* $H(x_i^j) \in S$ *und kommt* x_i^j *vollfrei in* $H(x_i^j)$ *vor, so ist* $\bigwedge x_i^j H(x_i^j) \in S$ *und* $\bigvee x_i^j H(x_i^j) \in S$.

Der Ausdruck $\bigwedge x_i^j H(x_i^j)$ ist zu lesen: Für alle x_i^j gilt $H(x_i^j)$. Der Ausdruck $\bigvee x_i^j H(x_i^j)$ ist zu lesen: Es gibt ein x_i^j mit der Eigenschaft $H(x_i^j)$. Den Übergang von einem Ausdruck $H(x_i^j)$ zu einem der Ausdrücke $\bigwedge x_i^j H(x_i^j)$ bzw. $\bigvee x_i^j H(x_i^j)$ bezeichnet man als *Quantifizierung*, wobei man genauer im ersten Fall von *Generalisierung*, im zweiten Fall von *Partikularisierung* spricht.[1]) Der Sinn der übrigen Bestandteile der Definitionen 1 und 3 dürfte nach den vorangegangenen Erklärungen bereits anschaulich klar sein. Insbesondere vergleiche man Definition 1 mit der Definition der aussagenlogischen Ausdrücke, die vom jetzigen Standpunkt aus eigentlich Terme einer einsortigen Sprache mit den Variablen p_i und den (ein- bzw. zweistelligen)

[1]) In der Literatur sind auch andere Symbole gebräuchlich, vor allem $\forall x$ und (x) statt $\wedge x$, $\exists x$ statt $\vee x$.

Operationssymbolen $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ sind. Diese Termsprache der Aussagenlogik dient zur Verständigung über die Struktur

$$(\{W, F\}; \Phi_3^1, \Phi_8^2, \Phi_2^2, \Phi_5^2, \Phi_7^2)$$

vom Typ $\big(1; (1; 1), (1, 1; 1), (1, 1; 1), (1, 1; 1), (1, 1; 1)\big)$.

Ausdrücke der Formen $t_1 = t_2$ (*Termgleichungen*) und $R_i^{i_1, \ldots, i_k}(t_1, \ldots, t_k)$ heißen *prädikative Ausdrücke*. Mit diesen Bezeichnungen kann man der Definition 3 offenbar auch folgende verbale Fassung geben: *Die Sprache S ist die kleinste Menge von Zeichenreihen, die alle prädikativen Ausdrücke (die mittels der vorgegebenen Basis B gebildet werden können) umfaßt und bezüglich aussagenlogischer Verknüpfungen und Quantifizierungen abgeschlossen ist.*

Beispiel 1. Wir betrachten die Basis $B = (2, \{R_1^{1,2}, R_2^{1,1,1}, R_3^{1,1,1,1}\}, \emptyset, \emptyset)$. Aus $B_f = B_c = \emptyset$ folgt nach Definition 1, daß die Variablen x_i^1 und x_i^2 ($i \geqq 0$) die einzigen Terme sind. Prädikative Ausdrücke der durch diese Basis gegebenen Sprache S haben daher eine der folgenden einfachen Formen:

$$x_i^1 = x_j^1, \qquad i, j \geqq 0;$$

$$x_i^2 = x_j^2, \qquad i, j \geqq 0;$$

$$R_1^{1,2}(x_i^1, x_j^2), \qquad i, j \geqq 0;$$

$$R_2^{1,1,1}(x_i^1, x_j^1, x_k^1), \qquad i, j, k \geqq 0;$$

$$R_3^{1,1,1,1}(x_i^1, x_j^1, x_k^1, x_l^1), \qquad i, j, k, l \geqq 0.$$

Durch aussagenlogische Verknüpfung entsteht hieraus unter anderem:

$$\Big(\big(R_3^{1,1,1,1}(x_1^1, x_2^1, x_3^1, x_4^1) \wedge R_3^{1,1,1,1}(x_3^1, x_4^1, x_5^1, x_6^1)\big) \rightarrow R_3^{1,1,1,1}(x_1^1, x_2^1, x_5^1, x_6^1)\Big), \tag{1}$$

$$\big(R_1^{1,2}(x_1^1, x_1^2) \wedge R_1^{1,2}(x_2^1, x_1^2)\big). \tag{2}$$

Die Variable x_1^2 kommt im Ausdruck (2) vollfrei vor. Daher ist auch

$$\bigvee x_1^2\big(R_1^{1,2}(x_1^1, x_1^2) \wedge R_1^{1,2}(x_2^1, x_1^2)\big)$$

ein Ausdruck. Da auch die Termgleichung $x_1^1 = x_2^1$ ein Ausdruck ist, ist

$$\Big(\neg\, x_1^1 = x_2^1 \rightarrow \bigvee x_1^2\big(R_1^{1,2}(x_1^1, x_1^2) \wedge R_1^{1,2}(x_2^1, x_1^2)\big)\Big) \tag{3}$$

ebenfalls ein Ausdruck. Analog erschließt man, daß die Zeichenreihen

$$\big(\neg\, x_1^1 = x_2^1 \rightarrow \bigvee x_3^1 R_2^{1,1,1}(x_1^1, x_2^1, x_3^1)\big), \tag{4}$$

$$\bigwedge x_1^2 \bigvee x_1^1 R_1^{1,2}(x_1^1, x_1^2) \tag{5}$$

Ausdrücke sind.

Die Basis $\boldsymbol{B}$ ist zu einem bestimmten Zweck eingeführt worden, nämlich um etwas über euklidische Ebenen in formalisierter Form aussagen zu können. Dies ist möglich, wenn man die Variablen x_i^1 als Variablen für Punkte und die Variablen x_i^2 als Variablen für Geraden, das Symbol $R_1^{1,2}$ als Symbol für die Relation R_{inz}, $R_2^{1,1,1}$ als Symbol für R_{zwi} und $R_3^{1,1,1,1}$ als Symbol für R_{kong} deutet. Die Ausdrücke (1) bis (5) können dann wie folgt gelesen werden:

(1′) Wenn das Punktepaar x_1^1, x_2^1 zum Punktepaar x_3^1, x_4^1 kongruent ist und das Punktepaar x_3^1. x_4^1 zum Punktepaar x_5^1, x_6^1 kongruent ist, so ist x_1^1, x_2^1 kongruent zu x_5^1, x_6^1.

(2′) Die Gerade x_1^2 geht durch die Punkte x_1^1 und x_2^1.

(3′) Wenn die Punkte x_1^1 und x_2^1 verschieden sind, gibt es eine Gerade x_1^2, die durch x_1^1 und x_2^1 geht.

(4′) Wenn die Punkte x_1^1 und x_2^1 verschieden sind, gibt es einen Punkt x_3^1, so daß x_2^1 zwischen x_1^1 und x_3^1 liegt.

(5′) Zu jeder Geraden x_1^2 gibt es einen Punkt x_1^1, so daß x_1^1 auf x_1^2 liegt.

Beispiel 1 ist in strenger Anlehnung an die Definitionen 1 bis 3 formuliert. Es zeigt, daß eine formalisierte Sprache im Sinne dieser Definitionen recht schwer lesbar und umständlich ist. Andererseits wurde schon in 3.2. bemerkt, daß man die allgemeine Definition der formalisierten Sprachen für konkrete Fälle modifizieren kann. Wir geben nun eine solche Modifizierung für Beispiel 1 an und empfehlen dem Leser, die Definition konkreter formalisierter Sprachen an weiteren Beispielen selbständig zu üben.

Beispiel 2. Definition einer modifizierten formalisierten Sprache für die ebene euklidische Geometrie nach HILBERT.

Variablen für Punkte: große lateinische Buchstaben, eventuell mit Indizes;

Variablen für Geraden: kleine lateinische Buchstaben, eventuell mit Indizes;

(Eine besondere Termdefinition entfällt, da keine Operationen oder Konstanten unter den Hilbertschen Grundbegriffen auftreten.)

Definition der Sprache $\boldsymbol{S}_{eukl}$:

(a) *Zeichenreihen der Formen* $X = Y$, $x = y$, X auf y, $[X, Y, Z]$ (*zu lesen:* „Y liegt zwischen X und Z“) *und* $XY \cong UV$ *seien prädikative Ausdrücke der Sprache* $\boldsymbol{S}_{eukl}$, *wenn für* X, Y, U, V *beliebige Punktvariablen und für* x, y *beliebige Geradenvariablen eingesetzt werden.*

(b) *Wenn* $H \in \boldsymbol{S}_{eukl}$, *so sei auch* $\neg H \in \boldsymbol{S}_{eukl}$; *wenn* $H_1, H_2 \in \boldsymbol{S}_{eukl}$, *so sei auch* $(H_1 \circ H_2) \in \boldsymbol{S}_{eukl}$, *falls für* $\circ$ *eines der Zeichen* $\wedge, \vee, \rightarrow, \leftrightarrow$ *eingesetzt wird.*

(c) *Wenn* $H(x) \in \boldsymbol{S}_{eukl}$ *und die Punkt- oder Geradenvariable* x *vollfrei in* H(x) *vorkommt, so sei auch* $\bigwedge x H(x) \in \boldsymbol{S}_{eukl}$ *und* $\bigvee x H(x) \in \boldsymbol{S}_{eukl}$.

Die Aussageformen bzw. Aussagen (1') bis (5') lassen sich in der Sprache $\boldsymbol{S}_{eukl}$ z. B. wie folgt schreiben:

(1'') $\big((\mathrm{AB} \cong \mathrm{CD} \wedge \mathrm{CD} \cong \mathrm{EF}) \rightarrow \mathrm{AB} \cong \mathrm{EF}\big)$;

(2'') $(\mathrm{A}$ auf $\mathrm{g} \wedge \mathrm{B}$ auf $\mathrm{g})$;

(3'') $\big(\neg\, \mathrm{P} = \mathrm{Q} \rightarrow \bigvee \mathrm{g}(\mathrm{P}$ auf $\mathrm{g} \wedge \mathrm{Q}$ auf $\mathrm{g})\big)$;

(4'') $(\neg\, \mathrm{P} = \mathrm{Q} \rightarrow \bigvee \mathrm{R}[\mathrm{P}, \mathrm{Q}, \mathrm{R}])$;

(5'') $\bigwedge \mathrm{g} \bigvee \mathrm{P}\; \mathrm{P}$ auf g.

Analoges wie für die Beziehungen zwischen den Beispielen 1 und 2 gilt auch für die Definition der Terme einer formalisierten Sprache bzw. deren Modifizierung im konkreten Fall unter Benutzung traditioneller oder einprägsamer Schreibweisen. Im folgenden Beispiel geben wir die Definition einer modifizierten formalisierten Sprache mit „echten" Termen und empfehlen dem Leser, diese Definition in die durch die Definition 1 und 3 vorgeschriebene Form zu übersetzen.

Beispiel 3. Eine Sprache für die Vektorrechnung.

Skalarvariablen: kleine griechische Buchstaben, eventuell mit Indizes;
Vektorvariablen: kleine lateinische Buchstaben, eventuell mit Indizes;
Bezeichnung des Nullvektors: o
Bezeichnung der Skalarnull: 0
Bezeichnung der Skalareins: 1
Bezeichnung der Vektoraddition: $\oplus$
Bezeichnung der Skalaraddition: $+$
Bezeichnung der Multiplikation im Skalarbereich: $\cdot$
Bezeichnung der Multiplikation von Skalaren mit Vektoren: $\circ$

Definition der Menge der Skalarterme und der Menge der Vektorterme:

(a) *Alle Skalarvariablen und die Symbole* 0, 1 *sind Skalarterme*;

(b) *alle Vektorvariablen und das Symbol* o *sind Vektorterme*;

(c) *sind* t_1 *und* t_2 *Vektorterme, so ist* $(t_1 \oplus t_2)$ *ein Vektorterm*;

(d) *sind* t_1 *und* t_2 *Skalarterme, so sind* $(t_1 + t_2)$ *und* $(t_1 \cdot t_2)$ *Skalarterme*;

(e) *ist* t_1 *ein Skalarterm und* t_2 *ein Vektorterm, so ist* $(t_1 \circ t_2)$ *ein Vektorterm.*

Die Komponente $\boldsymbol{B}_r$ der Sprache der Vektorrechnung ist leer, demnach sind Termgleichungen die einzigen prädikativen Ausdrücke. Der Rest der Ausdrucksdefinition kann wörtlich aus Beispiel 2 übernommen werden. Unter Benutzung der so definierten formalisierten Sprache kann man einige bekannte Axiome der Körpertheorie bzw. Vektorrechnung z. B. wie folgt formulieren:

$$(\alpha + 0) = \alpha,$$

$$(\mathrm{a} \oplus \mathrm{o}) = \mathrm{a},$$

$$(1 \circ \mathrm{a}) = \mathrm{a},$$

$$\begin{aligned}
\big((\alpha + \beta) \cdot \gamma\big) &= \big((\alpha \cdot \gamma) + (\beta \cdot \gamma)\big), \\
\big((\alpha + \beta) \circ \mathrm{a}\big) &= \big((\alpha \circ \mathrm{a}) \oplus (\beta \circ \mathrm{a})\big), \\
\big(\alpha \circ (\mathrm{a} \oplus \mathrm{b})\big) &= \big((\alpha \circ \mathrm{a}) \oplus (\alpha \circ \mathrm{b})\big), \\
\big(\alpha \circ (\beta \circ \mathrm{a})\big) &= \big((\alpha \cdot \beta) \circ \mathrm{a}\big).
\end{aligned}$$

Der Leser überprüfe in allen Fällen den vorschriftsmäßigen Aufbau der Terme laut (a) bis (e) und verifiziere, daß auf beiden Seiten der angegebenen Gleichungen jeweils Terme der gleichen Sorte stehen. Ferner mache sich der Leser klar, daß die ungewohnt vielen Klammern den Ersatz für gewisse üblicherweise verabredete Vorrangregeln darstellen.

Wir werden im folgenden allgemein von *kanonischen* und *modifizierten formalisierten Sprachen* reden. Die ersten sind die durch Definition 1 und 3 erfaßten, die letzten solche, wie sie in den Beispielen 2 und 3 behandelt wurden. Beide Arten von formalisierten Sprachen dienen verschiedenen Zwecken. Die Definition der kanonischen formalisierten Sprachen umfaßt im Prinzip alle vorkommenden Fälle und ist in Anbetracht dessen relativ einfach. Allgemeine Untersuchungen über beliebige formalisierte Sprachen lassen sich leicht auf diese Definition beziehen und werden auch im folgenden stets darauf bezogen. Formalisierte Sprachen sind jedoch nicht nur ein Hilfsmittel für metamathematische Untersuchungen (auch wenn dies in vielen Lehrbüchern der mathematischen Logik mehr oder weniger deutlich zum Ausdruck gebracht wird). Sie dringen vielmehr in immer stärkerem Maße in die mathematische Praxis ein, zunächst als eine Art mathematischer Stenographie, wie sie heute häufig schon von Studenten der ersten Semester mehr oder weniger richtig verwendet wird. Es zeigt sich, daß die Beherrschung formalisierter Sprachen, vor allem in ihrer modifizierten Form, zu größerer Sicherheit und Exaktheit auch bei der umgangssprachlichen Formulierung mathematischer und anderer Sachverhalte verhilft. Nicht zuletzt sind mathematische und andere Aussagen nur in formalisierter Form einer automatischen Verarbeitung zugänglich, und es gibt so zahlreiche Berührungspunkte zwischen den formalisierten prädikatenlogischen Sprachen und anderen Typen formalisierter Sprachen, insbesondere Programmiersprachen, daß eine intensive Beschäftigung mit den Formalisierungsmethoden auch unter diesem Gesichtspunkt von größtem Nutzen ist.

3.4. Abgekürzter Gebrauch formalisierter Sprachen

In Analogie zu den in 2.2. verabredeten Klammersparungsregeln der Aussagenlogik vereinbaren wir hier eine Reihe von Regeln zur vereinfachten Schreibung und leichteren Lesbarkeit konkreter Ausdrücke. Zeichenreihen, die unter Verwendung dieser Regeln gebildet werden, gelten nicht als Ausdrücke einer formalisierten Sprache, sondern als Abkürzungen solcher Ausdrücke. Insbesondere beziehen sich allgemeine Betrachtungen über formalisierte Sprachen immer auf die Ausdrücke in ihrer ursprünglichen (unabgekürzten) Form, auch wenn in einigen Fällen zur bequemeren Darstellung solcher allgemeinen Betrachtungen die hier vereinbarten Abkürzungsregeln benutzt werden.

(a) Alle in 2.2. vereinbarten aussagenlogischen Klammersparungsregeln werden für prädikatenlogische Sprachen übernommen. Es kann also z. B.

$$\big((\mathrm{AB} \cong \mathrm{CD} \wedge \mathrm{CD} \cong \mathrm{EF}) \to \mathrm{AB} \cong \mathrm{EF}\big)$$

(vgl. 3.3., Beispiel 2) zu

$$\mathrm{AB} \cong \mathrm{CD} \wedge \mathrm{CD} \cong \mathrm{EF} \to \mathrm{AB} \cong \mathrm{EF}$$

und

$$\begin{aligned}&\Big(\Big(\Big(\Big(\neg \bigvee g\big((\mathrm{A \text{ auf } g} \wedge \mathrm{B \text{ auf } g}) \wedge \mathrm{C \text{ auf } g}\big) \wedge \mathrm{AB} \cong \mathrm{A_1B_1}\Big) \wedge \mathrm{AC} \cong \mathrm{A_1C_1}\Big)\\ &\wedge \mathrm{BC} \cong \mathrm{B_1C_1}\Big)\\ &\to \neg \bigvee h\big((\mathrm{A_1 \text{ auf } h} \wedge \mathrm{B_1 \text{ auf } h}) \wedge \mathrm{C_1 \text{ auf } h}\big)\Big)\end{aligned} \tag{1}$$

zu

$$\begin{aligned}&\neg \bigvee g(\mathrm{A \text{ auf } g} \wedge \mathrm{B \text{ auf } g} \wedge \mathrm{C \text{ auf } g}) \wedge \mathrm{AB} \cong \mathrm{A_1B_1} \wedge \mathrm{AC} \cong \mathrm{A_1C_1} \wedge \mathrm{BC} \cong \mathrm{B_1C_1}\\ &\to \neg \bigvee h(\mathrm{A_1 \text{ auf } h} \wedge \mathrm{B_1 \text{ auf } h} \wedge \mathrm{C_1 \text{ auf } h})\end{aligned}$$

vereinfacht werden. (Der Leser prüfe zunächst den korrekten Aufbau des Ausdrucks (1) und „lese“ diesen dann nach dem Muster der Beispiele 1 und 2 aus 3.3.)

(b) Kommen im Ausdruck H die Variablen $x_{i_1}, \ldots, x_{i_k}$ vollfrei vor, so schreiben wir $H(x_{i_1}, \ldots, x_{i_k})$ und kürzen den Ausdruck $\bigwedge x_{i_1} \bigwedge x_{i_2} \ldots \bigwedge x_{i_k} H(x_{i_1}, \ldots, x_{i_k})$ durch $\bigwedge x_{i_1} \ldots x_{i_k} H(x_{i_1}, \ldots, x_{i_k})$ ab. Analog ist die Abkürzung $\bigvee x_{i_1} \ldots x_{i_k} H(x_{i_1}, \ldots, x_{i_k})$ zu verstehen. Wir schreiben also z. B. (vgl. 3.3., Beispiel 3)

$$\begin{aligned}&\bigwedge \alpha_1\alpha_2\alpha_3\alpha_4\gamma_1\gamma_2\big(\neg\, \alpha_1 \cdot \alpha_4 = \alpha_2 \cdot \alpha_3 \to \bigvee \beta_1\beta_2(\alpha_1 \cdot \beta_1 + \alpha_2 \cdot \beta_2 = \gamma_1\\ &\wedge \alpha_3 \cdot \beta_1 + \alpha_4 \cdot \beta_2 = \gamma_2)\big).\end{aligned}$$

(c) Kommt die Variable x im Ausdruck H(x) vollfrei vor und ist y eine in H(x) nicht vorkommende, zu x sortengleiche Variable, so bezeichnet H(y) die Zeichenreihe *Sub* (H(x), x, y). Durch Induktion über die Kompliziertheit von H läßt sich leicht zeigen, daß dann H(y) ebenfalls ein Ausdruck ist, in dem die Variable y vollfrei vorkommt.

Ein Ausdruck der Form

$$\begin{aligned}&\bigvee x_{i_1} \ldots x_{i_n}\big(\neg\,(x_{i_1} = x_{i_2} \vee x_{i_1} = x_{i_3} \vee \ldots \vee x_{i_1} = x_{i_n} \vee x_{i_2} = x_{i_3} \vee \ldots\\ &\vee x_{i_2} = x_{i_n} \vee \ldots \vee x_{i_{n-1}} = x_{i_n}) \wedge H(x_{i_1}) \wedge H(x_{i_2}) \wedge \ldots \wedge H(x_{i_n})\big)\end{aligned}$$

hat offenbar die Bedeutung: *Es gibt mindestens n verschiedene Dinge* x *mit der durch den Ausdruck* H(x) *dargestellten Eigenschaft.* Derartige Ausdrücke kürzen wir im folgenden durch $\bigvee_{n} x H(x)$ ab, wobei natürlich im konkreten Fall für x eine beliebige Variable der entsprechenden Sorte aus der betrachteten Sprache stehen kann. Wir schreiben also z. B. kurz

$$\bigwedge g \bigvee_{3} P\, P \text{ auf } g$$

statt

$$\bigwedge g \bigvee P_1P_2P_3(\neg(P_1 = P_2 \vee P_1 = P_3 \vee P_2 = P_3) \wedge P_1 \text{ auf } g \wedge P_2 \text{ auf } g \wedge P_3 \text{ auf } g) \tag{2}$$

und bringen damit in formalisierter Form zum Ausdruck: Auf jeder Geraden liegen mindestens drei verschiedene Punkte. (Der Leser schreibe (2) ohne Benutzung der Abkürzungsregeln (a) und (b)!)

(d) Bei Benutzung der Abkürzungsregel (c) können wir Aussagen der Bedeutung: *Es gibt höchstens n Dinge* x *mit der Eigenschaft* H(x) bzw. *Es gibt genau n Dinge* x *mit der Eigenschaft* H(x) offenbar durch

$$\neg \bigvee_{n+1} x H(x) \quad \text{bzw.} \quad \bigvee_{n} x H(x) \wedge \neg \bigvee_{n+1} x H(x)$$

formalisiert ausdrücken. Das erste kürzen wir durch $\bigvee_{n} !x H(x)$ und das letzte durch $\bigvee_{n} !!x H(x)$ ab. Insbesondere schreiben wir $\bigvee !x H(x)$ für $\neg \bigvee_{2} x H(x)$ und $\bigvee !!x H(x)$ für $\bigvee x H(x) \wedge \bigvee !x H(x)$.

Unter Verwendung der eingeführten Abkürzungen kann man z. B. sehr einfach formulieren

$$\bigwedge AB(\neg A = B \rightarrow \bigvee !!g(A \text{ auf } g \wedge B \text{ auf } g)), \tag{3}$$

$$\bigwedge g_1g_2(\bigvee P(P \text{ auf } g_1 \wedge P \text{ auf } g_2) \wedge \neg g_1 = g_2 \rightarrow \bigvee !!P(P \text{ auf } g_1 \wedge P \text{ auf } g_2)), \tag{4}$$

$$\bigwedge AB(\neg A = B \rightarrow \bigvee_{10000} P[A, P, B]). \tag{5}$$

Der Leser verwandle die Zeichenreihen (3) und (4) in korrekte Ausdrücke der Sprache S_{eukl} (vgl. 3.3., Beispiel 2) und mache sich klar, daß diese Aufgabe für die Zeichenreihe (5) praktisch undurchführbar ist.

4. Semantische Grundbegriffe der Metamathematik

4.1. Interpretationen

Wir haben formalisierte Sprachen bisher unter dem Aspekt betrachtet, daß alle zu ihrem Aufbau nötigen Grundbegriffe (Variablensorten, Relations-, Operations- und Konstantensymbole) eine feste Bedeutung haben. Unter dieser Voraussetzung ist anschaulich klar, daß jeder Term eine ganz bestimmte Operation darstellt, deren Stellenzahl gleich der Anzahl der in ihm vorkommenden Variablen ist, wobei der Variabilitätsbereich der einzelnen Variablen durch die den Variablensorten (stillschweigend) zugeordneten Grundbereiche gegeben ist. Ebenso stellt anscheinend jeder Ausdruck eine ganz bestimmte Aussageform (vgl. 3.2., (b)) dar, d. h. im allgemeinen eine Relation, deren Stellenzahl und Typ der Anzahl und Sorte der in dem Ausdruck frei (d. h. nicht quantifiziert) vorkommenden Variablen entspricht. Insbesondere hat ein Ausdruck ohne freie Variablen dann den Charakter einer Aussage.

Der entscheidende Schritt vom bloßen praktischen Gebrauch formalisierter Sprachen zu den Grundlagen der Mathematik besteht in der Erkenntnis, daß für eine beliebige formalisierte Sprache jede mit ihrer syntaktischen Struktur (Sortenzahl, Anzahl und Typ der Relations-, Operations- und Konstantensymbole) verträgliche Deutung möglich ist und daß das Wesen der axiomatischen Methode gerade darin besteht, aus den in einer (formalisierten) Sprache gegebenen Voraussetzungen (Axiomen) solche Schlüsse zu ziehen, die bei einer beliebigen Deutung der in den Axiomen vorkommenden Grundbegriffe gerechtfertigt sind. Die folgenden Abschnitte dienen der Präzisierung dieses Konzepts. Dabei zeigt sich, daß die Formalisierung der betrachteten Sprachen unentbehrliche Voraussetzung für eine solche Präzisierung ist.

Definition 1. Die n-sortige Sprache $\boldsymbol{S}$ sei durch ihre Basis $\boldsymbol{B} = (n, \boldsymbol{B}_r, \boldsymbol{B}_f, \boldsymbol{B}_c)$ gegeben. Eine *Interpretation* ω der Sprache $\boldsymbol{S}$ ist eine Abbildung mit folgenden Eigenschaften:

(a) ω *ordnet jeder Sortennummer* j $(1 \leqq j \leqq n)$ *eine Menge* $M_j = \omega(j)$ *zu, so daß die Mengen* M_j *nichtleer und paarweise disjunkt sind.*

(b) ω *ordnet jedem Relationssymbol* $\mathrm{R}_i^{j_1,\ldots,j_k} \in \boldsymbol{B}_r$ *eine Relation* $\varrho_i = \omega\ (\mathrm{R}_i^{j_1,\ldots,j_k})$ *vom Typ* $(j_1, \ldots, j_k)$ *im Bereich der Mengen* $M_1, \ldots, M_n$ *zu.*

(c) ω *ordnet jedem Operationssymbol* $\mathrm{F}_i^{j_1,\ldots,j_k;j} \in \boldsymbol{B}_f$ *eine im allgemeinen partielle*[1]) *Operation* $\varphi_i = \omega(\mathrm{F}_i^{j_1,\ldots,j_k;j})$ *vom Typ* $(j_1, \ldots, j_k; j)$ *im Bereich der Mengen* $M_1, \ldots, M_n$ *zu.*

(d) ω *ordnet jedem Konstantensymbol* $\mathrm{c}_i^j \in \boldsymbol{B}_c$ *ein Element* $\gamma_i = \omega(\mathrm{c}_i^j) \in M_j$ *zu.*

Ist ω eine Interpretation von $\boldsymbol{S}$, so bilden die Grundbereiche $\omega(1), \ldots, \omega(n)$ zusammen mit den Relationen $\omega(\mathrm{R})$ $(\mathrm{R} \in \boldsymbol{B}_r)$, den (partiellen) Operationen $\omega(\mathrm{F})$ $(\mathrm{F} \in \boldsymbol{B}_f)$ und den Konstanten $\omega(\mathrm{c})$ $(\mathrm{c} \in \boldsymbol{B}_c)$ eine Struktur, die wir, wenn sich keine Mißverständnisse daraus ergeben, ebenfalls mit ω bezeichnen.

Beispiel 1. Wir betrachten die Basis

$$\boldsymbol{B} = (2, \{\mathrm{R}_1^{1,2}, \mathrm{R}_2^{1,1,1}, \mathrm{R}_3^{1,1,1,1}\}, \emptyset, \emptyset)$$

(vgl. 3.3., Beispiel 1) und geben folgende (von der bisherigen Deutung dieser Sprache abweichende) Interpretation ω an:

$$\omega(1) := \mathbf{R}^2, \quad \omega(2) := \{\Gamma : \Gamma \in \mathbf{R}^2 \ \textit{und es existieren}\ \alpha, \beta, \delta \in \mathbf{R},\ \textit{so daß}\ \alpha^2 + \beta^2 > 0\ \textit{und}\ (\xi, \eta) \in \Gamma\ \textit{genau dann, wenn}\ \alpha\xi + \beta\eta + \delta = 0\}.$$

(Die so definierte Menge $\omega(2)$ der „Geraden" ist offenbar eine Menge von „Punkt"-mengen.) Das Symbol $\mathrm{R}_1^{1,2}$ wird durch die Elementrelation zwischen „Punkten" und „Geraden" interpretiert, d. h.

$$\omega(\mathrm{R}_1^{1,2}) := \{(\Pi, \Gamma) : \Pi \in \omega(1) \quad \textit{und} \quad \Gamma \in \omega(2) \quad \textit{und} \quad \Pi \in \Gamma\}.$$

Ferner definieren wir:

$$\omega(\mathrm{R}_2^{1,1,1}) := \Big\{\big((\xi_1, \eta_1), (\xi_2, \eta_2), (\xi_3, \eta_3)\big) : \xi_1, \xi_2, \xi_3, \eta_1, \eta_2, \eta_3 \in \mathbf{R}\ \textit{und es existiert ein}\ \tau \in \mathbf{R}\ \textit{mit}\ 0 < \tau < 1\ \textit{und}\ \xi_2 = \xi_1 + \tau(\xi_3 - \xi_1), \quad \eta_2 = \eta_1 + \tau(\eta_3 - \eta_1)\Big\},$$

$$\omega(\mathrm{R}_3^{1,1,1,1}) := \Big\{\big((\xi_1, \eta_1), (\xi_2, \eta_2), (\xi_3, \eta_3), (\xi_4, \eta_4)\big) : \xi_1, \ldots, \xi_4, \eta_1, \ldots, \eta_4 \in \mathbf{R}\ \textit{und}\ (\xi_1 - \xi_2)^2 + (\eta_1 - \eta_2)^2 = (\xi_3 - \xi_4)^2 + (\eta_3 - \eta_4)^2\Big\}.$$

Die Struktur $\big(\omega(1), \omega(2); \omega(\mathrm{R}_1^{1,2}), \omega(\mathrm{R}_2^{1,1,1}), \omega(\mathrm{R}_3^{1,1,1,1})\big)$ bezeichnen wir als die *reelle euklidische Ebene*. Sie ist dem Leser im Prinzip aus der analytischen Geometrie bekannt.

---

[1]) Hier weichen wir im Interesse einer besseren Anpassung des Interpretationsbegriffs an die Bedürfnisse konkreter mathematischer Theorien erheblich von dem in Lehrbüchern der mathematischen Logik üblichen Vorgehen ab.

Beispiel 2. Wir geben noch eine zweite Interpretation ω' der in Beispiel 1 betrachteten Sprache an. Es sei

$$\omega'(1) := \{1, 2, 3, 4\}, \quad \omega'(2) := \{a, b, c, d, e, f\},$$

$$\omega'(R_1^{1,2}) := \{(1, a), (2, a), (2, b), (4, b), (3, c), (4, c), (1, d), (3, d), (1, e), (4, e), (2, f), (3, f)\},$$

$$\omega'(R_1^{1,1,1}) := \{(1, 2, 4), (2, 4, 3), (4, 3, 1)\},$$

$$\omega'(R_3^{1,1,1,1}) := \{(1, 2, 3, 4), (2, 1, 3, 4), (1, 2, 4, 3), (2, 1, 4, 3), (3, 4, 1, 2), (3, 4, 2, 1), (4, 3, 1, 2), (4, 3, 2, 1)\}.$$

Es ist klar, daß man die hier für kanonische formalisierte Sprachen ausgeführte Definition des Begriffs Interpretation sinngemäß auf jede modifizierte formalisierte Sprache anwenden kann und daß eine konkrete Interpretation einer konkreten Sprache in vielen Fällen ohne großen Formelaufwand mit wenigen Sätzen hinreichend exakt angegeben werden kann. So ließe sich etwa die in Beispiel 1 ausführlich definierte Interpretation wie folgt für die modifizierte Sprache $\mathbf{S}_{euH}$ (vgl. 3.3., Beispiel 2) beschreiben:

„Punkte" seien alle Paare von reellen Zahlen.

„Geraden" seien alle Mengen von „Punkten" (ξ, η), die sich durch eine Gleichung der Form $\alpha\xi + \beta\eta + \delta = 0$ beschreiben lassen, wobei $\alpha, \beta, \delta \in \mathbf{R}$ und α, β nicht gleichzeitig gleich Null sind.

Das Relationssymbol auf *werde durch die Elementrelation zwischen „Punkten" und „Geraden" interpretiert.*

Die durch [A, B, C] *symbolisierte Zwischenrelation sei für Punkte (ξ_1, η_1), (ξ_2, η_2), (ξ_3, η_3) erfüllt, wenn die Vektorgleichung*

$$\begin{pmatrix} \xi_2 \\ \eta_2 \end{pmatrix} = \tau \begin{pmatrix} \xi_3 \\ \eta_3 \end{pmatrix} + (1 - \tau) \begin{pmatrix} \xi_1 \\ \eta_1 \end{pmatrix}$$

für ein gewisses reelles τ mit $0 < \tau < 1$ erfüllt ist.

Das Zeichen $\cong$ werde durch die Äquidistanz von „Punkten" im Sinne der Abstandsformel

$$\varrho\big((\xi_1, \eta_1), (\xi_2, \eta_2)\big) := \sqrt{(\xi_1 - \xi_2)^2 + (\eta_1 - \eta_2)^2}$$

interpretiert.

Die Tatsache, daß die Terme und Ausdrücke einer formalisierten Sprache vor der Fixierung einer bestimmten Interpretation völlig sinnleere Zeichenreihen sind, wird, wie schon gesagt, häufig dadurch verschleiert, daß man bei der Konstituierung einer formalisierten Sprache bereits an eine ganz bestimmte Interpretation denkt und dies womöglich dadurch unterstreicht, daß man prädikative Sprachbestandteile wählt, die mit einer anschaulichen Bedeutung vorbelastet sind. Man vergleiche hierzu noch einmal die hinsichtlich ihrer Interpretierbarkeit gleichwertigen Sprachen aus den Beispielen 1 und 2 in 3.3. Die Sprache S_{eukl} scheint von vornherein mit einer bestimmten Interpretation versehen zu sein, während hinter der gleichwertigen kanonischen Sprache ohne den erläuternden Kommentar kaum jemand einen bestimmten Sinn vermuten würde. Der Leser mache sich klar, daß dieser scheinbare Unterschied nicht von einem Betrachter wahrgenommen wird, der zwar mit dem Inhalt der ebenen euklidischen Geometrie vertraut ist, dem aber die uns geläufigen Vereinbarungen über Punkt- und Geradenvariablen, die Bedeutung des Wortes „auf" und die Zeichen für die Kongruenz- und Zwischenrelation unbekannt sind. (Man denke etwa an einen geometrisch gebildeten „Marsmenschen"!)

Ein durch eine Interpretation ω einer Sprache S zu einer Aussageform gewordener Ausdruck $\mathrm{H} \in S$ wird durch Einsetzen je eines Elements des zugehörigen Grundbereichs für alle in ihm vorkommenden Variablen zu einer Aussage (über die durch ω gegebene Struktur). Insbesondere bezeichnet jeder Term nach Einsetzung von Objekten entsprechender Sorte für die in ihm vorkommenden Variablen ein bestimmtes Element eines der Grundbereiche der Struktur. Dies wird durch die folgenden Definitionen präzisiert, wobei wir uns wieder auf kanonische formalisierte Sprachen beschränken.

Definition 2. Ist ω eine Interpretation der n-sortigen Sprache S, so heißt eine Abbildung f, die jeder Variablen $\mathbf{x}_i^j$ ($i \geqq 0, 1 \leqq j \leqq n$) ein Ding $f(\mathbf{x}_i^j) \in \omega(j)$ zuordnet, eine *Belegung bezüglich* ω. Ist f eine solche Belegung, $\mathbf{x}_i^j$ eine beliebige Variable und $\xi \in \omega(j)$ ein beliebiges Ding der Sorte j, so bezeichnet $f\begin{pmatrix} \mathbf{x}_i^j \\ \xi \end{pmatrix}$ die wie folgt abgeänderte Belegung:

$$f\begin{pmatrix} \mathbf{x}_i^j \\ \xi \end{pmatrix}(\mathbf{x}_l^k) := \begin{cases} \xi, & \textit{falls } j = k \textit{ und } i = l, \\ f(\mathbf{x}_l^k) & \textit{sonst}. \end{cases}$$

In der folgenden Definition wird durch Induktion über die Kompliziertheit der Terme das Objekt $\mathit{Wert}(\mathrm{t}, \omega, f)$ definiert, das der Term t (eventuell) bei der Interpretation ω und der Belegung f bezüglich ω darstellt. (Da wir die Interpretation von Operationssymbolen durch partielle Operationen zulassen, existiert $\mathit{Wert}(\mathrm{t}, \omega, f)$ im allgemeinen nicht für alle t, ω und f.[1])

Definition 3.

(a) $\mathit{Wert}(\mathbf{x}_i^j, \omega, f) := f(\mathbf{x}_i^j) \quad (i \geqq 0, 1 \leqq j \leqq n)$;

(b) $\mathit{Wert}(\mathrm{c}_i^j, \omega, f) := \omega(\mathrm{c}_i^j) \quad (\mathrm{c}_i^j \in \boldsymbol{B}_\mathrm{c})$;

[1]) In manchen Zusammenhängen ist es zweckmäßig, den Interpretationsbegriff dahin zu verallgemeinern, daß $\omega(\mathrm{c})$ nicht für alle Konstantensymbole c definiert sein muß. In der Umgangssprache sind z. B. Teufel und Mann im Mond solche nicht interpretierten Konstanten. In solchen Fällen existiert $\mathit{Wert}(\mathrm{c}, \omega, f)$ für keine Belegung f bezüglich ω.

(c) *wenn* $Wert(t_\varkappa, \omega, f)$ $(= \xi_\varkappa)$ *für* $t_\varkappa \in \boldsymbol{T}^{j_\varkappa}$ *und* $\varkappa = 1, \ldots, k$ *existiert und* $F_i^{j_1,\ldots,j_k;j} \in \boldsymbol{B}_t$ *ein Operationssymbol der betrachteten Sprache ist, so ist*

$$t = F_i^{j_1,\ldots,j_k;j}(t_1, \ldots, t_k) \in \boldsymbol{T}^j$$

(nach 3.3., Definition 1). *Falls die Operation* $\varphi = \omega(F_i^{j_1,\ldots,j_k;j})$ *an der Stelle* $(\xi_1, \ldots, \xi_k)$ *definiert ist, sei*

$$Wert(t, \omega, f) := \varphi(\xi_1, \ldots, \xi_k).$$

Andernfalls (d. h., falls einer der Werte $\xi_\varkappa$ *nicht definiert ist oder* $\varphi(\xi_1, \ldots, \xi_k)$ *nicht existiert) sei* $Wert(t, \omega, f)$ *nicht definiert.*

Jeweils durch Induktion über die Kompliziertheit der Terme kann man leicht beweisen:

Folgerung 1. $Wert(t, \omega, f)$ *existiert bei festem* ω *für alle* t *und* f, *falls die Interpretation* ω *jedem Operationssymbol eine volle Operation zuordnet.*

Folgerung 2. $Wert(t, \omega, f)$ *hängt nur von den Werten von* f *für diejenigen Variablen ab, die in* t *vorkommen. Insbesondere gilt*

Folgerung 3. *Kommen in einem Term* t *keine Variablen vor* (solche Terme bezeichnen wir kurz als *variablenfrei*), *d. h., ist* t *nur aus Konstanten- und Operationssymbolen aufgebaut, so hängt* $Wert(t, \omega, f)$ *nur von* ω *ab.*

In der folgenden Definition wird durch Induktion über die Kompliziertheit der Ausdrücke der Wahrheitswert $Wert(H, \omega, f)$ definiert, den die durch den Ausdruck H bei der Interpretation ω dargestellte Aussageform bei der Belegung f annimmt. Man beachte, daß diese Definition (Teil (a)) für den Fall der Termgleichungen mit sinnleerer linker oder rechter Seite eine Normierung des in diesem Fall nicht einheitlichen mathematischen Sprachgebrauchs herbeiführt.

Definition 4.

(a) $Wert(t_1 = t_2, \omega, f) = \boldsymbol{W}$ *genau dann, wenn* $Wert(t_i, \omega, f)$ *für* $i = 1, 2$ *existiert und*

$$Wert(t_1, \omega, f) = Wert(t_2, \omega, f)$$

ist. (D. h., $Wert(t_1 = t_2, \omega, f) = \boldsymbol{F}$, *falls einer der Werte* $Wert(t_i, \omega, f)$ *oder beide nicht existieren, oder falls beide existieren, aber verschieden sind.)*

(b) $Wert(R_i^{i_1,\ldots,i_k}(t_1, \ldots, t_k), \omega, f) = \boldsymbol{W}$ *genau dann, wenn* $\xi_\varkappa = Wert(t_\varkappa, \omega, f)$ *für* $\varkappa = 1, \ldots, k$ *existiert und* $(\xi_1, \ldots, \xi_k) \in \omega(R_i^{i_1,\ldots,i_k})$ *ist.*

(c) $Wert(\neg H, \omega, f) = \Phi_3^1\big(Wert(H, \omega, f)\big)$,

$Wert\big((H_1 \wedge H_2), \omega, f\big) = \Phi_8^2\big(Wert(H_1, \omega, f), Wert(H_2, \omega, f)\big)$,

$Wert\big((H_1 \vee H_2), \omega, f\big) = \Phi_2^2\big(Wert(H_1, \omega, f), Wert(H_2, \omega, f)\big)$,

$$Wert\big((H_1 \rightarrow H_2), \omega, f\big) = \Phi_5{}^2\big(Wert(H_1, \omega, f),\ Wert(H_2, \omega, f)\big),$$
$$Wert\big((H_1 \leftrightarrow H_2), \omega, f\big) = \Phi_7{}^2\big(Wert(H_1, \omega, f),\ Wert(H_2, \omega, f)\big).$$

(d) $Wert(\bigwedge x_i{}^j H(x_i{}^j), \omega, f) = W$ *genau dann, wenn*

$$Wert\left(H(x_i{}^j), \omega, f\begin{pmatrix} x_i{}^j \\ \xi \end{pmatrix}\right) = W \quad \textit{für alle } \xi \in \omega(j);$$

$Wert(\bigvee x_i{}^j H(x_i{}^j), \omega, f) = W$ *genau dann, wenn*

$$Wert\left(H(x_i{}^j), \omega, f\begin{pmatrix} x_i{}^j \\ \xi \end{pmatrix}\right) = W \quad \textit{für wenigstens ein } \xi \in \omega(j).$$

Der Leser durchdenke diese Definition, besonders Teil (d), sorgfältig und bestätige, daß sie genau die anschauliche Vorstellung vom Zutreffen einer durch einen formalisierten Ausdruck dargestellten Aussageform auf ein gewisses System von für die Variablen eingesetzten Dingen wiedergibt.

Definition 5. Eine Variable $x_i{}^j$ *kommt in einem Ausdruck* H *an einer bestimmten Stelle frei vor*, wenn diese Stelle nicht in einem Teilausdruck von H der Form $\bigwedge x_i{}^j H'(x_i{}^j)$ oder $\bigvee x_i{}^j H'(x_i{}^j)$ enthalten ist. Eine Variable *kommt in einem Ausdruck frei vor*, wenn sie in ihm an wenigstens einer Stelle frei vorkommt. (Kommt eine Variable in einem Ausdruck vollfrei vor (vgl. 3.3., Definition 2), so kommt sie offenbar in diesem Ausdruck frei vor, und zwar an allen Stellen ihres Vorkommens.) Eine Variable *kommt in einem Ausdruck an einer bestimmten Stelle gebunden vor*, wenn sie dort vorkommt, aber an dieser Stelle nicht frei vorkommt. Ein Ausdruck, in dem keine Variable frei vorkommt, heißt *abgeschlossen*. Die Menge der abgeschlossenen Ausdrücke einer Sprache S bezeichnen wir mit $\bar{\bar{S}}$.

Beispiel 3. In dem Ausdruck

$\bigvee$ P P auf $g \wedge \underline{\text{P}}$ auf h

der Sprache S_{exkl} kommt die Variable P an der unterstrichenen Stelle frei und an den beiden nicht unterstrichenen Stellen gebunden vor. Da sie somit im Gesamtausdruck nicht vollfrei vorkommt, ist nach 3.3., Definition 1 (d) die Zeichenreihe $\bigwedge$ P($\bigvee$ P P auf g $\wedge$ P auf h) kein Ausdruck. Der Ausdruck $\bigwedge$ gh($\bigvee$ P P auf g $\wedge$ $\bigvee$ P P auf h) ist abgeschlossen.

Satz 1. *$Wert(H, \omega, f)$ hängt nur von den Werten der Belegung f für diejenigen Variablen ab, die in* H *frei vorkommen, d. h., die Stellenzahl einer durch einen interpretierten Ausdruck dargestellten Aussageform ist gleich der Anzahl der verschiedenen in ihm frei vorkommenden Variablen. Insbesondere hängt $Wert(H, \omega, f)$ für $H \in \bar{\bar{S}}$ nur von ω (und nicht von f) ab, d. h., ein interpretierter abgeschlossener Ausdruck ist eine (wahre oder falsche) Aussage.*

Wir beweisen Satz 1 durch Induktion über die Kompliziertheit von H. Die Behauptung gilt für prädikative Ausdrücke auf Grund von Folgerung 2 und Definition 4 (a), (b). Aus der Gültigkeit der Behauptung für einen Ausdruck H folgt ihre Gültigkeit für $\neg$ H, da die Ausdrücke H und $\neg$ H die gleichen freien Variablen enthalten. Nach Definition 4(c) hängt $Wert\big((H_1 \wedge H_2), \omega, f\big)$ nur von $Wert(H_1, \omega, f)$ und $Wert(H_2, \omega, f)$ ab. Setzen wir die Gültigkeit der Behauptung für die Ausdrücke H_1 und H_2 voraus, so hängen diese Werte wiederum nur von den Werten der Belegung f für diejenigen Variablen ab, die in H_1 oder in H_2 frei vorkommen. Das sind aber genau diejenigen Variablen, die im Ausdruck $(H_1 \wedge H_2)$ frei vorkommen. Analog schließt man für die übrigen aussagenlogischen Verknüpfungen. Wir setzen nun voraus, daß die Behauptung für den Ausdruck H(x) bei beliebigen Belegungen f bezüglich einer festen Interpretation ω gilt. (Hier und im folgenden lassen wir die Indizes der Variablen fort, wenn sich daraus keine Mißverständnisse ergeben.) Es sei f_0 eine beliebige Belegung bezüglich ω. $Wert\big(\bigwedge x H(x), \omega, f_0\big)$ hängt nach Definition 4(d) nur von den Werten ab, die der Ausdruck H(x) bei gewissen abgeänderten Belegungen $f_0\begin{pmatrix} x \\ \xi \end{pmatrix}$ bezüglich der Interpretation ω annimmt. Nach Voraussetzung hängen diese Werte aber nur von den Werten der Belegungen $f_0\begin{pmatrix} x \\ \xi \end{pmatrix}$ für diejenigen Variablen ab, die in H(x) frei vorkommen. Die Abhängigkeit von der Variablen x ist dabei durch (d) eliminiert. Die restlichen in H(x) frei vorkommenden Variablen sind genau die in $\bigwedge$ xH(x) frei vorkommenden Variablen, d. h., $Wert\big(\bigwedge x H(x), \omega, f_0\big)$ hängt nur von $f(y)$ für die in $\bigwedge$ xH(x) frei vorkommenden Variablen y ab. Analog schließt man für Ausdrücke der Form $\bigvee$ xH(x).

Beispiel 4. Wir beziehen uns auf die in Beispiel 2 definierte Interpretation ω' der kanonischen Sprache für die ebene euklidische Geometrie und betrachten den Ausdruck

$$\neg \bigvee x_1^1\big(R_1^{1,2}(x_1^1, x_1^2) \wedge R_1^{1,2}(x_1^1, x_2^2)\big). \tag{1}$$

In ihm kommen genau die Variablen x_1^2 und x_2^2 frei (und sogar vollfrei) vor. Demnach stellt (1) bei der Interpretation ω' eine zweistellige Aussageform dar, der eine zweistellige „Geraden"relation in der Struktur ω' entspricht. Zur Berechnung des Wahrheitswertes $Wert\big((1), \omega', f\big)$ für eine Belegung f bezüglich ω' genügt nach Satz 1 die Vorgabe von f für die Variablen x_1^2 und x_2^2. Wir setzen z. B. $f(x_1^2) := a$, $f(x_2^2) := b$. Dann ergibt sich

$$Wert\big(R_1^{1,2}(x_1^1, x_1^2), \omega', f\big) = \begin{cases} W & \textit{für} \quad f(x_1^1) = 1, 2, \\ F & \textit{für} \quad f(x_1^1) = 3, 4, \end{cases}$$

$$Wert\big(R_1^{1,2}(x_1^1, x_2^2), \omega', f\big) = \begin{cases} W & \textit{für} \quad f(x_1^1) = 2, 4, \\ F & \textit{für} \quad f(x_1^1) = 1, 3, \end{cases}$$

$$Wert\big(R_1^{1,2}(x_1^1, x_1^2) \wedge R_1^{1,2}(x_1^1, x_2^2), \omega', f\big) = \begin{cases} W & \textit{für} \quad f(x_1^1) = 2, \\ F & \textit{für} \quad f(x_1^1) = 1, 3, 4. \end{cases}$$

Folglich gibt es ein $\xi \in \omega'(2)$, nämlich $\xi = 2$, so daß

$$Wert\left(R_1^{1,2}(x_1^1, x_1^2) \wedge R_1^{1,2}(x_1^1, x_2^2), \omega', f\left\langle\begin{matrix} x_1^1 \\ \xi \end{matrix}\right\rangle\right) = \boldsymbol{W}.$$

Demnach ist

$$Wert\left(\bigvee x_1^1\left(R_1^{1,2}(x_1^1, x_1^2) \wedge R_1^{1,2}(x_1^1, x_2^2)\right), \omega', f\right) = \boldsymbol{W}. \qquad (2)$$

Hieraus ergibt sich schließlich

$$Wert\left(\neg \bigvee x_1^1\left(R_1^{1,2}(x_1^1, x_1^2) \wedge R_1^{1,2}(x_1^1, x_2^2)\right), \omega, f\right) = \boldsymbol{F}. \qquad (3)$$

Veranschaulicht man sich die Inzidenzverhältnisse der Struktur ω' durch eine Skizze mit vier beliebig gelegenen Punkten „1", „2", „3" und „4" und den sie zu je zweien verbindenden „Geraden" a bis f, so wird die Bedeutung von (2) bzw. (3) sofort anschaulich klar. Von einer solchen Skizze liest man sofort ab, daß der besprochene Ausdruck H für die übrigen Belegungen der Variablen x_1^2 und x_2^2 folgende Werte annimmt:

$f(x_1^2)$	a	b	c	d	e	f	*sonst*
$f(x_2^2)$	c	d	a	b	f	e	
$Wert(\mathrm{H}, \omega', f)$	$\boldsymbol{W}$	$\boldsymbol{W}$	$\boldsymbol{W}$	$\boldsymbol{W}$	$\boldsymbol{W}$	$\boldsymbol{W}$	$\boldsymbol{F}$

Aus dieser Tabelle kann man wiederum ablesen, daß der abgeschlossene Ausdruck $\bigwedge x_1^2 \bigvee !! x_2^2 \mathrm{H}$, dessen Wahrheitswert bei der Interpretation ω' nicht mehr von einer Belegung abhängt, den Wert $\boldsymbol{W}$ annimmt, d. h. eine in der Struktur ω' gültige Aussage darstellt.

4.2. Prädikatenlogische Allgemeingültigkeit, prädikatenlogische Normalformen

Definition 1. Es sei S eine formalisierte Sprache, ω eine Interpretation von S. Ein Ausdruck $\mathrm{H} \in S$ heißt *allgemeingültig bei der Interpretation* ω (bzw. *allgemeingültig in der durch* ω *definierten Struktur*), wenn $Wert(\mathrm{H}, \omega, f) = \boldsymbol{W}$ für alle Belegungen f bezüglich ω gilt.

Aus den letzten Bemerkungen des vorigen Abschnitts folgt: Der Ausdruck

$$\bigwedge x_1^2 \bigvee !! x_2^2 \neg \bigvee x_1^1\left(R_1^{1,2}(x_1^1, x_1^2) \wedge R_1^{1,2}(x_1^1, x_2^2)\right)$$

ist bei der Interpretation ω' allgemeingültig. Dieses Resultat wurde auf Grund der Endlichkeit der Grundbereiche der Struktur ω' durch Berechnen von $Wert(\mathrm{H}, \omega', f)$ für alle möglichen Belegungen der in H frei vorkommenden Variablen erzielt. Es ist klar, daß dieses Verfahren bei unendlichen Grundbereichen nicht anwendbar ist.

Definition 2. Ein Ausdruck $\mathrm{H} \in S$ heißt *allgemeingültig*, wenn $Wert(\mathrm{H}, \omega, f) = \boldsymbol{W}$ bei jeder Interpretation ω der Sprache S und jeder Belegung f bezüglich ω ist, d. h., wenn H bei jeder Interpretation allgemeingültig ist. Die Menge aller allgemeingültigen Ausdrücke einer Sprache S bezeichnen wir mit ag_S.

Allgemeingültigkeit eines Ausdrucks H im Sinne der Definition 2 bedeutet, daß H ohne irgendwelche inhaltlichen Voraussetzungen (über die Bedeutung der in ihm vorkommenden Variablen, Relations-, Operations- und Konstantensymbole), d. h. allein auf Grund seiner sprachlichen Struktur stets wahr ist. Unter den allgemeingültigen Ausdrücken einer Sprache S kommen insbesondere diejenigen vor, die schon auf Grund ihrer aussagenlogischen Struktur allgemeingültig sind. Ist z. B. $\mathrm{H} \in S$ ein beliebiger Ausdruck, so ist $(\mathrm{H} \vee \neg \mathrm{H}) \in ag_S$. Allgemein gilt: Ist A ein allgemeingültiger Ausdruck der Aussagenlogik, der höchstens die Variablen $\mathrm{p}_1, \ldots, \mathrm{p}_n$ enthält (für hinreichend großes n ist dies stets erfüllt) und ist jeder dieser Variablen p_i ein Ausdruck $\mathrm{H}_i \in S$ zugeordnet, so ist

$$Sub(\mathrm{A}; \mathrm{p}_1, \mathrm{H}_1; \ldots; \mathrm{p}_n, \mathrm{H}_n) \in ag_S.$$

Ein typisches Beispiel eines Ausdrucks, dessen prädikatenlogische Allgemeingültigkeit sich nicht aussagenlogisch begründen läßt, ist jeder Ausdruck der Form

$$\bigwedge \mathrm{xH}(\mathrm{x}) \to \bigvee \mathrm{xH}(\mathrm{x}), \tag{1}$$

wobei für H(x) ein beliebiger Ausdruck gewählt werden kann, der eine Variable x vollfrei enthält. Der exakte Beweis der (anschaulich einleuchtenden) Allgemeingültigkeit von (1) ergibt sich wie folgt: Ist $Wert(\bigwedge \mathrm{xH}(\mathrm{x}), \omega, f) = \boldsymbol{F}$ (bei beliebigem ω und f), so ist wegen des Wertverlaufs der Implikation $Wert((1), \omega, f) = \boldsymbol{W}$. Ist $Wert(\bigwedge \mathrm{xH}(\mathrm{x}), \omega, f) = \boldsymbol{W}$, so ist (vgl. 4.1., Definition 4 (d))

$$Wert\left(\mathrm{H}(\mathrm{x}), \omega, f\begin{pmatrix}\mathrm{x}\\ \xi\end{pmatrix}\right) = \boldsymbol{W} \quad \text{für alle } \xi \in \omega(j),$$

wobei j die Sorte der Variablen x ist. Da nach Definition des Interpretationsbegriffs die Grundbereiche $\omega(j)$ stets nichtleer sind, gibt es ein $\xi \in \omega(j)$, und für dieses ist

$$Wert\left(\mathrm{H}(\mathrm{x}), \omega, f\begin{pmatrix}\mathrm{x}\\ \xi\end{pmatrix}\right) = \boldsymbol{W},$$

d. h. (wiederum nach 4.1., Definition 4)

$$Wert\left(\bigvee \mathrm{xH}(\mathrm{x}), \omega, f\right) = \boldsymbol{W}.$$

Es ist ohne weiteres klar, daß (1) bei Zulassung von Interpretationen mit leeren Grundbereichen nicht allgemeingültig wäre.

Definition 3. Ausdrücke H_1, H_2 einer Sprache S heißen *semantisch äquivalent*, wenn $H_1 \leftrightarrow H_2 \in ag_S$, d. h., sie sind genau dann semantisch äquivalent, wenn $Wert(H_1, \omega, f) = Wert(H_2, \omega, f)$ für jede Interpretation ω und jede Belegung f bezüglich ω gilt.

Die Untersuchung der Allgemeingültigkeit und semantischen Äquivalenz von prädikatenlogischen Ausdrücken ist Gegenstand der Prädikatenlogik. Der entscheidende Unterschied gegenüber der Aussagenlogik besteht darin, daß beide Fragen nicht durch Aufstellen endlicher Wertetabellen und, wie sich zeigen wird (vgl. Kapitel 8), überhaupt nicht durch ein auf beliebige Ausdrücke beliebiger Sprachen einheitlich anwendbares endliches Verfahren beantwortet werden können.

Im folgenden geben wir eine Übersicht über häufig benutzte semantisch äquivalente Umformungen an prädikatenlogischen Ausdrücken. Wir formulieren sie jeweils als Regeln der Form $S_1 \leftrightarrow S_2$, wobei S_1 und S_2 Schemata von Ausdrücken sind. Eine Regel $S_1 \leftrightarrow S_2$ bedeutet: Bei jeder Einsetzung konkreter Ausdrücke der durch S_1 bzw. S_2 beschriebenen Bauart für S_1 und S_2 wird $S_1 \leftrightarrow S_2$ ein allgemeingültiger Ausdruck, d. h., beide Seiten sind semantisch äquivalent. Die praktische Anwendung dieser Regeln beruht auf dem (zu 2.2., Satz 4, analogen)

Ersetzbarkeitstheorem. *Sind H_0, H_1, H_2 Ausdrücke einer gemeinsamen Sprache S, ist* $H_1 \leftrightarrow H_2 \in ag_S$ *(d. h.* H_1 *semantisch äquivalent zu* H_2*), enthält* H_0 *den Teilausdruck* H_1 *und ist die Zeichenreihe, die durch Ersetzen irgendeines Vorkommens von* H_1 *in* H_0 *durch* H_2 *entsteht, wieder ein Ausdruck* H_3*, so ist*

$$H_0 \leftrightarrow H_3 \in ag_S.$$

Der Beweis dieses Satzes (durch Induktion über die Kompliziertheit von H_0) sei dem Leser überlassen. Wir bemerken nur, daß die als Resultat der Ersetzung (nicht notwendig Substitution!) entstehende Zeichenreihe H_3 immer dann ein Ausdruck ist, falls nicht eine im eingesetzten Ausdruck H_2 schon gebundene Variable durch das Einsetzen von H_2 in H_0 nochmals gebunden wird.

Die „Richtigkeit" der folgenden Regeln, d. h. die Allgemeingültigkeit der durch sie beschriebenen Ausdrücke, ist in allen Fällen ebenso unmittelbar einleuchtend wie bei (1). Auch die exakten Beweise sind nicht schwerer als der Beweis von (1) und seien dem Leser als Übung empfohlen.

Regel der gebundenen Umbenennung. Voraussetzung: Die Variablen x und y sind sortengleich, y kommt nicht in H(x) vor, H(y) bezeichnet wie früher $Sub(H(x), x, y)$. Dann ist

$$\bigwedge xH(x) \leftrightarrow \bigwedge yH(y), \qquad \bigvee xH(x) \leftrightarrow \bigvee yH(y) \tag{2}$$

allgemeingültig.

Durch eine geeignete Umbenennung gebundener Variablen kann man also stets erreichen, daß alle in einem Ausdruck frei vorkommenden Variablen in diesem Ausdruck vollfrei vorkommen: Kommt x in H frei, jedoch nicht vollfrei vor, so existieren Teilausdrücke von H, die die Form $\bigwedge x H(x)$ oder $\bigvee x H(x)$ haben. Wenn man schrittweise in jedem derartigen Teilausdruck x durch eine im Gesamtausdruck noch nicht vorkommende Variable gleicher Sorte ersetzt, erhält man sclie ßlich einen zum ursprünglichen Ausdruck H semantisch äquivalenten Ausdruck, in dem x vollfrei vorkommt. In vielen Fällen ist es zweckmäßig, sich auf die Betrachtung solcher Ausdrücke zu beschränken, in denen alle frei vorkommenden Variablen vollfrei vorkommen, d. h. in denen diejenigen Variablen, von denen der Wert des Ausdrucks bei einer Interpretation und Belegung nur abhängt, zugleich die in diesem Ausdruck quantifizierbaren Variablen sind. Ist $H(x_1, \ldots, x_n)$ ein Ausdruck, der die Variablen $x_1, \ldots, x_n$ vollfrei und keine weiteren Variablen frei enthält, so bezeichnen wir den Ausdruck

$$\bigwedge x_1 \ldots x_n H(x_1, \ldots, x_n)$$

als *Generalisierte* des Ausdrucks $H(x_1, \ldots, x_n)$ (kurz *Gen* H). Offenbar ist *Gen* H stets abgeschlossen, und bei einer beliebigen Interpretation ω der betreffenden Sprache ist H genau dann allgemeingültig, wenn *Gen* H allgemeingültig ist.[1]) Dies präzisiert die anschaulich klare Behauptung, daß man sich bei der Formulierung von Sachverhalten grundsätzlich auf abgeschlossene Ausdrücke beschränken kann. Umgekehrt kann man nichtabgeschlossene Ausdrücke, wo sie als Formulierung von Aussagen (z. B. Axiomen) und nicht zur Darstellung von Relationen vorkommen, immer als Abkürzung ihrer Generalisierten auffassen.

Regeln der Quantorenvertauschung.

$$\bigwedge x \bigwedge y H(x, y) \leftrightarrow \bigwedge y \bigwedge x H(x, y), \tag{3}$$

$$\bigvee x \bigvee y H(x, y) \leftrightarrow \bigvee y \bigvee x H(x, y). \tag{4}$$

(3) und (4) liefern die Begründung für die in der mathematischen Umgangssprache (und auch beim Lesen formalisierter Ausdrücke) üblichen Redeweisen „Für alle x und y gilt ...“ (statt „Für alle x gilt, daß für alle y gilt: ...“) bzw. „Es gibt ein x und ein y, so daß ...“ (statt „Es gibt ein x, so daß es ein y gibt, so daß ...“). Zwei verschiedene Quantoren dürfen jedoch nicht vertauscht werden. Daß $\bigwedge x \bigvee y H(x, y)$ nicht semantisch äquivalent zu $\bigvee y \bigwedge x H(x, y)$ ist, mache der Leser sich am unterschiedlichen Sinn der beiden Aussagen $\bigwedge n \bigvee m\, n < m$ bzw. $\bigvee m \bigwedge n\, n < m$ klar, von denen im Bereich der natürlichen Zahlen die erste richtig und die zweite falsch ist.

[1]) H und *Gen* H sind jedoch nicht etwa semantisch äquivalent, da $H(x) \leftrightarrow \bigwedge x H(x)$ nicht allgemeingültig ist.

Regeln der Quantorenverschiebung.

$$\bigwedge x \neg H(x) \leftrightarrow \neg \bigvee x H(x), \tag{5}$$

$$\bigvee x \neg H(x) \leftrightarrow \neg \bigwedge x H(x), \tag{6}$$

$$\bigwedge x \big(H_1(x) \wedge H_2(x)\big) \leftrightarrow \bigwedge x H_1(x) \wedge \bigwedge x H_2(x), \tag{7}$$

$$\bigvee x \big(H_1(x) \vee H_2(x)\big) \leftrightarrow \bigvee x H_1(x) \vee \bigvee x H_2(x), \tag{8}$$

$$\bigwedge x \big(H_1(x) \wedge H_2\big) \leftrightarrow \bigwedge x H_1(x) \wedge H_2, \tag{9}$$

$$\bigvee x \big(H_1(x) \wedge H_2\big) \leftrightarrow \bigvee x H_1(x) \wedge H_2. \tag{10}$$

(9) und (10) sind allgemeingültig, falls die Variable x nicht in H_2 vorkommt. Die Bezeichnung „Regeln der Quantorenverschiebung" ist so zu verstehen, daß die Regeln (5) bis (10) — je nachdem, in welcher Richtung man sie anwendet — es gestatten, den *Wirkungsbereich eines Quantors* (d. h. denjenigen Teilausdruck, vor dem $\bigwedge x$ bzw. $\bigvee x$ steht) zu vergrößern oder zu verkleinern. (5) und (6) sind in gewissem Sinn Verallgemeinerungen der de Morganschen Regeln (vgl. 2.2.). Durch Übergang zur Negation auf beiden Seiten ergibt sich

$$\neg \bigwedge x \neg H(x) \leftrightarrow \bigvee x H(x), \tag{5'}$$

$$\neg \bigvee x \neg H(x) \leftrightarrow \bigwedge x H(x), \tag{6'}$$

d. h., man kann wahlweise entweder die Partikularisierung durch die Generalisierung oder umgekehrt ausdrücken. Es ist also einer der beiden Quantoren überzählig im gleichen Sinn, wie einige der verwendeten aussagenlogischen Grundfunktionen entbehrlich sind. Hier wie dort dient die Aufnahme der grundsätzlich entbehrlichen Sprachbestandteile der besseren Anpassung der formalisierten Sprachen an die mathematische Umgangssprache.

Definition 4. Ein Ausdruck der Form $Q_1x_1Q_2x_2 \ldots Q_nx_nH(x_1, \ldots, x_n)$, wobei $Q_i \in \{\bigwedge, \bigvee\}$ $(i = 1, \ldots, n)$ und der Ausdruck $H(x_1, \ldots, x_n)$ keine Quantoren enthält, heißt eine *pränexe Normalform.*[1]) Der Ausdruck $H(x_1, \ldots, x_n)$ heißt *Kern* der Normalform, der Bestandteil $Q_1x_1 \ldots Q_nx_n$ wird als *Präfix* der Normalform bezeichnet.

Satz 1. *Zu jedem Ausdruck gibt es eine semantisch äquivalente pränexe Normalform.*

Zum Beweis von Satz 1 beschreiben wir ein Verfahren, durch das man jeden Ausdruck in endlich vielen Schritten durch semantisch äquivalente Umformungen der oben behandelten Art in eine pränexe Normalform überführen kann. Zunächst können wir den gegebenen Ausdruck unter Benutzung aussagenlogischer Äquivalenzen so umformen, daß die Symbole $\leftrightarrow$, $\rightarrow$, $\vee$ nicht mehr vorkommen. Danach handelt es sich darum, alle vorkommenden Quantoren so weit wie möglich nach „außen"

[1]) Insbesondere ist also jeder quantorenfreie Ausdruck selbst eine pränexe Normalform.

zu treiben. Dies geschieht mit Hilfe der (von rechts nach links angewandten) Regeln (5), (6), (9) und (10). Ist die Voraussetzung für die Anwendung von (9) bzw. (10) zunächst nicht erfüllt, so kann man dies durch Anwendung der gebundenen Umbenennung erreichen.

Beispiel. Wir betrachten eine Sprache mit Variablen x, y, z und Relationssymbolen R (einstellig) und S (zweistellig) und geben uns den Ausdruck

$$\bigwedge xR(x) \to \bigvee x \bigwedge yS(x, y)$$

vor. Im ersten Schritt wird $\to$ eliminiert:

$$\neg\left(\underline{\bigwedge x}R(x) \wedge \neg \bigvee x \bigwedge yS(x, y)\right).$$

Um den Quantor an der unterstrichenen Stelle aus der Klammer herausziehen zu können, wird der entsprechende Teilausdruck durch gebundene Umbenennung verwandelt:

$$\neg\left(\bigwedge zR(z) \wedge \neg \bigvee x \bigwedge yS(x, y)\right).$$

Anwendung von (9) ergibt jetzt

$$\neg \bigwedge z\left(R(z) \wedge \neg \bigvee x \bigwedge yS(x, y)\right).$$

Die weiteren Schritte sind

$$\begin{array}{ll} \neg \bigwedge z\left(R(z) \wedge \bigwedge x \neg \bigwedge yS(x, y)\right) & \text{(Regel (5))}, \\ \neg \bigwedge z\left(R(z) \wedge \bigwedge x \bigvee y \neg S(x, y)\right) & \text{(Regel (6))}, \\ \neg \bigwedge z \bigwedge x \bigvee y\left(R(z) \wedge \neg S(x, y)\right) & \text{(Regeln (9), (10))}, \\ \bigvee z \bigvee x \bigwedge y \neg\left(R(z) \wedge \neg S(x, y)\right) & \text{(Regeln (5); (6))}. \end{array}$$

Der zuletzt erhaltene Ausdruck ist bereits eine pränexe Normalform, aber man kann dem Kern wieder eine aussagenlogisch übersichtlichere Gestalt geben:

$$\bigvee z \bigvee x \bigwedge y\left(R(z) \to S(x, y)\right).$$

In der Prädikatenlogik spielen neben den pränexen Normalformen verschiedene andere Normalformen (u. a. pränexe Normalformen mit speziellen Präfixen und sogenannte kontrapränexe Normalformen, bei denen alle Quantoren möglichst weit „innen" stehen) eine große Rolle. Oft läßt sich die Überführung eines beliebigen Ausdrucks in eine zu ihm äquivalente Normalform einer bestimmten Art nur für spezielle Sprachen erreichen, oder man muß die semantische Äquivalenz durch eine schwächere Äquivalenzrelation zwischen Ausdruck und Normalform ersetzen. Zum Beispiel wird verlangt: Der gegebene Ausdruck soll genau dann allgemeingültig sein, wenn die ihm zugeordnete (eventuell in einer anderen Sprache formulierte) Normalform allgemeingültig ist.[1]) Für weitere Informationen über prädikatenlogische Normalformen und ihre Anwendungen sei auf [3, § 9] verwiesen.

4.3. Modelle, Folgern, semantische Widerspruchsfreiheit, Unabhängigkeit und Vollständigkeit elementarer Theorien

Definition 1. Es sei S eine formalisierte Sprache, $A \subseteqq S$ eine beliebige (eventuell leere) Menge von Ausdrücken von S. Eine Interpretation ω von S (und auch die durch ω definierte Struktur) heißt ein *Modell von A*, wenn alle Ausdrücke $\mathrm{H} \in A$ bei der Interpretation ω (bzw. in der Struktur ω) allgemeingültig sind. (Insbesondere ist jede Interpretation von S ein Modell von $\emptyset \subset S$.)

Beispiel 1. Es sei S die durch Variablen a, b, c, ..., $\mathrm{x}_0, \mathrm{x}_1, \mathrm{x}_2, \ldots$ und ein zweistelliges Operationssymbol $+$ gegebene einsortige modifizierte Sprache. In dieser Sprache lassen sich die aus der Algebra bekannten Gruppenaxiome z. B. wie folgt formulieren:

$$(\mathrm{a} + \mathrm{b}) + \mathrm{c} = \mathrm{a} + (\mathrm{b} + \mathrm{c}), \tag{1}$$

$$\bigwedge \mathrm{ab} \bigvee \mathrm{c}\; \mathrm{a} + \mathrm{b} = \mathrm{c}, \tag{2}$$

$$\bigwedge \mathrm{ac} \bigvee \mathrm{b}\; \mathrm{a} + \mathrm{b} = \mathrm{c}, \tag{3}$$

$$\bigwedge \mathrm{bc} \bigvee \mathrm{a}\; \mathrm{a} + \mathrm{b} = \mathrm{c}. \tag{4}$$

((2) ist nötig, da wir generell die Interpretation von Operationssymbolen durch partielle Operationen zulassen und solche Interpretationen bei einer Definition des Gruppenbegriffs ausschließen müssen.) Eine *Gruppe*, in der Algebra üblicherweise definiert als „ein System $(G, +)$, wobei G eine beliebige nichtleere Menge und $+$ eine zweistellige Operation in G mit den Eigenschaften (1) bis (4) ist", ist nun nichts anderes als ein Modell der Menge {(1), (2), (3), (4)}, d. h. eine Interpretation ω der Sprache S, bei der die Ausdrücke (1) bis (4) allgemeingültig sind. Eine solche Interpretation besteht im Fall der Sprache S aus der Angabe des Grundbereichs und der Interpretation des Symbols $+$ durch eine zweistellige (zunächst eventuell partielle) Operation in diesem Grundbereich. Wie man sieht, besteht der einzige wesentliche Unterschied zwischen der „alten" und der „neuen" Definition des Gruppenbegriffs in einer sorgfältigeren Unterscheidung zwischen dem Operationssymbol $+$ und der

---

[1]) In diesem Sinne ist jeder aus einem Ausdruck H durch gebundene Umbenennungen und Generalisierungen entstehende abgeschlossene Ausdruck $\overline{\mathrm{H}}$ zu H allgemeingültigkeitsgleich, jedoch nicht semantisch äquivalent (vgl. Fußnote auf S. 59).

durch dieses Symbol dargestellten Operation. Völlig analog kann man die aus der Algebra bekannten Begriffe *Ring*, *Körper*, *Vektorraum* usw. neu definieren.

Definition 2. Es sei $\boldsymbol{S}$ eine formalisierte Sprache, $\boldsymbol{A} \subseteqq \boldsymbol{S}$, $\mathrm{H} \in \boldsymbol{S}$. H *folgt aus* $\boldsymbol{A}$, wenn H in jedem Modell von $\boldsymbol{A}$ allgemeingültig ist. (Insbesondere folgt H aus $\emptyset$, wenn H bei jeder Interpretation von $\boldsymbol{S}$ allgemeingültig ist; vgl. Definition 1.) Die Menge aller Ausdrücke H, die aus einer Menge $\boldsymbol{A} \subseteqq \boldsymbol{S}$ folgen, nennen wir die *Folgerungshülle von* $\boldsymbol{A}$ und bezeichnen sie mit $\boldsymbol{Fl_S}(\boldsymbol{A})$. Es gilt also $\mathrm{H} \in \boldsymbol{Fl_S}(\boldsymbol{A})$ genau dann, wenn H in allen Modellen von $\boldsymbol{A}$ allgemeingültig ist, insbesondere ist $\boldsymbol{Fl_S}(\emptyset) = \boldsymbol{ag_S}$.

Die durch das Wort *folgt* ausgedrückte Relation zwischen Ausdrücken und Ausdrucksmengen einer formalisierten Sprache geht in der in Definition 2 präzisierten Form auf den polnischen Logiker Alfred Tarski (geb. 1901)[1]) zurück und erweist sich als der zentrale Begriff im Bereich der semantischen Aspekte der Metamathematik. (Unter semantischen Aspekten verstehen wir hier mit Tarski die Beziehungen zwischen formalisierten Sprachen und ihren möglichen Bedeutungen, während alle rein formalen, vom Interpretationsbegriff unabhängigen Betrachtungen der mathematischen Logik unter der Bezeichnung Syntax zusammengefaßt werden.)

Satz 1. *Bei beliebiger Sprache* $\boldsymbol{S}$ *ist die in der Potenzmenge von* $\boldsymbol{S}$ *definierte einstellige Operation* $\boldsymbol{Fl_S}$ *ein Hüllenoperator, d. h.*

(a) $\boldsymbol{A} \subseteqq \boldsymbol{Fl_S}(\boldsymbol{A})$ *für* $\boldsymbol{A} \subseteqq \boldsymbol{S}$;

(b) *wenn* $\boldsymbol{A} \subseteqq \boldsymbol{B} \subseteqq \boldsymbol{S}$, *so ist* $\boldsymbol{Fl_S}(\boldsymbol{A}) \subseteqq \boldsymbol{Fl_S}(\boldsymbol{B})$;

(c) $\boldsymbol{Fl_S}(\boldsymbol{Fl_S}(\boldsymbol{A})) = \boldsymbol{Fl_S}(\boldsymbol{A})$ *für alle* $\boldsymbol{A} \subseteqq \boldsymbol{S}$.

Beweis. (a) Zu zeigen ist, daß jeder Ausdruck $\mathrm{H} \in \boldsymbol{A}$ in jedem Modell von $\boldsymbol{A}$ allgemeingültig ist. Dies folgt sofort aus Definition 1.

(b) Ist $\boldsymbol{A} \subseteqq \boldsymbol{B}$, so ist jedes Modell von $\boldsymbol{B}$ erst recht ein Modell von $\boldsymbol{A}$. Gilt folglich H in jedem Modell von $\boldsymbol{A}$, so gilt H in jedem Modell von $\boldsymbol{B}$.

(c) Zu zeigen ist wegen (a) nur die Inklusion $\boldsymbol{Fl_S}(\boldsymbol{Fl_S}(\boldsymbol{A})) \subseteqq \boldsymbol{Fl_S}(\boldsymbol{A})$. Es sei $\mathrm{H} \in \boldsymbol{Fl_S}(\boldsymbol{Fl_S}(\boldsymbol{A}))$, d. h., H gilt in jedem Modell von $\boldsymbol{Fl_S}(\boldsymbol{A})$. Aus Definition 1 folgt sofort, daß jedes Modell von $\boldsymbol{A}$ ein Modell von $\boldsymbol{Fl_S}(\boldsymbol{A})$ ist. Daher gilt H in jedem Modell von $\boldsymbol{A}$, d. h. $\mathrm{H} \in \boldsymbol{Fl_S}(\boldsymbol{A})$.

Definition 3. Eine Teilmenge $\boldsymbol{A}$ einer formalisierten Sprache $\boldsymbol{S}$ heißt *deduktiv abgeschlossen*, wenn $\boldsymbol{Fl_S}(\boldsymbol{A}) = \boldsymbol{A}$ gilt.

Unter Benutzung von Definition 3 kann man Satz 1 (c) offenbar auch wie folgt formulieren: *Die Folgerungshülle einer beliebigen Menge von Ausdrücken ist deduktiv abgeschlossen.*

---

[1]) Vgl. hierzu die in [4] wieder abgedruckten (sonst schwer zugänglichen) Arbeiten von Tarski. Dem Sinne nach war das Folgern bereits von Bernard Bolzano (1781—1848) richtig definiert und als zentraler Begriff der Logik erkannt worden. Vgl. Kapitel 9.

Definition 4. Eine *elementare Theorie* ist eine deduktiv abgeschlossene Menge von Ausdrücken einer formalisierten Sprache.

Nach Definition 4 wird eine elementare Theorie insbesondere durch eine formalisierte Sprache S und ein in dieser Sprache formuliertes Axiomensystem A gegeben. Die Theorie ist in diesem Fall gleich der Menge $Fl_S(A)$ aller Folgerungen, die man aus den gegebenen Axiomen ziehen kann. Unter einem *Axiomensystem* verstehen wir bis auf weiteres eine beliebige Teilmenge A einer formalisierten Sprache S. Zwei Axiomensysteme $A, B \subseteqq S$ definieren die gleiche elementare Theorie genau dann, wenn $Fl_S(A) = Fl_S(B)$ gilt. Wir zeigen unter alleiniger Voraussetzung der Hülleneigenschaften (a), (b), (c) des Folgerungsoperators Fl_S

Satz 2. *Für $A, B \subseteqq S$ gilt $Fl_S(A) = Fl_S(B)$ genau dann, wenn $A \subseteqq Fl_S(B)$ und $B \subseteqq Fl_S(A)$ ist.*

Beweis. Es sei $Fl_S(A) = Fl_S(B)$. Dann ist nach (a) $A \subseteqq Fl_S(B)$ und $B \subseteqq Fl_S(A)$. Es sei umgekehrt $A \subseteqq Fl_S(B)$ und $B \subseteqq Fl_S(A)$ vorausgesetzt. Dann ist

$$Fl_S(A) \underset{(b)}{\subseteqq} Fl_S\big(Fl_S(B)\big) \underset{(c)}{\subseteqq} Fl_S(B)$$

und ebenso $Fl_S(B) \subseteqq Fl_S(A)$, d. h. $Fl_S(A) = Fl_S(B)$.

(Satz 2 liefert die Rechtfertigung für die in der mathematischen Praxis intuitiv richtig gehandhabte Methode, die „Gleichwertigkeit" zweier Axiomensysteme zu beweisen. Die Beweise der Sätze 1 und 2 enthüllen zugleich, wie einfach sich viele in den Anwendungen weitreichende Sätze der Metamathematik beweisen lassen, sobald man über exakte Definitionen der in ihnen vorkommenden Begriffe verfügt. Wir werden im folgenden ähnlich triviale Beweise häufig dem Leser überlassen.)

Die wichtigsten Eigenschaften eines Axiomensystems und der hierdurch definierten elementaren Theorie sind Widerspruchsfreiheit, Unabhängigkeit, Vollständigkeit und Kategorizität. Die Untersuchung dieser Eigenschaften für eine konkrete mathematische Theorie ist Gegenstand der Grundlagen dieser Theorie (bzw. der Metatheorie dieser Theorie). Die präzise Definition der vier genannten Begriffe und die Untersuchung ihrer allgemeinen Eigenschaften gehört demgegenüber zu den Aufgaben der Grundlagen der Mathematik (im Sinne dieses Buches und seines Titels). In den folgenden Definitionen versehen wir die drei erstgenannten Begriffe mit dem Beiwort semantisch, um sie von den später zu behandelnden analogen syntaktischen Begriffen zu unterscheiden.

Definition 5. Eine Teilmenge X einer Sprache S heißt *semantisch widerspruchsfrei*, wenn X ein Modell besitzt. (Bis zur Einführung des Begriffs syntaktische Widerspruchsfreiheit bezeichnen wir semantisch widerspruchsfreie Ausdrucksmengen kurz als widerspruchsfrei.)

Satz 3. *Es sei* S *eine formalisierte Sprache und* $X \subseteq S$. *Dann gilt*

(a) X *ist widerspruchsfrei genau dann, wenn* $Fl_S(X) \subset S$ (*d. h. wenn* $Fl_S(X) \neq S$) *ist.*

(b) X *ist widerspruchsfrei genau dann, wenn für keinen Ausdruck* $\mathrm{H} \in S$ *zugleich* $\mathrm{H} \in Fl_S(X)$ *und* $\neg\, \mathrm{H} \in Fl_S(X)$ *gilt.*

(c) X *ist widerspruchsfrei genau dann, wenn* $Fl_S(X)$ *widerspruchsfrei ist.*

(d) *Ist* X *widerspruchsfrei und* $Y \subseteq X$, *so ist* Y *widerspruchsfrei.*

(a) und (b) bedeuten die Möglichkeit zweier zu Definition 5 äquivalenter Definitionen des Begriffs der semantischen Widerspruchsfreiheit. (c) begründet die häufig terminologisch undeutliche Unterscheidung zwischen der Widerspruchsfreiheit einer Theorie und der Widerspruchsfreiheit des diese Theorie definierenden Axiomensystems. Die leichten Beweise von (a) bis (d) seien dem Leser überlassen.

Der Nachweis der Widerspruchsfreiheit eines Axiomensystems erfolgt gemäß Definition 5 primär durch die Angabe eines Modells. Die Konstruktion einer Interpretation einer formalisierten Sprache und der Nachweis, daß diese Interpretation ein Modell der vorgegebenen Axiome ist, ist jedoch in einer direkten, unanfechtbaren und effektiv durchführbaren Weise nur im Fall der Existenz endlicher Modelle möglich. Zum Beispiel kann man die Widerspruchsfreiheit der Gruppentheorie (d. h. des Axiomensystems $\{(1), (2), (3), (4)\}$ aus Beispiel 1) ohne irgendwelche Voraussetzungen aus anderen Gebieten der Mathematik dadurch beweisen, daß man eine endliche Gruppe durch ihre Wertetafel vollständig beschreibt und die Gültigkeit der Gruppenaxiome durch Nachprüfen aller Fälle verifiziert. Die meisten interessanten mathematischen Theorien besitzen jedoch kein endliches Modell. (Man denke an die Theorie der natürlichen und die der reellen Zahlen, die Mengenlehre, die euklidische und die hyperbolische Geometrie usw.) In solchen Fällen kann man ein Modell nur im Rahmen eines als existent vorausgesetzten Modells einer anderen mathematischen Theorie nachweisen.

Definition 6. Es seien X und Y Teilmengen von eventuell verschiedenen formalisierten Sprachen. Dann heißt X *relativ* (*semantisch*) *widerspruchsfrei bezüglich* Y, wenn gilt: Wenn Y widerspruchsfrei ist, so ist auch X widerspruchsfrei.

Beispiel 2. Der Nachweis, daß die in 4.1., Beispiel 1, angegebene Interpretation der Sprache der ebenen euklidischen Geometrie ein Modell des in dieser Sprache formulierten Hilbertschen Axiomensystems ist, setzt die Gültigkeit aller derjenigen Eigenschaften der reellen Zahlen voraus, die bei diesem Nachweis benutzt werden. Man zeigt also in Wirklichkeit nur: Wenn das Axiomensystem der reellen Zahlen ein Modell besitzt, so besitzt auch die ebene euklidische Geometrie ein Modell, d. h., die letztere ist relativ widerspruchsfrei bezüglich der Theorie der reellen Zahlen.

Definition 7. Es sei S eine formalisierte Sprache, $X \subseteq S$, $H \in S$. Der Ausdruck H heißt (*semantisch*) *unabhängig von* X, wenn $H \notin Fl_S(X)$ ist. Eine Menge $X \subseteq S$ heißt (*semantisch*) *unabhängig*, wenn für alle $H \in X$

$$H \notin Fl_S(X \setminus \{H\})$$

gilt.

Satz 4. *Für abgeschlossene Ausdrücke* H (*und nur für solche, auf die man sich aber bei der Formulierung von Axiomensystemen beschränken kann*) *gilt:* H *ist unabhängig von* X *genau dann, wenn* $X \cup \{\neg H\}$ *widerspruchsfrei ist.*

Beweis. Es ist $H \notin Fl_S(X)$ gleichbedeutend damit, daß wenigstens ein Modell ω von X existiert, in dem H nicht allgemeingültig ist. Ist H abgeschlossen, so ist die Nichtallgemeingültigkeit von H in ω gleichbedeutend mit der Allgemeingültigkeit von $\neg H$ in ω, d. h., ω ist ein Modell von $X \cup \{\neg H\}$. Satz 4 bedeutet, daß man die Unabhängigkeit eines Ausdrucks (Axioms) von einem Axiomensystem ebenfalls durch Angabe eines Modells zu beweisen hat.

Beispiel 3. Es sei S die Sprache der ebenen euklidischen Geometrie, H das in dieser Sprache als abgeschlossener Ausdruck formulierte Parallelenaxiom

$$\wedge \, Pg\big(\neg \, P \text{ auf } g \rightarrow \vee!!h\big(P \text{ auf } h \wedge \neg \vee Q(Q \text{ auf } g \wedge Q \text{ auf } h)\big)\big),$$

X das System der restlichen Hilbertschen Axiome. Zum Nachweis, daß H unabhängig von X ist (also insbesondere mit den Voraussetzungen X allein nicht bewiesen werden kann), hat man ein Modell für das Axiomensystem $X \cup \{\neg H\}$ anzugeben. Ein solches Modell kann hier nur skizzenhaft beschrieben werden, zumal der Nachweis der Gültigkeit aller Axiome aus X erhebliche geometrische Kenntnisse voraussetzt. Man wähle in der euklidischen Ebene einen beliebigen Kreis k, interpretiere die Punktvariablen durch Punkte im Innern von k, die Geradenvariablen durch Sehnen von k, die Inzidenz durch die auf „Punkte" und „Geraden" der vorliegenden Interpretation eingeschränkte Inzidenz der euklidischen Geometrie, ebenso die Zwischenrelation durch die auf „Punkte" eingeschränkte Zwischenrelation der euklidischen Geometrie. (Man sagt in solchem Fall auch: *Die Begriffe* liegt auf *und* zwischen *werden standardinterpretiert.*) Es ist leicht zu verifizieren und sei dem Leser, falls er dieses Modell noch nicht kennt, zur Übung empfohlen, daß alle üblichen Inzidenz- und Anordnungsaxiome erfüllt sind. „Punkte" A, B, C, D dieses Modells mögen in der durch $AB \cong CD$ ausgedrückten Kongruenzrelation stehen, wenn eine projektive Abbildung φ der betrachteten Ebene in sich existiert, die den Kreis k und damit auch sein Inneres invariant läßt und für die $\varphi(A) = C$ und $\varphi(B) = D$ gilt (d. h., die nichteuklidischen Bewegungen werden durch die projektiven Abbildungen des Kreisinneren auf sich interpretiert). Bei dieser Interpretation sind, wie sich zeigt, alle Kongruenzaxiome sowie die geeignet formulierten

Stetigkeitsaxiome erfüllt. Daß das Parallelenaxiom H dagegen falsch bzw. seine Negation wahr wird, ist wieder sofort zu sehen.[1])

Jedes nicht unabhängige Axiomensystem ist unnötig umfangreich, denn wenn für einen Ausdruck $\mathrm{H} \in \boldsymbol{X}$

$$\mathrm{H} \in Fl_S(\boldsymbol{X} \setminus \{\mathrm{H}\})$$

gilt, ist $Fl_S(\boldsymbol{X}) = Fl_S(\boldsymbol{X} \setminus \{\mathrm{H}\})$. Dies folgt wiederum allein aus den Hülleneigenschaften (a), (b), (c) (Satz 1) des Folgerungsoperators: $Fl_S(\boldsymbol{X} \setminus \{\mathrm{H}\}) \subseteq Fl_S(\boldsymbol{X})$ gilt nach (b). Die umgekehrte Inklusion ergibt sich so: $\boldsymbol{X} \setminus \{\mathrm{H}\} \subseteq Fl_S(\boldsymbol{X} \setminus \{\mathrm{H}\})$ gilt nach (a), und nach Voraussetzung ist $\mathrm{H} \in Fl_S(\boldsymbol{X} \setminus \{\mathrm{H}\})$, also ist $\boldsymbol{X} \subseteq Fl_S(\boldsymbol{X} \setminus \{\mathrm{H}\})$, folglich gilt nach (b) und (c)

$$Fl_S(\boldsymbol{X}) \subseteq Fl_S\big((Fl_S(\boldsymbol{X} \setminus \{\mathrm{H}\})\big) = Fl_S(\boldsymbol{X} \setminus \{\mathrm{H}\}).$$

Während die Unabhängigkeit einzelner Axiome oft aus inhaltlichen oder auch nur historisch erklärbaren Gründen von großem Interesse ist, hat die Frage nach der Unabhängigkeit eines gesamten Axiomensystems im Sinne von Definition 7 keine besondere Bedeutung. Insbesondere kann einem übertriebenen Streben nach einer möglichst geringen Anzahl von Axiomen durch den Hinweis begegnet werden, daß man doch jedes endliche Axiomensystem $\{\mathrm{H}_1, \ldots, \mathrm{H}_n\}$ durch das einzelne Axiom $\mathrm{H}_1 \wedge \ldots \wedge \mathrm{H}_n$ ersetzen kann.

Definition 8. Eine Teilmenge $\boldsymbol{X}$ einer formalisierten Sprache $\boldsymbol{S}$ heißt (*semantisch*) *vollständig*, wenn $\mathrm{H} \in Fl_S(\boldsymbol{X})$ oder $\neg\, \mathrm{H} \in Fl_S(\boldsymbol{X})$ für jeden abgeschlossenen Ausdruck $\mathrm{H} \in \bar{\boldsymbol{S}}$ gilt.

Aus Definition 8 folgt sofort, daß jede widerspruchsvolle Menge vollständig ist. Eine widerspruchsfreie Menge hingegen ist genau dann vollständig, wenn für jeden abgeschlossenen Ausdruck $\mathrm{H} \in \bar{\boldsymbol{S}}$ genau eine der beiden Beziehungen $\mathrm{H} \in Fl_S(\boldsymbol{X})$, $\neg\, \mathrm{H} \in Fl_S(\boldsymbol{X})$ gilt. Wir zeigen

Satz 5. *Eine Teilmenge $\boldsymbol{X}$ einer formalisierten Sprache $\boldsymbol{S}$ ist genau dann widerspruchsfrei und vollständig, wenn $Fl_S(\boldsymbol{X})$ eine maximale widerspruchsfreie Teilmenge von $\boldsymbol{S}$ ist, d. h., wenn $Fl_S(\boldsymbol{X})$ widerspruchsfrei und jede Menge $\boldsymbol{Y}$ mit $Fl_S(\boldsymbol{X}) \subset \boldsymbol{Y} \subseteq \boldsymbol{S}$ widerspruchsvoll ist.*

Beweis. Wir setzen zunächst voraus, daß $\boldsymbol{X}$ widerspruchsfrei und vollständig ist. Dann ist nach Satz 3 (c) auch $Fl_S(\boldsymbol{X})$ widerspruchsfrei. Wir nehmen an, $\boldsymbol{Y}$ sei eine echte Obermenge von $Fl_S(\boldsymbol{X})$, die noch widerspruchsfrei ist. Es sei $\mathrm{H} \in \boldsymbol{Y} \setminus Fl_S(\boldsymbol{X})$. Falls H abgeschlossen ist, sei $\mathrm{H}' = \mathrm{H}$. Andernfalls sei H' ein Ausdruck, der aus H entsteht, indem man zunächst durch gebundene Umbenennungen alle frei vor-

[1]) Ausführliche Behandlungen des hier skizzierten Kleinschen Modells findet man u. a. in Enzyklopädie der Elementarmathematik, Band 5, DVW, Berlin 1971, sowie in A. P. Norden, Elementare Einführung in die Lobatschewskische Geometrie, DVW, Berlin 1958 (Übersetzungen aus dem Russischen).

kommenden Variablen vollfrei macht und dann zur Generalisierten übergeht. In beiden Fällen gilt für beliebige $\boldsymbol{Z} \subseteqq \boldsymbol{S}$, daß $\mathrm{H} \in Fl_{\boldsymbol{S}}(\boldsymbol{Z})$ genau dann, wenn $\mathrm{H}' \in Fl_{\boldsymbol{S}}(\boldsymbol{Z})$. Insbesondere ist daher

$$\mathrm{H}' \in Fl_{\boldsymbol{S}}(\boldsymbol{Y}), \tag{5}$$

$$\mathrm{H}' \notin Fl_{\boldsymbol{S}}(\boldsymbol{X}). \tag{6}$$

Da H′ abgeschlossen und $\boldsymbol{X}$ vollständig ist, ist (6) gleichbedeutend mit

$$\neg\, \mathrm{H}' \in Fl_{\boldsymbol{S}}(\boldsymbol{X}) \subset \boldsymbol{Y} \subseteqq Fl_{\boldsymbol{S}}(\boldsymbol{Y}).$$

Dies bedeutet zusammen mit (5), daß $Fl_{\boldsymbol{S}}(\boldsymbol{Y})$ und daher auch $\boldsymbol{Y}$ widerspruchsvoll ist, im Widerspruch zu unserer Annahme. Es sei nun $Fl_{\boldsymbol{S}}(\boldsymbol{X})$ eine maximale widerspruchsfreie Menge. Dann ist mit $Fl_{\boldsymbol{S}}(\boldsymbol{X})$ auch $\boldsymbol{X}$ widerspruchsfrei. Wir nehmen an, $\boldsymbol{X}$ wäre nicht vollständig, d. h., es gibt einen abgeschlossenen Ausdruck H, so daß weder $\mathrm{H} \in Fl_{\boldsymbol{S}}(\boldsymbol{X})$ noch $\neg\, \mathrm{H} \in Fl_{\boldsymbol{S}}(\boldsymbol{X})$ gilt. Das erste bedeutet, daß es ein Modell von $\boldsymbol{X}$ gibt, in dem H nicht allgemeingültig, d. h. wegen der Abgeschlossenheit $\neg$ H allgemeingültig ist. Demnach ist $Fl_{\boldsymbol{S}}(\boldsymbol{X}) \cup \{\neg\, \mathrm{H}\}$ widerspruchsfrei und wegen $\neg\, \mathrm{H} \notin Fl_{\boldsymbol{S}}(\boldsymbol{X})$ eine echte Obermenge von $Fl_{\boldsymbol{S}}(\boldsymbol{X})$.

Ist $\boldsymbol{X}$ ein in einer Sprache $\boldsymbol{S}$ formuliertes widerspruchsfreies und vollständiges Axiomensystem, so gilt nach Satz 5 für jeden abgeschlossenen Ausdruck $\mathrm{H} \in \bar{\boldsymbol{S}}$ entweder $\mathrm{H} \in Fl_{\boldsymbol{S}}(\boldsymbol{X})$ oder $\neg\, \mathrm{H} \in Fl_{\boldsymbol{S}}(\boldsymbol{X})$. Im ersten Fall gilt H in allen Modellen von $\boldsymbol{X}$, im zweiten Fall gilt H in keinem Modell von $\boldsymbol{X}$. Demnach gibt es keine in der Sprache $\boldsymbol{S}$ formulierbare Aussage, durch die sich zwei beliebige Modelle von $\boldsymbol{X}$ unterscheiden lassen. Wir sagen: Je zwei Modelle eines widerspruchsfreien und vollständigen Axiomensystems sind *elementar ununterscheidbar*.

4.4. Isomorphie von Interpretationen, Kategorizität

Definition 1. Es seien ω_1 und ω_2 Interpretationen einer Sprache $\boldsymbol{S}$, die durch die Basis $\boldsymbol{B} = (n, \boldsymbol{B}_r, \boldsymbol{B}_f, \boldsymbol{B}_c)$ gegeben ist. Ein *Isomorphismus von ω_1 auf ω_2* ist eine Abbildung φ, die für $j = 1, \ldots, n$ den Grundbereich $\omega_1(j)$ eineindeutig auf den Grundbereich $\omega_2(j)$ abbildet, so daß gilt:

(a) *Für* $\mathrm{R}_i^{j_1,\ldots,j_k} \in \boldsymbol{B}_r$ *gilt für beliebige* $\xi_\varkappa \in \omega_1(j_\varkappa)$ $\quad (\varkappa = 1, \ldots, k)$
$(\xi_1, \ldots, \xi_k) \in \omega_1(\mathrm{R}_i^{j_1,\ldots,j_k})$ *genau dann, wenn* $\big(\varphi(\xi_1), \ldots, \varphi(\xi_k)\big) \in \omega_2(\mathrm{R}_i^{j_1,\ldots,j_k})$.

(b) *Für* $\mathrm{F}_i^{j_1,\ldots,j_k;j} \in \boldsymbol{B}_f$ *gilt für beliebige* $\xi_\varkappa \in \omega_1(j_\varkappa)$ $\quad (\varkappa = 1, \ldots, k)$:
$\omega_1(\mathrm{F}_i^{j_1,\ldots,j_k;j})\,(\xi_1, \ldots, \xi_k)$ *existiert genau dann, wenn*
$\omega_2(\mathrm{F}_i^{j_1,\ldots,j_k;j})\,\big(\varphi(\xi_1), \ldots, \varphi(\xi_k)\big)$ *existiert und im Fall der Existenz gilt*
$\varphi\big(\omega_1(\mathrm{F}_i^{j_1,\ldots,j_k;j})\,(\xi_1, \ldots, \xi_k)\big) = \omega_2(\mathrm{F}_i^{j_1,\ldots,j_k;j})\,\big(\varphi(\xi_1), \ldots, \varphi(\xi_k)\big)$.

(c) *Für* $\mathrm{c}_i^j \in \boldsymbol{B}_c$ *gilt* $\varphi\big(\omega_1(\mathrm{c}_i^j)\big) = \omega_2(\mathrm{c}_i^j)$.

Der Leser prüfe nach, daß die Bedingungen (b) und (c) sich zwangsläufig aus (a) ergeben, wenn man Konstanten und (partielle) Operationen als spezielle Relationen auffaßt.

Beispiel 1. Es sei S die in 4.3., Beispiel 1, betrachtete Sprache (zur Formulierung „gruppentheoretischer" Aussagen), ω_1 und ω_2 seien Interpretationen dieser Sprache. Wir bezeichnen den Grundbereich von ω_i mit M_i ($i = 1, 2$) und die in M_i definierte Operation $\omega_i(+)$ mit „$+_i$". Dann ergibt sich aus Definition 1 (in diesem Fall wird nur (b) benötigt): Ein Isomorphismus der Interpretation ω_1 auf die Interpretation ω_2 (und damit von der Struktur $(M_1, +_1)$ auf die Struktur $(M_2, +_2)$) ist eine eineindeutige Abbildung φ von M_1 auf M_2, so daß für beliebige $\xi, \eta \in M_1$ genau dann $\xi +_1 \eta$ existiert, wenn $\varphi(\xi) +_2 \varphi(\eta)$ existiert, und im Fall der Existenz gilt

$$\varphi(\xi +_1 \eta) = \varphi(\xi) +_2 \varphi(\eta). \tag{1}$$

Der Leser wird (1) als die aus der Gruppentheorie bekannte Isomorphiebedingung erkennen. (Wenn man schon voraussetzt, daß $(M_1, +_1)$ und $(M_2, +_2)$ Gruppen sind, die entsprechenden Operationen also für alle Elementpaare ausführbar sind, entfällt natürlich die im Fall partieller Operationen zusätzlich gestellte Bedingung, daß die linke Seite von (1) genau dann definiert sein soll, wenn die rechte Seite definiert ist.)

Beispiel 2. Mit S_{inz} bezeichnen wir die durch die Basis $(2, \{R^{1,2}\}, \emptyset, \emptyset)$ gegebene Teilsprache der kanonischen Sprache der ebenen Geometrie. In dieser Sprache der ebenen Inzidenzgeometrie kann man natürlich nicht nur Aussagen der euklidischen, sondern auch der hyperbolischen, projektiven oder sonstigen ebenen Inzidenzgeometrie formulieren. Eine beliebige Interpretation ω der Sprache S_{inz} wird gegeben durch eine Menge $\mathcal{P}$ von „Punkten", eine Menge $\mathcal{G}$ von „Geraden" und eine Relation $Inz \subseteq \mathcal{P} \times \mathcal{G}$. Eine Struktur der Form $(\mathcal{P}, \mathcal{G}, Inz)$ nennen wir eine *(ebene) Inzidenzstruktur*. Dabei brauchen die Elemente von $\mathcal{P}$ durchaus nicht Punkte in irgendeinem geometrischen Sinn und die Elemente von $\mathcal{G}$ keine Geraden zu sein, sondern es kann sich, wie Hilbert einmal formulierte, um Tische, Bänke oder Bierseidel handeln. Eine Inzidenzstruktur $(\mathcal{P}, \mathcal{G}, Inz)$ wollen wir als *speziell* bezeichnen, wenn $\mathcal{G}$ eine Teilmenge der Potenzmenge von $\mathcal{P}$ ist (die „Geraden" also Mengen von „Punkten" sind) und Inz die auf $\mathcal{P} \times \mathcal{G}$ eingeschränkte Elementrelation ist. Nach Definition 1 sind zwei Inzidenzstrukturen $\omega_i = (\mathcal{P}_i, \mathcal{G}_i, Inz_i)$ ($i = 1, 2$) isomorph, wenn eine eineindeutige Abbildung φ von $\mathcal{P}_1$ auf $\mathcal{P}_2$ und von $\mathcal{G}_1$ auf $\mathcal{G}_2$ existiert, so daß für beliebige $\pi \in \mathcal{P}_1$ und $\gamma \in \mathcal{G}_1$ genau dann $(\pi, \gamma) \in Inz_1$ gilt, wenn $\big(\varphi(\pi), \varphi(\gamma)\big) \in Inz_2$ gilt. Wir beweisen folgenden für das anschauliche Arbeiten mit Inzidenzstrukturen wesentlichen

Satz 1. *Ist eine Inzidenzstruktur ein Modell der beiden Aussagen*

$$\wedge\, x_1{}^1 x_2{}^1\big(\neg\, x_1{}^1 = x_2{}^1 \rightarrow \vee!!x_1{}^2\big(R^{1,2}(x_1{}^1, x_1{}^2) \wedge R^{1,2}(x_2{}^1, x_1{}^2)\big)\big)$$

(d. h., zu je zwei verschiedenen Punkten gibt es genau eine mit beiden inzidierende Gerade),

$$\bigwedge x_1^{2} \bigvee_{2} x_1^{1} R^{1,2}(x_1^{1}, x_1^{2})$$

(d. h., auf jeder Geraden gibt es wenigstens zwei Punkte), *so existiert eine zu ihr isomorphe spezielle Inzidenzstruktur.*

Beweis. Es sei $\omega_1 = (\mathscr{P}_1, \mathscr{G}_1, Inz)$ ein beliebiges Modell der beiden im Satz genannten Aussagen. Wir definieren eine zu ω_1 isomorphe Interpretation ω_2 und zugleich einen Isomorphismus φ von ω_1 auf ω_2 durch

$$\mathscr{P}_2 = \mathscr{P}_1, \quad \varphi(\pi) = \pi \text{ für } \pi \in \mathscr{P}_1,$$

$$\varphi(\gamma) = \{\pi : (\pi, \gamma) \in Inz\}, \quad \mathscr{G}_2 = \{\varphi(\gamma) : \gamma \in \mathscr{G}_1\}.$$

Offenbar ist $(\mathscr{P}_2, \mathscr{G}_2, \in)$ eine spezielle Inzidenzstruktur und φ eine Abbildung von $\mathscr{P}_1$ auf $\mathscr{P}_2$ und von $\mathscr{G}_1$ auf $\mathscr{G}_2$. Aus der Gültigkeit der im Satz genannten Aussagen in ω_1 ergibt sich, daß jede „Gerade" dieser Struktur durch die Gesamtheit der mit ihr inzidierenden „Punkte" eindeutig bestimmt ist, d. h., φ ist nicht nur eineindeutig von $\mathscr{P}_1$ auf $\mathscr{P}_2$, sondern auch von $\mathscr{G}_1$ auf $\mathscr{G}_2$. Die Isomorphiebedingung für φ folgt nun sofort aus der Definition von $\varphi(\gamma)$.

Satz 1 und sein Beweis bedeuten: Sind die „Geraden" einer Inzidenzstruktur zunächst nicht Mengen von „Punkten" (sondern z. B. „geometrische Örter" im Sinne der antiken Geometrie), so kann man unter den in Satz 1 genannten (in fast allen in Frage kommenden geometrischen Theorien erfüllten) Voraussetzungen zu einer isomorphen Struktur mit den gleichen „Punkten" übergehen, in der jede Gerade gleich der Menge der mit ihr inzidierenden Punkte ist. In welchem Sinn die Betrachtung der isomorphen Struktur der Betrachtung der ursprünglichen Struktur gleichwertig ist, wird durch den folgenden Satz präzisiert.

Satz 2. *Sind ω_1 und ω_2 Interpretationen einer beliebigen Sprache* S, *ist φ ein Isomorphismus von ω_1 auf ω_2 und f eine Belegung bezüglich ω_1, so sei f_φ die durch* $f_\varphi(x) := \varphi(f(x))$ *definierte Belegung bezüglich ω_2. Es gilt*

(a) *für alle Terme* t *der Sprache* S: $Wert(t, \omega_1, f)$ *existiert genau dann, wenn auch* $Wert(t, \omega_2, f_\varphi)$ *existiert, und im Fall der Existenz ist* $\varphi(Wert(t, \omega_1, f)) = Wert(t, \omega_2, f_\varphi)$;

(b) *für alle Ausdrücke* $H \in S$: $Wert(H, \omega_1, f) = Wert(H, \omega_2, f_\varphi)$.

Wir beweisen zunächst (a) durch Induktion über die Kompliziertheit des Terms t. Ist t eine Variable, so gilt (a) nach Definition von f_φ. Ist t ein Konstantensymbol, so gilt (a) nach Definition 1(c). Es sei $t = F_i^{j_1,\ldots,j_k;j}(t_1, \ldots, t_k)$, und (a) sei schon für die Terme $t_1, \ldots, t_k$ bei beliebiger Belegung f bewiesen. Existiert $Wert(t, \omega_1, f)$, so existieren

$$\xi_\varkappa = Wert(t_\varkappa, \omega_1, f) \quad \text{für } \varkappa = 1, \ldots, k,$$

und die Operation $\omega_1(\mathrm{F}_i^{j_1,\ldots,j_k;j})$ ist auf das k-Tupel $\xi_1, \ldots, \xi_k$ anwendbar. Nach Induktionsannahme existieren dann auch die Elemente $\varphi(\xi_\varkappa)$ und sind für $\varkappa = 1, \ldots, k$ gleich $Wert(\mathrm{t}_\varkappa, \omega_2, f_\varphi)$. Da φ ein Isomorphismus von ω_1 auf ω_2 ist, folgt nach Definition 1 (b) die Gleichung

$$\begin{aligned}\varphi\big(Wert(\mathrm{t}, \omega_1, f)\big) &= \varphi\big(\omega_1(F_i^{j_1,\ldots,j_k;j})\,(\xi_1, \ldots, \xi_k)\big)\\ &= \omega_2(\mathrm{F}_i^{j_1,\ldots,j_k;j})\,\big(\varphi(\xi_1), \ldots, \varphi(\xi_k)\big) = Wert(\mathrm{t}, \omega_2, f_\varphi).\end{aligned}$$

Existiert $Wert(\mathrm{t}, \omega_1, f)$ nicht, so sind folgende zwei Fälle möglich:

a) $Wert(\mathrm{t}_\varkappa, \omega_1, f)$ existiert für wenigstens ein $\varkappa \in \{1, \ldots, k\}$ nicht. In diesem Fall existiert nach Induktionsannahme auch $Wert(\mathrm{t}_\varkappa, \omega_2, f_\varphi)$ nicht und daher erst recht nicht $Wert(\mathrm{t}, \omega_2, f_\varphi)$.

b) $Wert(\mathrm{t}_\varkappa, \omega_1, f)$ existiert zwar für $\varkappa = 1, \ldots, k$, aber die Operation $\omega_1(\mathrm{F}_i^{j_1,\ldots,j_k;j})$ ist auf das k-Tupel dieser Werte nicht anwendbar. In diesem Fall existieren nach Induktionsannahme zwar die Werte $Wert(\mathrm{t}_\varkappa, \omega_2, f_\varphi)$, aber da φ ein Isomorphismus ist, ist die Operation $\omega_2(\mathrm{F}_i^{j_1,\ldots,j_k;j})$ auf das k-Tupel dieser Werte nicht anwendbar, d. h., $Wert(\mathrm{t}, \omega_2. f_\varphi)$ existiert nicht.

Wir beweisen nun (b) durch Induktion über die Kompliziertheit des Ausdrucks H. Ist H eine Termgleichung $\mathrm{t}_1 = \mathrm{t}_2$, so ist $Wert(\mathrm{H}, \omega_1, f) = \boldsymbol{W}$ genau dann, wenn $Wert(\mathrm{t}_i, \omega_1, f)$ für $i = 1, 2$ existiert und diese Werte gleich sind. Auf Grund des bereits bewiesenen Teils (a) von Satz 2 ist dies aber genau dann der Fall, wenn $Wert(\mathrm{t}_i, \omega_2, f_\varphi)$ für $i = 1, 2$ existiert und diese Werte (wegen der Eineindeutigkeit von φ) ebenfalls übereinstimmen, d. h., wenn $Wert(\mathrm{H}, \omega_2, f_\varphi) = \boldsymbol{W}$ ist. Ist H ein prädikativer Ausdruck der Form $\mathrm{R}(\mathrm{t}_1, \ldots, \mathrm{t}_k)$, wobei R ein k-stelliges Relationssymbol der Sprache $\boldsymbol{S}$ bezeichnet, so ist $Wert(\mathrm{H}, \omega_1, f) = \boldsymbol{W}$ genau dann, wenn

$$Wert(\mathrm{t}_\varkappa, \omega_1, f) = \xi_\varkappa \quad \text{für } \varkappa = 1, \ldots, k$$

existiert und $(\xi_1, \ldots, \xi_k) \in \omega_1(\mathrm{R})$ gilt. Da φ ein Isomorphismus von ω_1 auf ω_2 ist, gilt dies genau dann, wenn $\big(\varphi(\xi_1), \ldots, \varphi(\xi_k)\big) \in \omega_2(\mathrm{R})$ ist. Auf Grund des bereits bewiesenen Teils (a) ist aber

$$\varphi(\xi_\varkappa) = Wert(\mathrm{t}_\varkappa, \omega_2, f_\varphi) \quad \text{für } \varkappa = 1, \ldots, k,$$

d. h., es gilt insgesamt $Wert(\mathrm{H}, \omega_1, f) = \boldsymbol{W}$ genau dann, wenn $Wert(\mathrm{H}, \omega_2, f_\varphi) = \boldsymbol{W}$. Ist schon $Wert(\mathrm{H}, \omega_1, f) = Wert(\mathrm{H}, \omega_2, f_\varphi)$ für einen beliebigen Ausdruck H, so gilt trivialerweise

$$Wert(\neg\, \mathrm{H}, \omega_1, f) = Wert(\neg\, \mathrm{H}, \omega_2, f_\varphi).$$

Analog schließt man für die übrigen aussagenlogischen Verknüpfungen.

Wir setzen nun voraus, daß die Behauptung für den Ausdruck H(x) bei beliebigen Belegungen f bezüglich ω_1 bewiesen ist. Es gilt

$$Wert(\wedge\, \mathrm{x}\mathrm{H}(\mathrm{x}), \omega_1, f) = W \tag{2}$$

genau dann, wenn für alle Elemente ξ des bei der Interpretation ω_1 zur Variablensorte von x gehörigen Grundbereichs M_1

$$Wert\left(\mathrm{H}(\mathrm{x}),\ \omega_1, f\left\langle \begin{matrix} \mathrm{x} \\ \xi \end{matrix} \right\rangle\right) = W \tag{3}$$

gilt. Nach Induktionsannahme ist (3) gleichbedeutend mit

$$Wert\left(\mathrm{H}(\mathrm{x}),\ \omega_2, \left(f\left\langle \begin{matrix} \mathrm{x} \\ \xi \end{matrix} \right\rangle\right)_\varphi\right) = W. \tag{4}$$

Offenbar ist $\left(f\left\langle \begin{matrix} \mathrm{x} \\ \xi \end{matrix} \right\rangle\right)_\varphi = f_\varphi\left\langle \begin{matrix} \mathrm{x} \\ \varphi(\xi) \end{matrix} \right\rangle$. Da φ eine eineindeutige Abbildung von M_1 auf den entsprechenden Grundbereich M_2 der Interpretation ω_2 ist, ist die Gültigkeit von (4) für alle $\xi \in M_1$ gleichbedeutend mit der Gültigkeit von

$$Wert\left(\mathrm{H}(\mathrm{x}), \omega_2, f_\varphi\left\langle \begin{matrix} \mathrm{x} \\ \eta \end{matrix} \right\rangle\right) = W \quad \text{für alle } \eta \in M_2, \text{ d. h. mit}$$

$$Wert(\wedge\, \mathrm{x}\mathrm{H}(\mathrm{x}), \omega_2, f_\varphi) = W. \tag{5}$$

Folglich gilt (2) genau dann, wenn (5) gilt. Der hierzu weitgehend analoge letzte Beweisschritt für Ausdrücke der Form $\vee\, \mathrm{x}\mathrm{H}(\mathrm{x})$ sei dem Leser überlassen.

Aus Satz 2(b) ergibt sich sofort

Folgerung. *Sind ω_1 und ω_2 isomorphe Interpretationen einer beliebigen Sprache* S, *so ist ein beliebiger Ausdruck* H *bei ω_1 allgemeingültig genau dann, wenn er bei ω_2 allgemeingültig ist, d. h., isomorphe Interpretationen sind erst recht elementar ununterscheidbar.*

Beispiel 3. Als Anwendungen der obigen Folgerung ergeben sich Sätze folgender Art: Ist $(M_1, +_1)$ eine Gruppe und $(M_2, +_2)$ eine Struktur, von der nur vorausgesetzt wird, daß die mit $+_2$ bezeichnete (eventuell partielle) Operation in der Menge M_2 zweistellig ist, so folgt aus der Existenz eines Isomorphismus von $(M_1, +_1)$ auf $(M_2, +_2)$, daß auch $(M_2, +_2)$ eine Gruppe ist. Unter den in $(M_1, +_1)$ und folglich auch in $(M_2, +_2)$ allgemeingültigen Ausdrücken befinden sich nämlich die Gruppenaxiome (vgl. 4.3., Beispiel 1). Analog erhält man: Jede zu einem Ring isomorphe Struktur ist ein Ring; jede zu einem Körper isomorphe Struktur ist ein Körper; eine zu einer kommutativen Gruppe isomorphe Gruppe ist ebenfalls kommutativ; ein zu einem Körper der Charakteristik p isomorpher Körper hat ebenfalls die Charakteristik p usw.

Definition 2. Ein in einer formalisierten Sprache S formuliertes widerspruchsfreies Axiomensystem $X \subset S$ (und die durch X definierte elementare Theorie $Fl_S(X)$) heißt *kategorisch*, wenn je zwei Modelle von X zueinander isomorph sind. (In diesem Fall sagen wir auch: *X besitzt bis auf Isomorphie nur ein einziges Modell.*)

Definition 2 gestattet es, über zwei wesentlich verschiedene Ziele der axiomatischen Methode in verschiedenen Teilgebieten der Mathematik zu sprechen. Manche Axiomensysteme (wie z. B. die in der Gruppen-, Ring-, Körper- und Verbandstheorie benutzten) sind absichtlich so „schwach" gewählt, daß sie sehr viele wesentlich verschiedene (d. h. paarweise nicht zueinander isomorphe) Modelle besitzen. Hier sollen gewisse gemeinsame Gesetzmäßigkeiten, die in verschiedensten Strukturen auftreten, unter einem gemeinsamen Gesichtspunkt untersucht werden bzw. gewisse Sätze, deren Beweise andernfalls in verschiedenen Teilgebieten der Mathematik jeweils wörtlich zu wiederholen wären, werden als z. B. gruppentheoretische Sätze erkannt und nur einmal, nämlich in der Gruppentheorie, bewiesen. Demgegenüber verfolgen gewisse andere axiomatische Theorien das Ziel, im wesentlichen (d. h. bis auf Isomorphie) eine einzige Struktur durch innere Eigenschaften zu charakterisieren. Hierher gehören z. B. die axiomatischen Theorien der verschiedenen Zahlbereiche und die ebene und räumliche euklidische Geometrie. In diesen Theorien wird also ein kategorisches Axiomensystem gesucht, dessen (bis auf Isomorphie) einziges Modell die zu charakterisierende Struktur ist. Allerdings zeigt sich, daß, abgesehen von Trivialfällen, eine solche axiomatische Charakterisierung mit den bisher behandelten Mitteln nicht möglich ist. Es gilt der

Satz 3 (Satz von Tarski). *Besitzt ein Axiomensystem ein unendliches Modell (d. h. ein Modell mit wenigstens einem unendlichen Grundbereich), so ist dieses Axiomensystem nicht kategorisch.*

Auf den Beweis dieses Satzes müssen wir hier verzichten. In 6.4. wird der Weg, auf dem dieser Satz erhalten werden kann, skizziert werden. Der folgende Abschnitt ist einer solchen Modifizierung des bisher behandelten Sprach- und Folgerungsbegriffs gewidmet, mit deren Hilfe kategorische Axiomensysteme auch für unendliche Strukturen erhalten werden können.

4.5. Nichtelementare Sprachen und Theorien

Definition 1. Ist S eine beliebige formalisierte Sprache und σ eine beliebige nichtleere Klasse von Interpretationen von S, so heißt das Paar (S, σ) eine *Sprache*. Mit σ_0 bezeichnen wir stets die Klasse aller Interpretationen von S und nennen die

Sprachen (S, σ_0) *elementar*. Ist σ eine echte Teilklasse von σ_0, so heißt die Sprache (S, σ) *nichtelementar*. Die Elemente von σ werden als *zulässige* Interpretationen von (S, σ) bezeichnet.

Der Begriff der elementaren Sprache stimmt im wesentlichen mit dem bisher betrachteten Sprachbegriff überein. Für konkrete nichtelementare Sprachen wird die Klasse σ der zulässigen Interpretationen meist durch eine verbale Beschreibung dieser Interpretationen gegeben. (Zum Beispiel bezeichne $\sigma_{(1)}$ stets die Klasse aller Interpretationen von S, bei denen alle Operationssymbole durch volle Operationen interpretiert werden.) Diese Beschreibung bezeichnen wir allgemein als eine *Interpretationsvorschrift* für S (und identifizieren sie zuweilen mit der Klasse σ der beschriebenen Interpretationen, wenn keine Mißverständnisse möglich sind). Der hier definierte Begriff der nichtelementaren Sprache ist etwas allgemeiner und abstrakter als die in der Literatur unter diesem Namen behandelten speziellen nichtelementaren Sprachen (bei denen die Interpretationsvorschrift mehr oder weniger stillschweigend verabredet und häufig durch ganz spezielle Modifizierungen der benutzten formalisierten Sprache zum Ausdruck gebracht wird). Er umfaßt jedoch, wie die folgenden Beispiele zeigen werden, alle diese speziellen nichtelementaren Sprachen, verallgemeinert sie in naheliegender Weise und gestattet es zugleich, viele für elementare Sprachen formulierte Begriffe und Sachverhalte fast wörtlich auf beliebige nichtelementare Sprachen zu übertragen.

Beispiel 1. Die zweisortige Sprache S enthalte die Variablen $\mathbf{x}_i^2$ der zweiten Sorte nur in prädikativen Ausdrücken der Form $\mathbf{x}_j^1 \in \mathbf{x}_i^2$, so daß also $\in$ ein Relationssymbol vom Typ $(1, 2)$ ist. Es sei $\sigma_{(2)}$ die Klasse aller Interpretationen von S, bei denen der Grundbereich $\omega(2)$ die Potenzmenge des Grundbereichs $\omega(1)$ ist und das Symbol $\in$ durch die Elementrelation interpretiert wird. Nichtelementare Sprachen der Form $(S, \sigma_{(2)})$ bezeichnet man als *Sprachen zweiter Stufe*. Analog zu den Sprachen zweiter Stufe kann man für $n > 2$ *nichtelementare Sprachen n-ter Stufe* bilden. Dazu wähle man n Variablensorten x_i^j $(1 \leqq j \leqq n)$ und prädikative Ausdrücke der Formen $\mathbf{x}_i^j \in \mathbf{x}_k^{j+1}$ $(1 \leqq j \leqq n-1$; als Prädikatensymbol kann man stets das gleiche Zeichen $\in$ nehmen, da die jeweilige Bedeutung aus den Sortenindizes der dabeistehenden Variablen hervorgeht). Es sei $\sigma_{(n)}$ die Klasse aller Interpretationen dieser Sprache S, bei denen $\omega(j+1)$ durch die Potenzmenge von $\omega(j)$ $(1 \leqq j \leqq n-1)$ und $\in$ durch die Elementrelation interpretiert wird. Eine Interpretation $\omega \in \sigma_{(n)}$ von S ist daher jeweils durch die Wahl einer beliebigen nichtleeren Menge $\omega(1)$ bereits eindeutig bestimmt.

Beispiel 2. Enthält eine Sprache S Variablen der Sorten $\mathbf{x}_i$ und $\mathbf{f}_i$ und ist bei beliebigen i, k die Zeichenreihe $\mathbf{f}_i(\mathbf{x}_k)$ ein Term der Sorte $\mathbf{x}$ (d. h., ...(...) ist ein modifiziertes Operationssymbol vom Typ $(1, 2; 2)$), so ist damit (meist stillschweigend) folgende Interpretationsvorschrift verbunden: Ist bei einer Interpretation der zu den Variablen $\mathbf{x}_i$ gehörige Grundbereich M, so ist der zu den Variablen $\mathbf{f}_i$ gehörige Grundbereich eine Menge von eindeutigen Abbildungen von M in sich, und dem

Operationssymbol ... (...) wird die durch $\Phi(\varphi, \xi) := \varphi(\xi)$ definierte sogenannte *Anwendungsoperation* zugeordnet.

Definition 2. Ist $(\boldsymbol{S}, \sigma)$ eine nichtelementare Sprache, so verstehen wir unter einem *σ-Modell* eines Axiomensystems $\boldsymbol{X} \subseteqq \boldsymbol{S}$ eine Interpretation $\omega \in \sigma$, die im früher definierten Sinn ein Modell von $\boldsymbol{X}$ ist. Ferner sei

$$Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X}) := \{\mathrm{H} : \mathrm{H} \in \boldsymbol{S} \textit{ und } \mathrm{H} \textit{ allgemeingültig in allen } \sigma\textit{-Modellen von } \boldsymbol{X}\}$$

die *σ-Folgerungshülle* von $\boldsymbol{X}$.

Wie man leicht zeigt (vgl. 4.3., Satz 1), ist $Fl_{\boldsymbol{S}}^{\sigma}$ bei beliebigem $\boldsymbol{S}$ und σ ein Hüllenoperator in der Potenzmenge von $\boldsymbol{S}$. Als Spezialfall des durch die Operation $Fl_{\boldsymbol{S}}^{\sigma}$ gegebenen nichtelementaren σ-Folgerns ergibt sich für $\sigma = \sigma_0$ das in 4.3. betrachtete Folgern, das wir von nun an als *elementar* bezeichnen werden. Sind σ, σ' Klassen von Interpretationen einer Sprache $\boldsymbol{S}$ mit $\sigma \subseteqq \sigma'$, so gilt für beliebige $\boldsymbol{X} \subseteqq \boldsymbol{S}$ stets

$$Fl_{\boldsymbol{S}}^{\sigma'}(\boldsymbol{X}) \subseteqq Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X}),$$

d. h., die Menge der aus einem Axiomensystem erhältlichen Folgerungen wächst mit der Verkleinerung der Klasse der zulässigen Interpretationen. Insbesondere ist

$$Fl_{\boldsymbol{S}}(\boldsymbol{X}) \subseteqq Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X})$$

für beliebige nichtelementare Sprachen $(\boldsymbol{S}, \sigma)$ und $\boldsymbol{X} \subseteqq \boldsymbol{S}$. Als einen zum elementaren Folgern entgegengesetzten Ausnahmefall betrachten wir eine Klasse $\sigma = \{\omega_0\}$ von Interpretationen einer Sprache $\boldsymbol{S}$, die nur ein einziges Element ω_0 enthält. Ist für ein Axiomensystem $\boldsymbol{X} \subseteqq \boldsymbol{S}$ die Interpretation ω_0 kein Modell von $\boldsymbol{X}$, so ist jeder Ausdruck $\mathrm{H} \in \boldsymbol{S}$ in allen σ-Modellen von $\boldsymbol{X}$ allgemeingültig, d. h., es ist $Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X}) = \boldsymbol{S}$. Andernfalls besteht $Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X})$ gerade aus den in ω_0 allgemeingültigen Ausdrücken der Sprache $\boldsymbol{S}$. Demnach besteht das σ-Folgern in diesem Fall in der Überprüfung, ob der betrachtete Ausdruck bei der einzig zugelassenen Interpretation ω_0 von $\boldsymbol{S}$ allgemeingültig ist.

Die in 4.3. und 4.4. betrachteten Begriffe lassen sich ohne weiteres auf nichtelementare Sprachen übertragen, indem man in den jeweiligen Definitionen das elementare Folgern $Fl_{\boldsymbol{S}}$ durch das σ-Folgern $Fl_{\boldsymbol{S}}^{\sigma}$ ersetzt:

Definition 3. Eine *nichtelementare Theorie* $\boldsymbol{T}$ ist eine Menge der Form $\boldsymbol{T} = Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X})$, wobei $(\boldsymbol{S}, \sigma)$ eine nichtelementare Sprache und $\boldsymbol{X} \subseteqq \boldsymbol{S}$ ein Axiomensystem ist.

Definition 4. Ist $(\boldsymbol{S}, \sigma)$ eine nichtelementare Sprache, so heißt ein Axiomensystem $\boldsymbol{X} \subseteqq \boldsymbol{S}$ (und die durch $\boldsymbol{X}$ definierte nichtelementare Theorie $Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X})$)

σ-widerspruchsfrei, wenn $\boldsymbol{X}$ ein σ-Modell besitzt (genau dann, wenn $Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X}) \subset \boldsymbol{S}$ bzw. genau dann, wenn für kein $\mathrm{H} \in \boldsymbol{S}$ sowohl $\mathrm{H} \in Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X})$ als auch $\neg\, \mathrm{H} \in Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X})$ gilt, Beweise analog wie in 4.3.);

σ-unabhängig, wenn für kein $\mathrm{H} \in \boldsymbol{X}$ gilt $\mathrm{H} \in \boldsymbol{Fl_S}^{\sigma}(\boldsymbol{X} \setminus \{\mathrm{H}\})$;

σ-vollständig, wenn für $\mathrm{H} \in \overline{\boldsymbol{S}}$ gilt $\mathrm{H} \in \boldsymbol{Fl_S}^{\sigma}(\boldsymbol{X})$ oder $\neg\, \mathrm{H} \in \boldsymbol{Fl_S}^{\sigma}(\boldsymbol{X})$;

σ-kategorisch, wenn $\boldsymbol{X}$ σ-widerspruchsfrei ist und je zwei σ-Modelle isomorph sind (d. h., wenn $\boldsymbol{X}$ bis auf Isomorphie genau ein σ-Modell besitzt).

Aus der bereits erwähnten Antimonotonie

$$\boldsymbol{Fl_S}^{\sigma'}(\boldsymbol{X}) \subseteqq \boldsymbol{Fl_S}^{\sigma}(\boldsymbol{X}) \qquad \textit{für } \sigma \subseteqq \sigma' \tag{1}$$

ergibt sich sofort

Satz 1. *Sind σ, σ' Interpretationsklassen von* $\boldsymbol{S}$ *mit* $\sigma \subseteqq \sigma'$, *so ist*

(a) *jede σ-widerspruchsfreie Menge σ'-widerspruchsfrei,*

(b) *jede σ-unabhängige Menge σ'-unabhängig,*

(c) *jede σ'-vollständige Menge σ-vollständig,*

(d) *jede σ'-kategorische und σ-widerspruchsfreie Menge σ-kategorisch.*

(1) bedeutet: Eine Verkleinerung der Klasse der zulässigen Interpretationen wirkt sich im Sinne einer Vergrößerung der Folgerungshülle aus. Im gleichen Sinne wirkt aber auch eine Vergrößerung des benutzten Axiomensystems. Dies führt uns zu

Definition 5. Eine nichtelementare Sprache $(\boldsymbol{S}, \sigma)$ heißt *unwesentlich nichtelementar*, wenn ein Axiomensystem $\boldsymbol{Y} \subseteqq \boldsymbol{S}$ existiert, so daß

$$\boldsymbol{Fl_S}^{\sigma}(\boldsymbol{X}) = \boldsymbol{Fl_S}(\boldsymbol{X} \cup \boldsymbol{Y}) \qquad \textit{für alle } \boldsymbol{X} \subseteqq \boldsymbol{S}$$

gilt (d. h., wenn man das nichtelementare σ-Folgern gleichwertig durch elementares Folgern aus einem stärkeren Axiomensystem ersetzen kann).

Offenbar ist eine Sprache $(\boldsymbol{S}, \sigma)$ insbesondere dann unwesentlich nichtelementar, wenn es ein Axiomensystem $\boldsymbol{Y} \subseteqq \boldsymbol{S}$ gibt, so daß σ die Klasse aller Modelle von $\boldsymbol{Y}$ ist (oder wenigstens zu jeder Interpretation $\omega \in \sigma$ ein elementar ununterscheidbares Modell von $\boldsymbol{Y}$ und umgekehrt zu jedem Modell von $\boldsymbol{Y}$ eine elementar ununterscheidbare Interpretation $\omega \in \sigma$ existiert).

Beispiel 3. Die Sprachen $(\boldsymbol{S}, \sigma_{(1)})$ sind unwesentlich nichtelementar. Es sei nämlich $\boldsymbol{Y}$ die Menge aller Ausdrücke der Form

$$\bigwedge \mathrm{x}_1^{j_1} \ldots \mathrm{x}_1^{j_k} \bigvee \mathrm{x}_1^{j} \mathrm{F}_i^{j_1,\ldots,j_k;j}(\mathrm{x}_1^{j_1}, \ldots, \mathrm{x}_1^{j_k}) = \mathrm{x}_1^{j}, \tag{2}$$

wobei $\mathrm{F}_i^{j_1,\ldots,j_k;j}$ die Operationssymbole der Sprache $\boldsymbol{S}$ durchläuft. Dann ist $\sigma_{(1)}$ gerade die Klasse aller Modelle von $\boldsymbol{Y}$. An dieser Stelle sei bemerkt, daß in den meisten Lehrbüchern der mathematischen Logik $\boldsymbol{Fl_S}^{\sigma_{(1)}}$ statt $\boldsymbol{Fl_S}^{\sigma_\bullet}$ als elementare Folgerungsoperation definiert wird.

Beispiel 4. Die in Beispiel 2 betrachteten Sprachen sind unwesentlich nichtelementar. Dazu betrachten wir folgende in der betreffenden Sprache $\boldsymbol{S}$ selbst formulierbare

Axiome über die Dinge der Sorte f_i und die Terme der Form $f_i(x_j)$:

$$\bigwedge f_1 f_2 \left(\bigwedge x\, f_1(x) = f_2(x) \rightarrow f_1 = f_2\right), \tag{3}$$

(d. h., die Dinge f_i sind durch ihren „Werteverlauf" eindeutig bestimmt),

$$\bigwedge f_1 x_1 \bigvee x_2\, f_1(x_1) = x_2 \tag{4}$$

(vgl. (2), d. h., die f_i sind volle Abbildungen). Es sei Y die Menge der beiden Axiome (3), (4). Unter den Modellen von Y kommen alle zulässigen Interpretationen der betrachteten Sprache vor. Umgekehrt brauchen in einem beliebigen Modell von Y die f_i nicht selbst Abbildungen zu sein. Man kann jedoch zu jedem solchen Modell ein isomorphes zulässiges Modell konstruieren (vgl. hierzu die Konstruktion zum Beweis von Satz 1 in 4.4.). Es sei bemerkt, daß man weitere Einschränkungen der Klasse der zulässigen Interpretationen, die sich in diesem Fall anbieten, ebenfalls durch Axiome zum Ausdruck bringen kann. Zum Beispiel bedeutet

$$\bigwedge f_1 x_1 x_2 \left(f_1(x_1) = f_1(x_2) \rightarrow x_1 = x_2\right),$$

daß die f_i eineindeutig sind, während das Axiom

$$\bigwedge f_1 x_1 \bigvee x_2\, f_1(x_2) = x_1$$

sichert, daß die f_i Abbildungen **auf** den entsprechenden Grundbereich sind. Es ist jedoch nicht möglich, durch elementar formulierte Axiome auszudrücken, daß der den Variablen f_i zugeordnete Grundbereich aus **allen** Abbildungen des den Variablen x_i zugeordneten Grundbereichs in sich bestehen soll.

Als wichtiges Beispiel einer wesentlich nichtelementaren Sprache behandeln wir im folgenden die Formalisierung des bekannten Axiomensystems von Peano zur Charakterisierung der natürlichen Zahlen mittels der Grundbegriffe Null und Nachfolgeroperation (vgl. MfL, Band 1, 3.2.).

Beispiel 5. Die Axiome von Peano kann man — zunächst unformalisiert — etwa wie folgt formulieren:

(a) Null ist eine natürliche Zahl.

(b) Jede natürliche Zahl hat genau einen Nachfolger.

(c) Null hat keinen Vorgänger.

(d) Jede natürliche Zahl hat höchstens einen Vorgänger.

(e) Enthält eine beliebige Menge M von natürlichen Zahlen die Zahl 0 und mit einer beliebigen Zahl auch deren Nachfolger, so enthält M alle natürlichen Zahlen.

Zur Formalisierung dieser Axiome benötigt man Variablen n_i für natürliche Zahlen, Variablen M_i für Mengen von natürlichen Zahlen, ein Konstantensymbol o zur Bezeichnung der Zahl 0, ein Operationszeichen $'$ zur Bezeichnung der Nachfolgeroperation und ein Symbol $\in$ zur Bezeichnung der Elementrelation (zwischen Dingen der Sorte n_i und Dingen der Sorte M_i). Die Termdefinition der entsprechenden modifizierten zweisortigen Sprache lautet demnach:

M_i *(und nur diese) sind Terme der Sorte „Menge".*

o *ist ein Term der Sorte „Zahl".*

n_i *sind Terme der Sorte „Zahl".*

Ist t *ein Term der Sorte „Zahl", so ist auch* t′ *ein Term der Sorte „Zahl".*

Prädikative Ausdrücke sind $\mathrm{t}_i = \mathrm{t}_j$ (t_i, t_j sortengleiche Terme), $\mathrm{t}_i \in \mathrm{M}_j$ (t_i Term der Sorte „Zahl") und keine weiteren. Wir kommen zur Formalisierung der Axiome (a) bis (e):

(a′) entfällt auf Grund der syntaktischen Festlegungen.

(b′) $\bigwedge \mathrm{n}_1 \bigvee \mathrm{n}_2\, \mathrm{n}_1' = \mathrm{n}_2$.

(Daß der Nachfolger einer natürlichen Zahl eindeutig bestimmt ist, ergibt sich aus der generellen Festlegung über die Interpretation von Operationssymbolen. Daß — wie in (b′) verlangt — jede natürliche Zahl wenigstens einen Nachfolger hat, muß gefordert werden, da wir allgemein die Interpretation von Operationssymbolen durch partielle Operationen zulassen. (b′) entfällt natürlich, wenn man, wie meist üblich, $Fl^{\sigma(1)}$ als elementares Folgern ansieht.)

(c′) $\neg \bigvee \mathrm{n}_1\, \mathrm{n}_1' = \mathrm{o}$.

(d′) $\mathrm{n}_1' = \mathrm{n}_2' \rightarrow \mathrm{n}_1 = \mathrm{n}_2$.

(e′) $\mathrm{o} \in \mathrm{M}_1 \wedge \bigwedge \mathrm{n}_1(\mathrm{n}_1 \in \mathrm{M}_1 \rightarrow \mathrm{n}_1' \in \mathrm{M}_1) \rightarrow \bigwedge \mathrm{n}_2\, \mathrm{n}_2 \in \mathrm{M}_1$.

Die Kategorizität dieses Axiomensystems beruht darauf, daß die Variablen M_i (die nur in (e′), dem *Induktionsaxiom*, vorkommen) und das Symbol $\in$ in der in Beispiel 1 angegebenen Weise interpretiert werden, daß also die verwendete Sprache durch diese (mehr oder weniger stillschweigend vereinbarte) Interpretationsvorschrift zu einer nichtelementaren Sprache $(S, \sigma_{(2)})$ zweiter Stufe wird. Ein „naiver" Kategorizitätsbeweis kann dann leicht geführt werden, indem man zu zwei beliebigen Modellen $(\mathbf{N}_1, {}'^1, \mathrm{o}_1)$ und $(\mathbf{N}_2, {}'^2, \mathrm{o}_2)$ induktiv einen Isomorphismus φ von $\mathbf{N}_1$ auf $\mathbf{N}_2$ definiert:

$$\varphi(\mathrm{o}_1) := \mathrm{o}_2, \quad \varphi(\nu'^1) := \big(\varphi(\nu)\big)'^2. \tag{5}$$

Der exakte Beweis allerdings, daß durch (5) eine eindeutig bestimmte eineindeutige Abbildung von $\mathbf{N}_1$ auf $\mathbf{N}_2$ definiert ist, erfordert neben (e′) die von der axiomatischen Mengenlehre bereitgestellten sogenannten Rechtfertigungssätze für induktive Definitionen und deren Beweise und soll daher hier nicht geführt werden. Nötig sind jedoch noch folgende Bemerkungen.

Zum Beweis der Kategorizität (hier genauer: $\sigma_{(2)}$-Kategorizität) ist (vgl. Definition 4) zunächst die $\sigma_{(2)}$-Widerspruchsfreiheit des betrachteten Axiomensystems zu zeigen. Dies kann (wegen der Unendlichkeit des angestrebten Modells) natürlich nur im Rahmen einer anderen mathematischen Theorie geschehen. Üblicherweise konstruiert man in einer (als widerspruchsfrei vorausgesetzten) axiomatischen Mengenlehre einen Bereich **N** von natürlichen Zahlen, die (a′) bis (e′) erfüllen, indem man z. B. die Kardinalzahlen der endlichen Mengen als natürliche Zahlen nimmt. Man zeigt also nur die relative Widerspruchsfreiheit der Peanoschen Axiome in bezug auf das vorausgesetzte Axiomensystem der Mengenlehre, und nur unter der Voraussetzung der Widerspruchsfreiheit dieses Axiomensystems gelingt auch der exakte Kategorizitätsbeweis.

Man kann in naheliegender Weise ein elementares Axiomensystem für die natürlichen Zahlen erhalten, indem man das Induktionsaxiom (e′) durch ein Schema ersetzt, in welchem die in (e′) für beliebige Mengen von natürlichen Zahlen formulierte Aussage nur für alle in der Sprache selbst beschreibbaren Mengen gefordert wird. Es sei also S diejenige einsortige elementare Sprache, in der die Axiome (a′) bis (d′) formuliert sind. X sei das Axiomensystem, das aus (a′) bis (d′) und allen Ausdrücken der Form

$$\text{(e}''\text{)} \qquad \mathrm{H(o)} \wedge \bigwedge \mathrm{n_1}\big(\mathrm{H(n_1)} \rightarrow \mathrm{H(n_1{}')}\big) \rightarrow \bigwedge \mathrm{n_2 H(n_2)}$$

besteht, wobei $\mathrm{H} \in S$ ein beliebiger Ausdruck ist, der eine Variable vollfrei enthält. Die Theorie $Fl_S(X)$ wird als *elementare Nachfolgertheorie der natürlichen Zahlen* bezeichnet. Aus dem Satz von Tarski folgt, daß sie nicht kategorisch ist (also „Nichtstandardmodelle" besitzt) und auch durch Hinzunahme beliebig vieler weiterer in S formulierbarer Axiome nicht kategorisch gemacht werden kann.

4.6. Definitorische Spracherweiterungen

Definition 1. Sind (S_1, σ_1) und (S_2, σ_2) beliebige Sprachen, so heißt (S_2, σ_2) eine *Erweiterung* von (S_1, σ_1) und (S_1, σ_1) eine *Teilsprache* von (S_2, σ_2), wenn gilt:

(a) $S_1 \subset S_2$;

(b) die Einschränkung einer beliebigen Interpretation $\omega \in \sigma_2$ auf S_1 ist eine zulässige Interpretation von S_1 (d. h. Element von σ_1).

Dabei heißt (S_2, σ_2) eine *Erweiterung erster Art* von (S_1, σ_1), wenn die Sprachen S_1 und S_2 die gleichen Variablensorten besitzen, also in S_2 nur zusätzliche Relations-, Operations- oder Konstantensymbole auftreten. Andernfalls (d. h., wenn S_2 wenigstens eine Variablensorte enthält, die in S_1 nicht vorkommt) heißt (S_2, σ_2) eine *Erweiterung zweiter Art*.

Beschränkt man sich auf die Betrachtung elementarer Sprachen, so werden die Begriffe Teil- bzw. Erweiterungssprache durch die einfache Beziehung (a) vollständig erfaßt. Die Unterscheidung zwischen Erweiterungen erster und zweiter

Art ist auch für diesen Fall von Bedeutung. Wir beschäftigen uns in diesem Abschnitt mit Spracherweiterungen spezieller Art, die man als *definitorisch* (oder auch als *unwesentlich*) bezeichnet. Zunächst behandeln wir den Fall, daß die Erweiterung sich nur durch ein zusätzliches Relationssymbol von der ursprünglichen Sprache unterscheidet. Da man sich grundsätzlich auf die Betrachtung reiner Relationsstrukturen beschränken kann (vgl. 3.2.), umfaßt dieser Fall im Prinzip alle denkbaren Fälle und wird daher in den meisten Lehrbüchern der mathematischen Logik als einzige Form der definitorischen Spracherweiterung angesehen. Wir werden jedoch im Anschluß an diesen Spezialfall eine sehr allgemeine Definition des Begriffs definitorische Spracherweiterung angeben und zeigen, daß einige für die praktische Handhabung formalisierter Sprachen wichtige Definitionsmethoden sich als Spezialfälle dieses Begriffs erweisen.

Es sei $\boldsymbol{S}$ eine formalisierte Sprache, $\mathrm{H}(\mathrm{x}_1, \ldots, \mathrm{x}_n)$ ein Ausdruck dieser Sprache, der genau die angegebenen Variablen vollfrei und keine weiteren freien Variablen enthält, R ein (eventuell modifiziertes) Relationssymbol, das in $\boldsymbol{S}$ noch nicht vorkommt. Unter einer *korrekten Definition* verstehen wir eine Zeichenreihe der Form

$$\mathrm{R}(\mathrm{x}_1, \ldots, \mathrm{x}_n) :\leftrightarrow \mathrm{H}(\mathrm{x}_1, \ldots, \mathrm{x}_n). \tag{1}$$

Zu Definitionen dieser Form sind folgende Standpunkte möglich:

(a) Das im folgenden verwendete Symbol R hat lediglich den Charakter einer Abkürzung, d. h., „Ausdrücke“, in denen dieses Symbol vorkommt, sind nicht im strengen Sinn Ausdrücke einer formalisierten Sprache, sondern in ähnlicher Weise Abkürzungen korrekter Ausdrücke, wie dies für Zeichenreihen zutrifft, die unter Benutzung der in 3.4. vereinbarten Regeln geschrieben werden.

(b) Es bezeichne $\boldsymbol{S}'$ diejenige Erweiterung von $\boldsymbol{S}$, die durch Hinzunahme des Relationssymbols R entsteht. (Der Typ von R ist durch Anzahl und Sorte der im definierenden Ausdruck H vollfrei vorkommenden Variablen bestimmt.) Bei der Behandlung eines beliebigen Axiomensystems $\boldsymbol{X} \subseteq \boldsymbol{S}$ werde vom Zeitpunkt der Niederschrift der Definition (1) an die Sprache $\boldsymbol{S}'$ benutzt und gleichzeitig das Axiomensystem

$$\boldsymbol{X}' = \boldsymbol{X} \cup \{\mathrm{R}(\mathrm{x}_1, \ldots, \mathrm{x}_n) \leftrightarrow \mathrm{H}(\mathrm{x}_1, \ldots, \mathrm{x}_n)\}$$

zugrunde gelegt. Das hinzugefügte Axiom bezeichnet man in diesem Fall als *Einführungsaxiom* für den Begriff R. Ist ω ein Modell dieses Einführungsaxioms, so ist die Relation $\omega(\mathrm{R})$ offenbar eindeutig bestimmt, wenn die Interpretation der ursprünglichen Sprache $\boldsymbol{S}$ gegeben ist. Ist umgekehrt in einer (elementaren oder nichtelementaren) Theorie $\boldsymbol{T}$ ein Ausdruck der Form

$$\mathrm{R}(\mathrm{x}_1, \ldots, \mathrm{x}_n) \leftrightarrow \mathrm{H}(\mathrm{x}_1, \ldots, \mathrm{x}_n)$$

allgemeingültig, wobei das Relationssymbol R nicht in $\mathrm{H}(\mathrm{x}_1, \ldots, \mathrm{x}_n)$ vorkommt, so kann man das Symbol R in jedem Ausdruck eliminieren: H' enthalte einen Teilausdruck der Form $R(t_{i_1}, \ldots, t_{i_n})$, wobei t_{i_ν} Terme entsprechender Sorte sind. Durch gebundene Umbenennungen erhält man einen zu $H(x_1, \ldots, x_n)$ semantisch äquivalenten Ausdruck $H_1(x_1, \ldots, x_n)$, in dem keine in H' vorkommende Variable gebunden vorkommt. H_2 bezeichne den Ausdruck $Sub\,(H_1; x_1, t_{i_1}; \ldots; x_n, t_{i_n})$, der aus H_1 entsteht, indem man jede der Variablen x_ν $(\nu = 1, \ldots, n)$ durch den Term t_{i_ν} ersetzt. Dann sind in der betrachteten Theorie $\boldsymbol{T}$ auch die Ausdrücke $R(t_{i_1}, \ldots, t_{i_n}) \leftrightarrow H_2$ und

$$H' \leftrightarrow Sub\left(H', R(t_{i_1}, \ldots, t_{i_n}), H_2\right) \tag{2}$$

allgemeingültig, wobei das Relationssymbol R im rechten Teilausdruck von (2) mindestens einmal weniger als in H' vorkommt. Indem man eventuelle weitere Vorkommen von R (mit anderen Termkombinationen) in gleicher Weise ersetzt, erhält man in endlich vielen Schritten einen zu H' in der Theorie $\boldsymbol{T}$ gleichwertigen Ausdruck H'' (d. h., $H' \leftrightarrow H''$ ist in $\boldsymbol{T}$ allgemeingültig), in dem R nicht vorkommt. Das bedeutet, daß der dem Symbol R entsprechende Begriff beim axiomatischen Aufbau dieser Theorie prinzipiell entbehrlich ist. Man kann zu einer um das Symbol R reduzierten Sprache übergehen und R mit Hilfe von (1) sofort wieder definitorisch einführen, wobei man sich nach Belieben auf den Standpunkt (a) oder (b) stellen kann.

Beispiele. Es sei $\boldsymbol{S}$ die durch Punktvariablen (große lateinische Buchstaben), Geradenvariablen (kleine lateinische Buchstaben) und prädikative Ausdrücke der Form P auf g charakterisierte modifizierte Sprache der Inzidenzgeometrie (vgl. 4.4., Beispiel 2). Wir definieren

$$P \neq Q :\leftrightarrow \neg\, P = Q; \tag{3}$$

$$A, B, C \text{ paarweise verschieden } :\leftrightarrow A \neq B \wedge A \neq C \wedge B \neq C; \tag{4}$$

$$A, B, C \text{ kollinear } :\leftrightarrow \bigvee g(A \text{ auf } g \wedge B \text{ auf } g \wedge C \text{ auf } g); \tag{5}$$

$$A, B, C \text{ nicht kollinear } :\leftrightarrow \neg\, A, B, C \text{ kollinear}; \tag{6}$$

$$g \| h :\leftrightarrow \neg \bigvee P(P \text{ auf } g \wedge P \text{ auf } h). \tag{7}$$

Bei (3) kommt der Abkürzungscharakter der Definitionen besonders klar zum Ausdruck. Demgegenüber ist bei (4) der definitorisch eingeführte Ausdruck sogar länger als der ihn definierende Ausdruck. (4) zahlt sich als Abkürzung erst aus, wenn man nach dem Muster dieser Definition fortfährt:

$$\begin{aligned} &P_1, \ldots, P_n \text{ paarweise verschieden } :\leftrightarrow P_1 \neq P_2 \wedge P_1 \neq P_3 \wedge \ldots \\ &\wedge P_1 \neq P_n \wedge P_2 \neq P_3 \wedge \ldots \wedge P_{n-1} \neq P_n. \end{aligned} \tag{8}$$

Es sei jedoch nachdrücklich betont, daß die Zeichenreihe (8) keinesfalls als eine korrekte Definition im Sinne von (1) aufzufassen ist, sondern lediglich als eine (allerdings wohl unmißverständliche) Zusammenfassung unendlich vieler zu (4) ähnlicher korrekter Einzeldefinitionen. Ähnlich könnte man (5) verallgemeinern zu

$$P_1, \ldots, P_n \text{ kollinear} :\leftrightarrow \bigvee g(P_1 \text{ auf } g \wedge \ldots \wedge P_n \text{ auf } g). \tag{9}$$

Das Problem der Definitionsschemata nach Art von (8) und (9) besteht darin, daß die Art und Weise, wie der definierende Ausdruck in Abhängigkeit von der Zahl der Variablen zu bilden ist, von Fall zu Fall verschieden ist, so daß die richtige Deutung der in ihm vorkommenden „...“ beim Leser stets Intelligenz und guten Willen voraussetzt. Definition (6) zeigt, daß man durch hinreichend viele Definitionen auch in formalisierten Sprachen eine gewisse sprachliche Glätte erreichen kann. Schließlich wirft (7) das Problem der „richtigen“ Definition auf. In diesem Fall ist eine zweistellige Geradenrelation Parallelität inhaltlich schon vorgegeben, und es handelt sich eigentlich darum, ihre Entbehrlichkeit im Rahmen der Grundbegriffe einer axiomatischen Geometrie dadurch nachzuweisen, daß man sie durch eine formal korrekte Definition inhaltlich richtig erfaßt. Formal korrekt wäre (7) auch in der räumlichen Inzidenzgeometrie. Der so definierte Parallelitätsbegriff für Geraden deckt sich dann aber nicht mehr mit dem anschaulichen Parallelitätsbegriff, da er auch windschiefe Geraden umfaßt. Inhaltlich richtig ist in diesem Fall z. B.

$$g \| h :\leftrightarrow \neg \bigvee P(P \text{ auf } g \wedge P \text{ auf } h) \wedge \bigvee \alpha \bigwedge Q(Q \text{ auf } g \vee Q \text{ auf } h \rightarrow Q \text{ auf } \alpha),$$

wobei α eine Ebenenvariable ist und „auf“ in diesem Zusammenhang die Inzidenz zwischen Punkten und Ebenen bezeichnet.

Um von der zunächst behandelten definitorischen Einführung von Relationssymbolen zum angekündigten allgemeinen Begriff der definitorischen Spracherweiterung zu gelangen, bemerken wir zunächst, daß wir uns im folgenden durchweg auf den Standpunkt (b) stellen werden (ohne dem anderen Standpunkt seine grundsätzliche Berechtigung abzusprechen). Bei der Verallgemeinerung dieses Standpunktes auf andere Definitionsmethoden macht es sich allerdings störend bemerkbar, daß die Definitionen, die ja eigentlich nur den Übergang zu einer anderen Sprache betreffen, durch die jeweiligen Einführungsaxiome mit der jeweils betrachteten Theorie verknüpft werden. Dem naiven Gebrauch von Definitionen in der mathematischen Praxis entspricht besser die Vorstellung, daß eine Definition im Übergang zu einer Erweiterungssprache besteht, wobei den neuen Sprachbestandteilen eine Interpretationsvorschrift mitgegeben wird, die Erweiterungssprache also in jedem Fall zunächst nichtelementar ist. Die Einführungsaxiome zeigen dann nachträglich, daß die so erhaltenen Sprachen unwesentlich nichtelementar sind.

Definition 2. Eine Erweiterung (S_2, σ_2) einer Sprache (S_1, σ_1) heißt *definitorisch*, wenn gilt:

(a) Jede zulässige Interpretation $\omega \in \sigma_1$ der ursprünglichen Sprache S_1 läßt sich zu einer zulässigen Interpretation $\omega' \in \sigma_2$ der erweiterten Sprache S_2 fortsetzen, die bis auf Isomorphie eindeutig durch ω bestimmt ist, d. h., sind $\omega', \omega'' \in \sigma_2$ Interpretationen von S_2, deren Einschränkungen auf die Sprache S_1 mit $\omega \in \sigma_1$ übereinstimmen, so existiert ein Isomorphismus φ von ω' auf ω'' (vgl. S. 68) mit $\varphi(\xi) = \xi$ für alle ξ aus den gemeinsamen (schon zu ω gehörenden) Grundbereichen.

(Für den Fall einer definitorischen Erweiterung erster Art läßt sich Bedingung (a) wesentlich kürzer formulieren: Jede zulässige Interpretation $\omega \in \sigma_1$ der ursprünglichen Sprache S_1 läßt sich auf genau eine Weise zu einer zulässigen Interpretation $\omega' \in \sigma_2$ von S_2 fortsetzen.)

(b) (S_2, σ_2) ist unwesentlich nichtelementar in bezug auf (S_1, σ_1), d. h., es existiert ein System $E \subset S_2$ von *Einführungsaxiomen*, so daß σ_2 die Klasse aller Fortsetzungen von Interpretationen aus σ_1 ist, die Modelle von E sind.

(c) Zu jedem Ausdruck $H \in S_2$, der höchstens solche Variablen frei enthält, die bereits zu S_1 gehören, existiert ein Ausdruck $[H] \in S_1$ mit den gleichen freien Variablen (als *Rückübersetzung* von H bezeichnet), so daß

$$H \leftrightarrow [H] \in Fl_{S_2}^{\sigma_2}(\emptyset)$$

gilt.

Ist insbesondere die Ausgangssprache S_1 elementar und E das System der Einführungsaxiome, so läßt sich die in (c) an die Rückübersetzung gestellte Forderung in der Form

$$H \leftrightarrow [H] \in Fl_{S_2}(E)$$

schreiben. Wie man sieht, ist dies trivialerweise erfüllt, wenn man bei gegebener Rückübersetzungsvorschrift die Gesamtheit der Ausdrücke der Form $H \leftrightarrow [H]$ als Einführungsaxiome nimmt. In den folgenden Spezialfällen definitorischer Erweiterungen werden wir dieses System jedoch immer auf ein endliches System von Einführungsaxiomen (häufig sogar auf ein einzelnes Axiom) reduzieren können.

Wie man sieht, bringt Bedingung (a) der Definition 2 den Definitionsaspekt der definitorischen Erweiterungen zum Ausdruck, d. h. die Forderung, daß die Bedeutung der definierten Begriffe im wesentlichen eindeutig festliegt, wenn man eine beliebige Interpretation der ursprünglichen Sprache vorgibt. Demgegenüber bringt (c) den Aspekt der Entbehrlichkeit zum Ausdruck, d. h. die Forderung, daß man prinzipiell ohne die definitorisch eingeführten Sprachbestandteile auskommen kann. Es sei bemerkt, daß bei der praktischen Handhabung halbformalisierter Sprachen zum Aufbau konkreter Theorien häufig nur der Aspekt (a) beachtet wird, so daß im Laufe der Zeit Begriffe „definiert“ werden, die die Ausdrucksfähigkeit der benutzten Sprache wesentlich erweitern (mit anderen Worten, die der Forderung (c) nicht genügen). Zum Beispiel kann man in einen Aufbau der ebenen euklidischen Geometrie scheinbar unverfänglich die „Definition“ einbauen:

> Variablen M_i bezeichnen im folgenden beliebige Punktmengen, und $P \in M$ bedeute das Liegen des Punktes P in der Menge M.

Es ist klar, daß sich (zumindest von einem metatheoretisch naiven Standpunkt aus) jede Interpretation der elementaren Ausgangssprache zu einer eindeutig bestimmten Interpretation der neuen Sprachbestandteile fortsetzen läßt. Die Erweiterungssprache ist jedoch nichtelementar und wesentlich ausdrucksfähiger, da man in ihr z. B. das Dedekindsche Stetigkeitsaxiom formulieren und damit ein kategorisches Axiomensystem für die euklidische Ebene aufstellen kann.

Als Folgerung aus Definition 2 formulieren wir (hier ohne Beweis)

Satz 1. *Ist* (S_2, σ_2) *eine definitorische Erweiterung von* (S_1, σ_1) *und* $X \subseteq S_1$ $(\subseteq S_2)$ *ein* σ_1*-kategorisches Axiomensystem, so ist* X *auch* σ_2*-kategorisch.*

Der Beweis sei dem Leser zur Übung empfohlen.

Es dürfte klar sein, in welcher Weise die bisher behandelten definitorischen Einführungen neuer Relationssymbole Spezialfälle des in Definition 2 erfaßten allgemeinen Begriffs der definitorischen Spracherweiterung sind. Wir behandeln anschließend verschiedene Formen der definitorischen Einführung von Operations- und Konstantensymbolen und einen häufig vorkommenden Typ definitorischer Erweiterungen zweiter Art. Während die diskutierten Formen der definitorischen Erweiterung erster Art bei mehrfacher Wiederholung im wesentlichen den Begriff der definitorischen Erweiterung erster Art ausschöpfen, sind die Möglichkeiten für definitorische Erweiterungen zweiter Art in gewissem Sinn uferlos, und der behandelte Typ ist nicht mehr als ein Beispiel.

Es sei S eine formalisierte Sprache, $t_1(x_1, \ldots, x_n)$, $t_2(x_1, \ldots, x_n)$ seien sortengleiche Terme dieser Sprache, die genau die angegebenen Variablen enthalten, H sei ein Ausdruck, der höchstens die Variablen $x_1, \ldots, x_n$ vollfrei und keine weiteren Variablen frei enthält, F sei ein in S noch nicht vorkommendes Operationssymbol. Eine *korrekte Definition durch Superposition,* zum Ausdruck gebracht durch

$$F(x_1, \ldots, x_n) := t_1(x_1, \ldots, x_n),$$

besteht im Übergang zu der durch das Symbol F erweiterten Sprache und im Übergang von der Klasse der bisher zulässigen Interpretation zur Teilklasse derjenigen Interpretationen, die außerdem Modelle des Einführungsaxioms

$$F(x_1, \ldots, x_n) = t_1(x_1, \ldots, x_n)$$

für F sind. (Auf Grund der in 4.1., Definitionen 3 und 4, getroffenen Wertfestsetzung für Terme und Termgleichungen ist die so definierte Funktion F bei beliebigen Interpretationen der ursprünglichen Sprache im allgemeinen partiell definiert, und zwar für genau diejenigen Werte, für die der definierende Term einen Wert hat.) Eine *korrekte Definition durch Fallunterscheidung*, zum Ausdruck gebracht durch die Gleichung

$$F(x_1, \ldots, x_n) := \begin{cases} t_1(x_1, \ldots, x_n), & \textit{falls } H, \\ t_2(x_1, \ldots, x_n), & \textit{falls } \neg H, \end{cases}$$

besteht im Übergang zu der durch das Symbol F erweiterten Sprache und im Übergang von der Klasse der bisher zulässigen Interpretationen zur Teilklasse derjenigen Interpretationen, die außerdem Modelle des Einführungsaxioms

$$y = F(x_1, \ldots, x_n) \leftrightarrow \big(H \wedge y = t_1(x_1, \ldots, x_n)\big) \vee \big(\neg H \wedge y = t_2(x_1, \ldots, x_n)\big)$$

für F sind. (Für y ist eine von $x_1, \ldots, x_n$ verschiedene Variable zu nehmen, deren Sorte sich aus der Sorte der Terme t_1, t_2 ergibt.)

Die Rückübersetzbarkeit ist im Fall der Definition durch Superposition unmittelbar klar. Bei den Definitionen durch Fallunterscheidung hat man nur zu überlegen, wie man einen beliebigen Ausdruck, der das Symbol F enthält, semantisch äquivalent in einen Ausdruck umformen kann, in dem F nur noch in Gleichungen der Form $y = F(t_1, \ldots, t_n)$ auftritt (deren Eliminierbarkeit sich wieder sofort aus dem Einführungsaxiom für F ergibt). Dazu ersetze man jeden prädikativen Teilausdruck H, der einen Term der Form $F(t_1, \ldots, t_n)$ enthält, durch den semantisch äquivalenten Teilausdruck

$$\bigvee y\big(y = F(t_1, \ldots, t_n) \wedge \mathit{Sub}\,\big(H, F(t_1, \ldots, t_n), y\big)\big),$$

wobei für y eine im Gesamtausdruck nicht vorkommende Variable entsprechender Sorte zu nehmen ist.

Eine Möglichkeit der definitorischen Einführung von Konstanten ergibt sich als Spezialfall ($n = 0$) der Definition durch Superposition. Ist t ein variablenfreier Term einer Sprache S, c ein in S noch nicht vorkommendes Konstantensymbol, so ist $c := t$ eine korrekte Definition. Sie besteht im Übergang zu der durch c erweiterten Sprache und in der Einschränkung der bisher zulässigen Interpretationen auf solche, die Modelle des Einführungsaxioms $c = t$ sind.[1]) Als Anwendung der letzten Methode erhält man z. B. in der Sprache der Peano-Arithmetik (vgl. 4.5., Beispiel 5) aus der zunächst einzigen Konstanten o fortlaufend

$$1 := o', \quad 2 := 1', \quad 3 := 2', \ldots$$

Die bisher betrachteten Definitionsmethoden für Operationen und Konstanten sind nur anwendbar, falls die ursprüngliche Sprache schon solche Symbole enthält.

[1]) Spätestens hier muß man die Forderung fallenlassen, daß jede Interpretation jedem Konstantensymbol ein Objekt zuordnet (vgl. die Fußnote auf S. 52). Ist ω eine Interpretation von S, für die *Wert* (t, ω) nicht existiert, so läßt sich ω nicht zu einer Interpretation von c fortsetzen, die das Einführungsaxiom $c = t$ erfüllt.

Wir kommen nun zur Definition von Operationen und Konstanten aus Relationen, die die oben behandelten Definitionsarten und viele weitere als Spezialfälle umfaßt.

Es sei $\boldsymbol{S}$ eine Sprache, die das Symbol F noch nicht enthält, $H_1 \in \boldsymbol{S}$ ein Ausdruck, der höchstens die Variablen $x_1, \ldots, x_n$ vollfrei und keine weiteren freien Variablen enthält, $H_2 \in \boldsymbol{S}$ ein Ausdruck, der die Variable y vollfrei und außerdem höchstens $x_1, \ldots, x_n$ vollfrei und keine weiteren freien Variablen enthält. Dann ist die Zeichenreihe

$$F(x_1, \ldots, x_n) := \iota y H_2, \quad \textit{falls } H_1 \tag{10}$$

(zu lesen: ... dasjenige Ding y mit der Eigenschaft H_2) eine *korrekte Definition einer Operation durch bestimmten Artikel.*

Durch eine solche Definition wird die Sprache $\boldsymbol{S}$ um das Symbol F zu einer Sprache $\boldsymbol{S}'$ erweitert und gleichzeitig jede (zulässige) Interpretation ω von $\boldsymbol{S}$ zu derjenigen eindeutig bestimmten Interpretation des Symbols F fortgesetzt, die das Einführungsaxiom

$$y = F(x_1, \ldots, x_n) \leftrightarrow \textit{Gen}\,(H_1 \to \bigvee!!y H_2) \land H_1 \land H_2 \tag{11}$$

erfüllt, d. h. ein Modell dieses Axioms ist. Die tatsächliche Anwendung einer Definition der Form (10) erfolgt meist im Rahmen einer in der Sprache $\boldsymbol{S}$ formulierten Theorie $\boldsymbol{T}$, zu deren gültigen Sätzen der abgeschlossene Ausdruck $\textit{Gen}\,(H_1 \to \bigvee!!y H_2)$ gehört. In diesem Fall ist (11) äquivalent zu

$$y = F(x_1, \ldots, x_n) \leftrightarrow H_1 \land H_2,$$

d. h., das Symbol F soll bei jeder Interpretation der ursprünglichen Sprache, die ein Modell der betrachteten Theorie ist, diejenige Operation bezeichnen, deren Definitionsbereich durch den Ausdruck H_1 beschrieben wird und für die die Relation zwischen Argumenten und Wert durch den Ausdruck H_2 dargestellt wird. In einer Interpretation, in der der abgeschlossene Ausdruck $\textit{Gen}\,(H_1 \to \bigvee!!y H_2)$ falsch ist, stellt F auf Grund des Einführungsaxioms (11) offenbar eine nirgends definierte Operation dar.

Beispiel 1. Es sei $\boldsymbol{S}$ wieder die bereits betrachtete modifizierte Sprache der ebenen Inzidenzgeometrie. Unter Bezug auf die in der ebenen euklidischen Geometrie gültigen Ausdrücke

$$A \neq B \to \bigvee!!g(A \text{ auf } g \land B \text{ auf } g),$$
$$\neg P \text{ auf } g \to \bigvee!!h(P \text{ auf } h \land \neg \bigvee Q(Q \text{ auf } g \land Q \text{ auf } h))$$

definiert man

$$AB := \iota g(A \text{ auf } g \land B \text{ auf } g), \quad \textit{falls } A \neq B$$

und

$$\text{Parallele durch } P \text{ zu } g := \iota h(P \text{ auf } h \land \neg \bigvee Q(Q \text{ auf } g \land Q \text{ auf } h)),$$
$$\textit{falls } \neg P \text{ auf } g.$$

(In beiden Fällen ist der logische Charakter der so eingeführten neuen Sprachbestandteile als Operationssymbole durch den hier zunächst benutzten üblichen Sprachgebrauch etwas verwischt. Statt AB bzw. Parallele durch P zu g könnte man aber auch z. B. L(A, B) bzw. Par(P, g) schreiben.)

Beispiel 2. Es bezeichne S die durch das zweistellige Operationssymbol $+$ erzeugte einsortige Sprache (vgl. 4.3., Beispiel 1). Unter Bezug auf den in der Gruppentheorie gültigen Ausdruck

$$\bigvee !!\mathrm{x}\; \mathrm{a} + \mathrm{x} = \mathrm{b}$$

definiert man die zweistellige Operation *Subtraktion* durch

$$\mathrm{b} - \mathrm{a} := \iota\mathrm{x}\; \mathrm{a} + \mathrm{x} = \mathrm{b}\,.$$

(Die einschränkende Bedingung „..., *falls* H_1" kann hier offenbar unterdrückt werden, da man als Ausdruck H_1 einen beliebigen allgemeingültigen Ausdruck nehmen kann, der keine freien Variablen außer a, b enthält.) Demgegenüber führt die sonst analoge Definition der Division in der Körpertheorie schon wieder auf eine echt partielle Operation. Hier definiert man unter Bezug auf

$$\mathrm{a} \neq 0 \rightarrow \bigvee !!\mathrm{x}\; \mathrm{a} \cdot \mathrm{x} = \mathrm{b}$$

die mit dem Symbol : bezeichnete Division durch

$$\mathrm{b} : \mathrm{a} := \iota\mathrm{x}\; \mathrm{a} \cdot \mathrm{x} = \mathrm{b}, \quad \textit{falls}\; \mathrm{a} \neq 0\,.$$

Die Elimination von Operationssymbolen, die durch den bestimmten Artikel ι definitorisch eingeführt worden sind, kann man unter Benutzung des Einführungsaxioms (11) offenbar in analoger Weise durchführen, wie dies für den Fall der Definition durch Fallunterscheidung erklärt wurde. Die letzte Definitionsmethode ergibt sich als Spezialfall der Definition durch bestimmten Artikel, wenn man für H_2 den Ausdruck

$$\big(\mathrm{H} \wedge \mathrm{y} = \mathrm{t}_1(\mathrm{x}_1, \ldots, \mathrm{x}_n)\big) \vee \big(\neg\, \mathrm{H} \wedge \mathrm{y} = \mathrm{t}_2(\mathrm{x}_1, \ldots, \mathrm{x}_n)\big)$$

und für H_1 den Ausdruck $\bigvee \mathrm{y}\mathrm{H}_2$ setzt. Analog ergibt sich die Definition durch Superposition, indem man für H_2 den Ausdruck $\mathrm{y} = \mathrm{t}_1(\mathrm{x}_1, \ldots, \mathrm{x}_n)$ und für H_1 den Ausdruck $\bigvee \mathrm{y}\mathrm{H}_2$ setzt.

Die *Definition von Konstanten durch bestimmten Artikel* kann man als einen gewissen Ausartungsfall ($n = 0$) der Definition von Operationen durch bestimmten Artikel ansehen. In diesem Fall enthält der Ausdruck H_2 genau die Variable y vollfrei und keine weiteren freien Variablen. Da H_1 dementsprechend ein abgeschlossener Ausdruck sein müßte, können wir uns auf den Fall beschränken, daß seine Gültigkeit in der betrachteten Theorie vorausgesetzt wird. Die korrekte Definition eines neuen Konstantensymbols lautet dann

$$c := \iota y H_2(y).$$

Die zulässigen Interpretationen der entsprechenden Erweiterungssprache werden durch das Einführungsaxiom

$$y = c \leftrightarrow \bigvee !! y H_2(y) \wedge H_2(y)$$

charakterisiert. Hier macht sich wieder (vgl. die Fußnote auf S. 85) störend bemerkbar, daß das Symbol c in solchen Interpretationen ω der ursprünglichen Sprache, bei denen der Ausdruck $\bigvee !! y H_2(y)$ falsch wird, nicht interpretiert werden kann. Terme, in denen diese Konstante vorkommt, haben dann bei keiner Belegung einen Wert. Setzt man andererseits die Gültigkeit des Ausdrucks $\bigvee !! y H_2(y)$ voraus, so ist ein beliebiger Ausdruck H(c), der c enthält, gleichbedeutend mit dem Ausdruck

$$\bigvee y \big(H_2(y) \wedge \mathit{Sub}\,(H(c), c, y)\big),$$

c ist demnach sehr leicht zu eliminieren. (Für y ist eine in H(c) nicht vorkommende Variable zu nehmen.)

Beispiel 3. Wir nehmen an, die Sprache der Peano-Arithmetik (vgl. 4.5., Beispiel 5) sei zunächst ohne das Konstantensymbol o gegeben. Dann kann man unter Bezug auf das neu aufzunehmende Axiom

$$\bigvee !! n_1 \neg \bigvee n_2\, n_2' = n_1 \tag{12}$$

definieren

$$o := \iota n_1 \neg \bigvee n_2\, n_2' = n_1 .$$

Die Elimination der so definierten Konstanten o aus dem Axiom (c′)

$$\neg \bigvee n_1\, n_1' = o$$

nach der oben angegebenen Regel liefert

$$\bigvee n_3 (\neg \bigvee n_2\, n_2' = n_3 \wedge \neg \bigvee n_1\, n_1' = n_3),$$

was sich semantisch äquivalent zu

$$\bigvee n_3 \neg \bigvee n_2\, n_2' = n_3$$

vereinfachen läßt. In dieser Form ist es aber offenbar eine Folgerung aus dem neuen Axiom (12). Damit haben wir folgendes Ergebnis erhalten: Der Grundbegriff o ist für die Formulierung des kategorischen Axiomensystems der natürlichen Zahlen entbehrlich. Ersetzt man im ursprünglichen Axiomensystem (vgl. 4.5., Beispiel 5) das Axiom (c′) durch das Axiom (12) und eliminiert die Konstante o aus dem Induktionsaxiom nach der Regel, so erhält man ein gleichwertiges Axiomensystem, in dem die Konstante o nicht vorkommt. Führt man diese Konstante unter Berufung auf Axiom (12) definitorisch ein, so gibt es zu jedem Ausdruck H der Erweiterungssprache einen Ausdruck H', der die Konstante o nicht enthält, so daß $H \leftrightarrow H'$ aus dem Axiomensystem folgt.

Als Beispiel für definitorische Erweiterungen zweiter Art betrachten wir eine Definitionsmethode, die vor allem in der Geometrie eine wichtige Rolle spielt. Es sei S eine Sprache, $H_2 \in S$ ein Ausdruck, der genau die Variablen $y, x_1, \ldots, x_n$ ($n \geqq 1$) vollfrei und keine weiteren freien Variablen enthält, $H_1 \in S$ enthalte höchstens $x_1, \ldots, x_n$ vollfrei und keine weiteren freien Variablen. Wir erweitern S durch Variablen Y_0, Y_1, Y_2 einer neuen Sorte und prädikative Ausdrücke der Form $y_i \in Y_j$, wobei y_i zu y sortengleiche Variablen sind. Die Interpretation dieser neuen Sprachbestandteile wird durch folgende Einführungsaxiome festgelegt:

$$\bigwedge y(y \in Y_1 \leftrightarrow y \in Y_2) \rightarrow Y_1 = Y_2$$

(d. h., die Dinge der Sorte Y sind durch ihre „Elemente" eindeutig bestimmt, sie sind bis auf Isomorphie Mengen von Dingen der Sorte y).

$$H_1 \rightarrow \bigvee Y_1 \bigwedge y\big(y \in Y_1 \leftrightarrow H_2(y, x_1, \ldots, x_n)\big)$$

(d. h., daß sich alle „Mengen" der Form $\{y: H_2(y, x_1, \ldots, x_n)\}$ unter den Dingen der Sorte Y befinden, wobei die Parameter $x_1, \ldots, x_n$ die Bedingung H_1 erfüllen).

$$\bigwedge Y_1 \bigvee x_1 \ldots x_n\Big(H_1 \wedge \bigwedge y\big(y \in Y_1 \leftrightarrow H_2(y, x_1, \ldots, x_n)\big)\Big)$$

(d. h., außer den oben genannten Mengen gibt es keine weiteren unter den Dingen der Sorte Y).

Offensichtlich kann man jede Interpretation ω der Sprache S auf genau eine Weise so zu einer Interpretation ω' der erweiterten Sprache S' fortsetzen, daß diese drei Einführungsaxiome erfüllt sind. Daher ist die an beliebige definitorische Erweiterungen gestellte Bedingung (a) erfüllt. Rückübersetzbarkeit nichtabgeschlossener Ausdrücke von S' in S kann jetzt natürlich nicht mehr erreicht werden. Quantifizierte Variablen der Sorte Y können nach folgendem Verfahren eliminiert werden: Enthält H(Y) die Variable Y vollfrei, so seien x_ν' zu x_ν sortengleiche Variablen ($\nu = 1, \ldots, n$), die in H_1, H_2 und H nicht vorkommen. Es sei

$$H_i' := \mathit{Sub}\,(H_i; x_1, x_1'; \ldots; x_n, x_n') \qquad \text{für } i = 1, 2,$$

und H' entsteht aus H(Y), indem man jeden prädikativen Teilausdruck der Form $y_j \in Y$ durch den Ausdruck $\mathit{Sub}\,(H_2', y, y_j)$ ersetzt. Dann folgt aus den Einführungsaxiomen

$$\bigvee YH(Y) \leftrightarrow \bigvee x_1' \ldots x_n'(H_1' \wedge H'),$$

$$\bigwedge YH(Y) \leftrightarrow \bigwedge x_1' \ldots x_n'(H_1' \rightarrow H').$$

Beispiel 4. Wir erweitern die modifizierte Sprache der ebenen euklidischen Geometrie durch Variablen k_i ($i = 0, 1, 2, 3, \ldots$) für Kreise und prädikative Ausdrücke der Form $P \in k_i$ (P Punktvariable). Die Einführungsaxiome lauten in diesem Fall

$$\bigwedge P(P \in k_1 \leftrightarrow P \in k_2) \rightarrow k_1 = k_2,$$

$$B \neq C \rightarrow \bigvee k \bigwedge P(P \in k \leftrightarrow PA \cong BC),$$

$$\bigwedge k \bigvee BC(B \neq C \wedge \bigwedge P(P \in k \leftrightarrow PA \cong BC).$$

Sie bedeuten, daß die Variablen k_i bei einer beliebigen Interpretation der Grundbegriffe Punkt und $\cong$ (Streckenkongruenz) die Gesamtheit aller Punktmengen der Form $\{P : PA \cong BC\}$, wobei $B \neq C$ ist, als Grundbereich haben.

Beispiel 5. Wir gehen von einer Formalisierung der ebenen euklidischen Geometrie aus, bei der die Begriffe Gerade und Inzidenz zunächst nicht zu den Grundbegriffen gehören, dafür die dreistellige Punktrelation Kollinearität (die man ihrerseits durch die Zwischenrelation definitorisch einführen kann). Dann kann man die Geraden als Punktmengen der Form $\{P : P, Q, R \text{ kollinear}\}$, wobei $Q \neq R$, zusammen mit der Punkt-Geraden-Inzidenz $\in$ nach dem Muster des Beispiels 4 definitorisch einführen. Der Leser formuliere die Einführungsaxiome für diesen Fall! Als Beispiel für die Rückübersetzung abgeschlossener Ausdrücke eliminieren wir g und $\in$ aus dem wie folgt formulierten Axiom von Pasch

$$\bigwedge gABC\big(\neg\, A, B, C \text{ kollinear} \wedge \bigvee D(D \in g \wedge D \text{ zwischen } A, B) \wedge \neg\, A \in g$$
$$\wedge \neg\, B \in g \wedge \neg\, C \in g \rightarrow \bigvee E\big(E \in g \wedge (E \text{ zwischen } B, C \vee E \text{ zwischen } A, C)\big)\big).$$

Die nach der oben angegebenen Vorschrift gebildete Rückübersetzung lautet

$$\bigwedge PQ\big(P \neq Q \rightarrow \bigwedge ABC\big(\neg\, A, B, C \text{ kollinear} \wedge \bigvee D(D, P, Q \text{ kollinear}$$
$$\wedge\, D \text{ zwischen } A, B) \wedge \neg\, A, P, Q \text{ kollinear} \wedge \neg\, B, P, Q \text{ kollinear}$$
$$\wedge \neg\, C, P, Q \text{ kollinear} \rightarrow \bigvee E(E, P, Q \text{ kollinear} \wedge (E \text{ zwischen } B, C$$
$$\vee\, E \text{ zwischen } A, C))\big)\big).$$

Beispiel 6. Wir zeigen als komplexe Anwendung der definitorischen Spracherweiterungen, daß die Streckenkongruenz allein genügt, alle anderen Grundbegriffe der nach Hilbert formalisierten ebenen euklidischen Geometrie zu definieren. Auf Grund des Resultats aus Beispiel 5 genügt es zu zeigen:

(a) Kollinearität ist durch die Zwischenrelation definierbar.

(b) Die Zwischenrelation ist durch die Streckenkongruenz definierbar.

Man definiert:

$$A, B, C \text{ kollinear} :\leftrightarrow A = B \vee A = C \vee B = C \vee A \text{ zwischen } B, C$$
$$\vee\, B \text{ zwischen } A, C \vee C \text{ zwischen } A, B.$$

Zur bequemeren Formulierung einer Definition der Zwischenrelation definieren wir mittels Kongruenz zunächst den Hilfsbegriff R(A, B, C) (Bedeutung: A, B, C bilden ein bei B rechtwinkliges Dreieck):

$$\begin{aligned} \mathrm{R}(\mathrm{A}, \mathrm{B}, \mathrm{C}) :\leftrightarrow \bigvee \mathrm{DE}(&\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}, \mathrm{E} \text{ paarweise verschieden} \wedge \mathrm{AB} \cong \mathrm{BD} \\ &\wedge \mathrm{CB} \cong \mathrm{BE} \wedge \mathrm{AC} \cong \mathrm{CD} \wedge \mathrm{CD} \cong \mathrm{DE} \wedge \mathrm{DE} \cong \mathrm{EA}) \end{aligned}$$

$$\mathrm{B} \text{ zwischen } \mathrm{A}, \mathrm{C} :\leftrightarrow \bigvee \mathrm{D}\big(\mathrm{R}(\mathrm{A}, \mathrm{B}, \mathrm{D}) \wedge \mathrm{R}(\mathrm{C}, \mathrm{B}, \mathrm{D}) \wedge \mathrm{R}(\mathrm{A}, \mathrm{D}, \mathrm{C})\big)$$

(Inhaltliche Rechtfertigung durch den Satz des THALES. Der Leser veranschauliche sich den geometrischen Sinn beider Definitionen durch Skizzen!)

Der allgemeine Begriff der definitorischen Spracherweiterung ermöglicht es, die Gleichwertigkeit verschiedener Systeme von Grundbegriffen für den Aufbau einer axiomatischen Theorie genauso exakt zu fassen wie die Gleichwertigkeit zweier in der gleichen Sprache $(\boldsymbol{S}, \sigma)$ formulierter Axiomensysteme $\boldsymbol{X}, \boldsymbol{Y}$ (die sich ja einfach durch $Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X}) = Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{Y})$ ausdrücken ließ).

Definition 3. Zwei Sprachen $(\boldsymbol{S}_1, \sigma_1)$ und $(\boldsymbol{S}_2, \sigma_2)$ heißen *verträglich*, wenn gilt:

a) Es existiert die kleinste gemeinsame Erweiterung $\boldsymbol{S}_1\boldsymbol{S}_2$ der formalisierten Sprachen $\boldsymbol{S}_1, \boldsymbol{S}_2$, d. h. eine Sprache $\boldsymbol{S} \supset \boldsymbol{S}_1 \cup \boldsymbol{S}_2$, die genau die in $\boldsymbol{S}_1$ oder $\boldsymbol{S}_2$ vorkommenden Variablensorten, Relations-, Operations- und Konstantensymbole enthält.

b) Die Klasse $\sigma_1\sigma_2$ aller Interpretationen von $\boldsymbol{S}_1\boldsymbol{S}_2$, deren Einschränkung auf $\boldsymbol{S}_1$ zu σ_1 und deren Einschränkung auf $\boldsymbol{S}_2$ zu σ_2 gehört, ist nicht leer.

Kanonische elementare Sprachen sind stets verträglich, und ihre kleinste gemeinsame Erweiterung ist ebenfalls elementar. Für modifizierte formalisierte Sprachen ist jedoch im allgemeinen nicht einmal a) erfüllt: Kommt etwa das einstellige Relationssymbol R in $\boldsymbol{S}_1$ mit Variablen x_i und in $\boldsymbol{S}_2$ mit Variablen y_i vor, so müßte man, um zu einer gemeinsamen Erweiterung $\boldsymbol{S}_1\boldsymbol{S}_2$ zu gelangen, entweder $\mathrm{R}(\mathrm{x}_i)$ und $\mathrm{R}(\mathrm{y}_j)$ als prädikative Ausdrücke zur Darstellung zweier verschiedener Relationen ansehen oder die Variablensorten x_i und y_i identifizieren.

Definition 4. Es seien $\boldsymbol{T}_1 = Fl_{\boldsymbol{S}_1}^{\sigma_1}(\boldsymbol{X}_1)$ und $\boldsymbol{T}_2 = Fl_{\boldsymbol{S}_2}^{\sigma_2}(\boldsymbol{X}_2)$ zwei in verträglichen Sprachen $(\boldsymbol{S}_1, \sigma_1)$ bzw. $(\boldsymbol{S}_2, \sigma_2)$ formulierte Theorien. Dann heißt $\boldsymbol{T}_1$ *semantisch einbettbar* in $\boldsymbol{T}_2$, wenn die gemeinsame Erweiterung $(\boldsymbol{S}, \sigma)$ der beiden Sprachen eine definitorische Erweiterung von $(\boldsymbol{S}_2, \sigma_2)$ ist und $\boldsymbol{X}_1 \subseteq Fl_{\boldsymbol{S}}^{\sigma}(\boldsymbol{X}_2)$ gilt, d. h., wenn die Grundbegriffe von $\boldsymbol{T}_1$, soweit sie nicht schon selbst in der Sprache von $\boldsymbol{T}_2$ vorkommen, in dieser Sprache so definierbar sind, daß in der entsprechenden definitorischen Erweiterung die Axiome von $\boldsymbol{T}_1$ aus den Axiomen von $\boldsymbol{T}_2$ (und den Definitionen der Grundbegriffe) folgen. Die Theorien $\boldsymbol{T}_1$ und $\boldsymbol{T}_2$ heißen *äquivalent*, wenn jede der beiden semantisch in die andere einbettbar ist.

Ein wichtiger Anwendungsfall dieser Begriffe wird im nächsten Abschnitt behandelt. Hier sei noch erwähnt

Satz 2. *Ist* $T_1 = Fl^{\sigma_1}_{S_1}(X_1)$ *semantisch einbettbar in* $T_2 = Fl^{\sigma_2}_{S_2}(X_2)$, *so ist* T_1 *relativ widerspruchsfrei bezüglich* T_2.

(Nachträglich stellt sich heraus, daß die Methode, eine Theorie semantisch in eine andere einzubetten, fast stets den Beweisen der relativen Widerspruchsfreiheit zugrunde liegt.)

Beweis. Nach Voraussetzung ist die gemeinsame Erweiterung (S, σ) von (S_1, σ_1) und (S_2, σ_2) eine definitorische Erweiterung von (S_2, σ_2). Ist also T_2 widerspruchsfrei und ω ein beliebiges σ_2-Modell von X_2, so läßt sich ω zu einer σ-Interpretation ω' von S so fortsetzen, daß ω' ein Modell von X_2 bleibt, mithin X_2 auch als in (S, σ) formuliertes Axiomensystem widerspruchsfrei bleibt. Wegen der Voraussetzung

$$X_1 \subseteqq Fl_S{}^\sigma(X_2)$$

ist ω' ein σ-Modell von X_1, und wegen der Verträglichkeit von σ_1 und σ_2 ist die Einschränkung ω'' von ω' auf S_1 ein σ_1-Modell von X_1, also ist X_1 σ_1-widerspruchsfrei.

4.7. Äquivalenz der Grundbegriffe Kongruenz und Bewegung für die ebene euklidische Geometrie

Dieser Abschnitt ist vor allem als ein komplexes Beispiel für die im vorigen Abschnitt behandelten Begriffe der definitorischen Spracherweiterung und der semantischen Einbettbarkeit von Theorien aufzufassen. Davon unabhängig ist sein Inhalt wichtig für die Grundlagen des modernen Geometrieunterrichts. Wir schließen mit diesem Abschnitt die in vielen Beispielen der vorangehenden Abschnitte verstreuten Bemerkungen zu den Grundlagen der Elementargeometrie ab.

Wir erinnern an die in 3.3., Beispiel 2, eingeführte modifizierte formalisierte Sprache S_{eukl} für die Formalisierung der ebenen euklidischen Geometrie nach dem Vorbild Hilberts (d. h. mittels der Grundbegriffe Punkt, Gerade, Inzidenz, Zwischen- und Kongruenzrelation). Eine beliebige Interpretation dieser Sprache wird durch eine Struktur der Form $(M_1, M_2; R_{inz}, R_{zwi}, R_{kong})$ gegeben (vgl. 3.2., Beispiel 2), deren Signatur gleich $(2; (1, 2), (1, 1, 1), (1, 1, 1, 1))$ ist. Eine solche Struktur heißt eine *euklidische Ebene*, wenn sie ein Modell des folgenden Axiomensystems ist, das aus drei Inzidenzaxiomen I1 bis I3, vier Anordnungsaxiomen 01 bis 04 und neun Kongruenzaxiomen K1 bis K9 besteht. (Zur Formulierung der sogenannten Maß- oder Stetigkeitsaxiome ist eine nichtelementare Spracherweiterung nötig. Diese Axiome bewirken letzten Endes die Kategorizität des gesamten Axiomensystems. Ein großer Teil der Elementargeometrie läßt sich jedoch unabhängig von diesen Axiomen aufbauen, und wir wollen sie hier nicht in den Begriff der euklidischen Ebene einbeziehen. Demzufolge gibt es zahlreiche paarweise nicht isomorphe Modelle unseres Axiomensystems, d. h. zahlreiche wesentlich verschiedene euklidische Ebenen.)

I1 $\bigvee$ ABC A, B, C nicht kollinear.[1])

I2 $P \neq Q \to \bigvee!!g(P \text{ auf } g \wedge Q \text{ auf } g)$.

I3 $\neg\, A \text{ auf } g \to \bigvee!!h(A \text{ auf } h \wedge h \| g)$.

01 $[A, B, C] \to A, B, C \text{ kollinear} \wedge A \neq B \wedge A \neq C \wedge B \neq C$.

02 $[A, B, C] \to [C, B, A] \wedge \neg\, [B, A, C]$.

03 $A \neq B \to \bigvee C[A, B, C]$.

04 $A, B, C \text{ nicht kollinear} \wedge \bigvee D(D \text{ auf } g \wedge [A, D, B]) \wedge \neg\, A \text{ auf } g \wedge \neg\, B \text{ auf } g \wedge \neg\, C \text{ auf } g \to \bigvee E\big(E \text{ auf } g \wedge ([A, E, C] \vee [B, E, C])\big)$.

K1 $AB \cong AB$.

K2 $AB \cong BA$.

K3 $AB \cong CD \to CD \cong AB$.

K4 $AB \cong CD \wedge CD \cong EF \to AB \cong EF$.

K5 $AB \cong CC \to A = B$.

K6 $A \neq B \wedge C \neq D \to \bigvee!!E([A, B, E] \wedge BE \cong CD)$.

K7 $A, B, C \text{ nicht kollinear} \wedge D, E, F \text{ nicht kollinear} \wedge AB \cong DE \to \bigvee!!G\big(AC \cong DG \wedge BC \cong EG \wedge \neg \bigvee H([G, H, F] \wedge D, E, H \text{ kollinear})\big)$.

K8 $[A, B, C] \wedge [D, E, F] \wedge AB \cong DE \wedge BC \cong EF \to AC \cong DF$.

K9 $[A, B, C] \wedge [E, F, G] \wedge A, B, D \text{ nicht kollinear} \wedge AB \cong EF \wedge BC \cong FG \wedge AD \cong EH \wedge BD \cong FH \to CD \cong GH$.

Will man den Grundbegriff (Strecken-) Kongruenz im Rahmen einer elementaren Theorie durch den Grundbegriff Bewegung ersetzen, so sind die euklidischen Ebenen als Strukturen der Form

$$(M_1, M_2, M_3; R_{inz}, R_{zwi}; F)$$

aufzufassen, wobei M_3 eine zu M_1 und M_2 disjunkte Menge von Dingen einer dritten Sorte (*Bewegungen* genannt) und F eine Abbildung von $M_3 \times M_1$ in M_1 ist (die je einer Bewegung φ und einem Punkt P den Punkt $\varphi(P)$ zuordnet). Derartige Strukturen werden wir im folgenden als *Bewegungsebenen* bezeichnen. Eine passende Sprache S_{bew} zur Beschreibung dieser Strukturen entsteht aus S_{eukl}, indem man das Relationssymbol $\cong$ aus der Basis entfernt, statt dessen Variablen φ, φ_i einer dritten Sorte für Bewegungen aufnimmt und die Spracherzeugungsregeln durch folgende Termbildungsvorschrift ergänzt:

---

[1]) Hier und im folgenden werden abkürzungshalber die Definitionen (5), (6), (7) aus 4.6. benutzt.

Ist t_1 *ein Term der Sorte Bewegung und* t_2 *ein Term der Sorte Punkt, so ist* $t_1(t_2)$ *ein Term der Sorte Punkt.*

In der so erhaltenen Sprache S_{bew} kann man die (im wesentlichen von F. SCHUR aufgestellten) Bewegungsaxiome etwa wie folgt formulieren:

B1 $\quad \bigwedge \varphi_1\varphi_2 \bigvee \varphi_3 \bigwedge \mathrm{P}\ \varphi_2(\varphi_1(\mathrm{P})) = \varphi_3(\mathrm{P})$,

(d. h., die Hintereinanderausführung zweier Bewegungen ist eine Bewegung).

B2 $\quad \bigwedge \varphi_1 \bigvee \varphi_2 \bigwedge \mathrm{P}\ \varphi_2(\varphi_1(\mathrm{P})) = \mathrm{P}$,

(d. h., zu jeder Bewegung gibt es eine inverse Bewegung).

B3 $\quad [\mathrm{A}, \mathrm{B}, \mathrm{C}] \leftrightarrow [\varphi(\mathrm{A}), \varphi(\mathrm{B}), \varphi(\mathrm{C})]$,

(d. h., bei einer beliebigen Bewegung stehen drei Punkte genau dann in der Zwischenrelation, wenn ihre Bildpunkte in der gleichen Relation stehen. Allgemein bezeichnet man Abbildungen φ mit der Eigenschaft B3 als *affin*).

B4 $\quad$ A, B, C nicht kollinear $\wedge$ D, E, F nicht kollinear $\rightarrow \bigvee !!\varphi(\varphi(\mathrm{A}) = \mathrm{D}$ $\wedge ([\mathrm{D}, \varphi(\mathrm{B}), \mathrm{E}] \vee \varphi(\mathrm{B}) = \mathrm{E} \vee [\mathrm{D}, \mathrm{E}, \varphi(\mathrm{B})]) \wedge \neg \bigvee \mathrm{P}([\varphi(\mathrm{C}), \mathrm{P}, \mathrm{F}] \wedge$ D, E, P kollinear$))$,

(d. h., eine Bewegung ist jeweils eindeutig bestimmt, wenn man für drei nicht kollineare Punkte den Bildpunkt des ersten, von dort ausgehend den Strahl, auf dem der Bildpunkt des zweiten Punktes liegen soll, und für den dritten die Halbebene des Bildpunktes willkürlich vorgibt).

B5 $\quad \bigwedge \mathrm{PQ} \bigvee \varphi(\varphi(\mathrm{P}) = \mathrm{Q} \wedge \varphi(\mathrm{Q}) = \mathrm{P})$.

B6 $\quad$ A, B, C nicht kollinear $\wedge \bigvee \varphi(\varphi(\mathrm{A}) = \mathrm{A} \wedge \varphi(\mathrm{B}) = \mathrm{C})$ $\rightarrow \bigvee \varphi(\varphi(\mathrm{A}) = \mathrm{A} \wedge \varphi(\mathrm{B}) = \mathrm{C} \wedge \varphi(\mathrm{C}) = \mathrm{B})$.

Um die Bewegungen im Rahmen einer elementaren Sprache in der von SCHUR beabsichtigten Weise zu beschreiben, muß nun noch die üblicherweise vor der Formulierung der Axiome B1 bis B6 mitgeteilte Interpretationsvorschrift, daß die Dinge der Sorte φ Abbildungen des Bereichs der Punkte auf sich sein sollen, durch weitere Axiome ersetzt werden (d. h., die in den üblichen Darstellungen benutzte nichtelementare Sprache ist unwesentlich nichtelementar; vgl. 4.5.).

B7 $\quad \bigwedge \varphi_1\varphi_2(\bigwedge \mathrm{P}\ \varphi_1(\mathrm{P}) = \varphi_2(\mathrm{P}) \rightarrow \varphi_1 = \varphi_2)$.

B8 $\quad \bigwedge \varphi\mathrm{P} \bigvee \mathrm{Q}\ \varphi(\mathrm{P}) = \mathrm{Q}$.

B9 $\quad \bigwedge \varphi\mathrm{P} \bigvee \mathrm{Q}\ \varphi(\mathrm{Q}) = \mathrm{P}$.

Die Axiome B7 bis B9 sichern, daß die Objekte, durch die die Variablen φ interpretiert werden, bis auf Isomorphie gewisse Abbildungen des Bereichs der Punkte auf sich sind. Die Axiome B4 bis B6 sichern, daß alle Bewegungen im inhaltlichen Sinn zu den mit φ bezeichneten Objekten gehören. Die Axiome B1 bis B3 bedeuten zusammen, daß die Bewegungen eine Untergruppe der Gruppe der affinen Abbildungen bilden.

In der Sprache S_{bew} ist die Kongruenz nach dem gewöhnlichen Verfahren der Relationsdefinition definierbar:

D1 $\quad \mathrm{AB} \cong \mathrm{CD} :\leftrightarrow \bigvee \varphi\big(\varphi(\mathrm{A}) = \mathrm{C} \wedge \varphi(\mathrm{B}) = \mathrm{D}\big).$

Für die so definierte Kongruenz folgen aus dem Axiomensystem

$$\boldsymbol{X} = \{\mathrm{I}\,1, \mathrm{I}\,2, \mathrm{I}\,3, 01, 02, 03, 04, \mathrm{B}\,1, \ldots, \mathrm{B}\,9\}$$

der Bewegungsebenen die Kongruenzaxiome K 1, ..., K 9. (Die Beweise gehören der Elementargeometrie an und werden hier nicht durchgeführt.) Bezeichnet $\boldsymbol{Y}$ das System der Axiome der euklidischen Ebenen, so ist demnach die euklidische ebene Geometrie, aufgefaßt als $Fl_{S_{eukl}}(\boldsymbol{Y})$, semantisch einbettbar in die Theorie $Fl_{S_{bew}}(\boldsymbol{X})$ der Bewegungsebenen, d. h., es gilt

$$\boldsymbol{Y} \subseteqq Fl_{S}(\boldsymbol{X} \cup \{\mathrm{E}\}),$$

wobei S die durch D 1 aus S_{bew} entstehende Erweiterungssprache und E das D 1 entsprechende Einführungsaxiom ist. Wir zeigen nun, daß „im wesentlichen" auch die Umkehrung dieses Sachverhalts richtig ist.

Ausgehend von der Sprache S_{eukl} und der in dieser Sprache formulierten Theorie $Fl_{S_{eukl}}(\boldsymbol{Y})$ möchte man, inhaltlichen Erwägungen folgend, etwa „definieren":

Im folgenden sollen die neuen Variablen φ, φ_i (bis auf Isomorphie) Abbildungen des Bereichs der Punkte auf sich sein, für die

$$\varphi(\mathrm{A})\,\varphi(\mathrm{B}) \cong \mathrm{AB} \tag{1}$$

gilt, und $\varphi(\mathrm{P})$ soll den Bildpunkt von P bei der Abbildung φ bezeichnen.

Dies bedeutet zunächst den Übergang zu einer Erweiterungssprache (wiederum S, der kleinsten gemeinsamen Erweiterung von S_{eukl} und S_{bew}). Damit diese Erweiterung definitorisch im Sinne von 4.6., Definition 2, sein kann, muß zunächst die dortige Bedingung (a) erfüllt sein, daß sich jede Interpretation von S_{eukl} zu einer Interpretation von S fortsetzen läßt. Daß es aber überhaupt Abbildungen φ des Punktbereichs auf sich mit der Eigenschaft (1) gibt, hängt wesentlich von den bei der unformalen „Definition" vorausgesetzten Axiomen $\boldsymbol{Y}$ ab. (Aus der Sicht dieses Falles und vieler anderer scheint es manchmal zweckmäßig zu sein, neben dem Begriff der definitorischen Spracherweiterung einen Begriff der definitorischen Erweiterung von Theorien zur Verfügung zu haben.) Man kann sich in diesem Fall (und in allen Fällen, in denen eine Definition, um sinnvoll zu sein, die Gültigkeit eines gewissen Axiomensystems $\boldsymbol{Y}$ voraussetzt) durch folgenden Kunstgriff helfen: Es sei σ die Klasse aller Interpretationen von S_{eukl}, die Modelle von $\boldsymbol{Y}$ sind. Statt der Sprache S_{eukl} wird nun die (unwesentlich nichtelementare) Sprache (S_{eukl}, σ) definitorisch erweitert. Damit ist die Bedingung (a) sofort erfüllt. Um (b) zu erfüllen, kann man als Einführungsaxiome für die definierten Bewegungen z. B. das aus (1), B7, B8, B9 und B4 bestehende System E' nehmen, wobei die Rolle von B4 darin besteht, zu sichern, daß unter den definitorisch eingeführten Bewegungen alle Bewegungen im inhaltlichen Sinn vorkommen. Wesentlich interessanter für die geometrische Praxis ist jedoch das Verfahren der Rückübersetzung von Ausdrücken $H \in S$, in denen höchstens quantifizierte Bewegungsvariablen vorkommen, in semantisch äquivalente Ausdrücke [H] der Sprache S_{eukl}, mit anderen Worten: die Eliminierbarkeit quantifizierter Bewegungsvariablen mittels des Kongruenzbegriffes. Eine Bewegungsvariable φ kann in Termen der Sorte Punkt (im folgenden mit t_φ bezeichnet) und in Termgleichungen $\varphi = \varphi_i$ (bzw. $\varphi_i = \varphi$) vorkommen. Wir geben zunächst drei vorbereitende Reduktionsschritte an.

1. Schritt. Vorkommen von φ in der Form $\varphi = \varphi_i$ werden auf Vorkommen von φ in Punkttermen reduziert, indem man unter Benutzung einer im gesamten betrachteten Ausdruck nicht vorkommenden Punktvariablen P den Ausdruck $\varphi = \varphi_i$ semantisch äquivalent durch $\bigwedge P\ \varphi(P) = \varphi_i(P)$ ersetzt.

2. Schritt. Vorkommen von φ in Punkttermen t_φ werden auf solche der Form $t_\varphi = P$ reduziert, indem man (wieder unter Benutzung einer noch nicht vorkommenden Punktvariablen P) $H(t_\varphi)$ semantisch äquivalent durch $\bigvee P\big(t_\varphi = P \wedge H(P)\big)$ ersetzt.

3. Schritt. Vorkommen der Form $\varphi_1\big(\ldots \varphi_k\big(\varphi(t)\big) \ldots\big) = P$ werden auf solche der Form $\varphi(t) = P$ reduziert, wobei φ in t nicht mehr vorkommt. Ist φ ein „innerstes" Vorkommen im Term $\varphi_1\big(\ldots \varphi_k\big(\varphi(t)\big) \ldots\big)$, so wird die entsprechende Termgleichung unter Benutzung einer noch nicht vorkommenden Punktvariablen Q semantisch äquivalent durch $\bigvee Q\big(\varphi_1\big(\ldots \varphi_k(Q) \ldots\big) = P \wedge \varphi(t) = Q\big)$ ersetzt.

Bei der eigentlichen Elimination gehen wir von dem (aus Y und den Einführungsaxiomen E' folgenden) Sachverhalt aus, daß eine Bewegung durch drei beliebige, nicht kollineare Punkte A, B, C und deren Bildpunkte A', B', C' in der Weise bestimmt ist, daß dann für jeden Punkt P die Strecken PA, PB, PC gleich den Strecken P'A', P'B', P'C' sind (P' Bildpunkt von P). Es sei $H(\varphi) \in S$ ein Ausdruck, der die Variable φ vollfrei enthält, ferner seien A, B, C, A', B', C' in $H(\varphi)$ nicht vorkommende Punktvariablen. Dann ist $\bigvee \varphi H(\varphi)$ semantisch äquivalent zu

$$\bigvee ABCA'B'C'(A, B, C \text{ nicht kollinear} \wedge AB \cong A'B' \wedge AC \cong A'C' \wedge BC \cong B'C' \wedge H^*),$$

wobei H^* aus $H(\varphi)$ entsteht, indem man jeden prädikativen Teilausdruck der Form $\varphi(t) = P$ (φ nicht in t) durch den Ausdruck $tA \cong PA' \wedge tB \cong PB' \wedge tC \cong PC'$ ersetzt. Auf Grund der Reduktionsschritte 1, 2, 3 können wir annehmen, daß es in $H(\varphi)$ keine anderen Vorkommen der Variablen φ gibt. Den Fall $\bigwedge \varphi H(\varphi)$ kann man analog behandeln oder durch Umformung in $\neg \bigvee \varphi \neg H(\varphi)$ auf den behandelten Fall zurückführen.

Damit haben wir erhalten: $\boldsymbol{S}$ mit der durch die Axiome $\boldsymbol{Y}$ und $\boldsymbol{E}'$ gegebenen Interpretationsvorschrift ist eine definitorische Erweiterung zweiter Art von $(\boldsymbol{S}_{eukl}, \sigma)$. Es liegt nun wiederum im Rahmen der Elementargeometrie, die semantische Einbettbarkeit der Theorie $Fl_{\boldsymbol{S}_{bew}}(\boldsymbol{X})$ der Bewegungsebenen in die ebene euklidische Geometrie $Fl_{\boldsymbol{S}_{eukl}}(\boldsymbol{Y})$ zu zeigen, d. h. $\boldsymbol{X} \subseteqq Fl_{\boldsymbol{S}}(\boldsymbol{Y} \cup \boldsymbol{E}')$. Wegen I1, I2, I3, 01, 02, 03, 04 $\in \boldsymbol{Y}$ und B4, B7, B8, B9 $\in \boldsymbol{E}'$ ist dabei in Wahrheit nur nachzuweisen, daß B1, B2, B3, B5, B6 $\in Fl_{\boldsymbol{S}}(\boldsymbol{Y} \cup \boldsymbol{E}')$ gilt.

5. Eine Formalisierung der Mengenlehre

Es gibt gegenwärtig mehrere axiomatische Systeme der Mengenlehre, die sich zum Teil nicht nur in mehr oder weniger formalen Einzelheiten, sondern wesentlich in den ihnen zugrunde liegenden inhaltlichen Vorstellungen unterscheiden. Dieses Kapitel führt in eines der verbreitetsten dieser Systeme ein, das im wesentlichen von J. v. NEUMANN, P. BERNAYS und K. GÖDEL (siehe hierzu [5, 9]) entwickelt wurde. Die hier behandelte Variante unterscheidet sich von der ursprünglichen Form vor allem durch die Einbeziehung von sogenannten Urelementen in die Betrachtung. Das Anliegen des Kapitels besteht einerseits in der Vermittlung von Grundkenntnissen über axiomatische Mengenlehre, andererseits in einer komplexen Wiederholung und Anwendung der in den vorangehenden Kapiteln behandelten Formalisierungstechniken.

Der Wahl einer geeigneten modifizierten formalisierten Sprache sind (wie stets) einige inhaltliche Überlegungen vorauszuschicken. Wie bekannt, wünscht man in der Mengenlehre nach Möglichkeit mit der Elementrelation und eventuell Identität als einzigen Grundbegriffen auszukommen. Ist H(x) eine beliebige, mittels dieser Begriffe formulierbare Eigenschaft von Dingen x (d. h. ein Ausdruck der entsprechenden Sprache mit der vollfreien Variablen x), so soll die Menge aller Dinge x mit der Eigenschaft H(x) existieren. Bekanntlich läßt sich das von G. CANTOR, dem Begründer der Mengenlehre, naiv formulierte Mengenbildungsprinzip nicht in dieser Allgemeinheit widerspruchsfrei verwirklichen. Zum Beispiel folgt aus der Annahme der Existenz der Menge $M = \{x: \neg\, x \in x\}$ die Antinomie von RUSSELL: Wenn $M \in M$ gilt, so ist $\neg\, M \in M$; wenn $\neg M \in M$, so ist $M \in M$. J. v. NEUMANN bemerkte als erster[1]), daß diese Antinomie und andere sich noch nicht aus der

[1]) Gewisse Briefe von CANTOR an DEDEKIND deuten allerdings darauf hin, daß CANTOR selbst schon vor 1900 den Begriff der Unmenge ins Auge gefaßt hatte. Vgl. hierzu O. BECKER, Grundlagen der Mathematik, Verlag Karl Alber, Freiburg—München 1954, S. 308—309.

Annahme der Existenz einer gewissen Gesamtheit (z. B. $\{x: \neg\, x \in x\}$) an sich ergeben, sondern erst, wenn man annimmt, daß für diese Gesamtheit ihrerseits die Frage sinnvoll ist, ob sie Element gewisser Mengen ist oder nicht. Er ergänzte daher das System aller Mengen durch die Annahme gewisser Unmengen, die zwar — wie die Mengen — Gesamtheiten von Dingen sind, aber im Gegensatz zu den Mengen selbst nicht Element anderer Gesamtheiten sein können.

Ein besonderer Zug des axiomatischen Systems von v. NEUMANN, BERNAYS und GÖDEL wie auch anderer axiomatischer Systeme (z. B. des Systems von ZERMELO-FRAENKEL) besteht darin, daß eine Fülle von Mengen, einschließlich der verschiedenen Zahlbereiche, sozusagen aus dem Nichts erschaffen wird: Betrachtet man die Elemente der Elemente der ... der Elemente der Elemente beliebiger Mengen, so stößt man letzten Endes immer auf die leere Menge als einzigen „Baustein". Natürlich ist es einerseits interessant, daß die gedankliche Konstruktion aller mathematischen Begriffe auf einer so schwachen Grundlage grundsätzlich möglich ist, und von vielen Mathematikern wurde ein solcher Aufbau der Mengenlehre sogar als besonders „rein" bzw. befriedigend gerühmt, aber als Materialisten erwarten wir von einem axiomatischen System der Mengenlehre unter anderem gerade die Modellierung der Beziehungen zwischen mathematischen Begriffen einerseits und Dingen der objektiven Realität andererseits. Ein uns befriedigendes System der Mengenlehre soll nicht nur die tatsächliche historische Entstehung der mathematischen Begriffe aus dem Gegenständlichen heraus im wesentlichen richtig widerspiegeln, sondern auch die theoretische Begründung für die Anwendungen der Mathematik in der Praxis liefern.

Diesen Überlegungen entsprechend, gehen wir bei der folgenden Axiomatisierung der Mengenlehre davon aus, daß der Bereich aller betrachteten Dinge Objekte von dreierlei Art umfaßt:

(a) *Urelemente,* d. h. Dinge, die zwar als Elemente von Gesamtheiten auftreten können, aber selbst keine Elemente enthalten,

(b) *Mengen,* d. h. Dinge, die sowohl Gesamtheiten als auch Elemente gewisser anderer Gesamtheiten sind,

(c) *Unmengen,* d. h. Gesamtheiten, die nicht als Elemente anderer Gesamtheiten auftreten können.

Urelemente und Mengen bezeichnen wir mit dem gemeinsamen Namen *Elemente,* Mengen und Unmengen mit dem gemeinsamen Namen *Klassen.*

Wenn man nicht auf die Existenz einer leeren Menge verzichten will, verursacht die Zulassung von Urelementen die Schwierigkeit, daß die charakteristische Eigenschaft der leeren Menge, kein Element zu besitzen, auch auf alle Urelemente zutrifft, d. h., ohne einen zusätzlichen Grundbegriff läßt sich die leere Menge nicht von den Urelementen unterscheiden. Anders gesagt: Unter allen Dingen, die kein Element besitzen, ist genau eines als eine Menge auszuzeichnen. Vom Standpunkt der formalisierten Sprachen ist das am einfachsten durch die Aufnahme eines Individuensymbols für die leere Menge möglich.

Wir benutzen demnach folgende modifizierte einsortige Sprache $\boldsymbol{S}_m$: Die Konstante $\emptyset$ (zur Bezeichnung der leeren Menge) und kleine lateinische Variablen (eventuell mit Indizes) für Dinge sind die Terme dieser Sprache. Prädikative Ausdrücke haben die Form $x = y$ oder $x \in y$ (wobei x, y Variable oder gleich $\emptyset$ sind). Im folgenden notieren wir (manchmal mit Erläuterungen) in zweckmäßiger Reihenfolge formalisierte Definitionen (bezeichnet mit D1, D2, D3, ...), Axiome (bezeichnet mit A1, A2, A3, ...) und Folgerungen aus den Axiomen und Definitionen (bezeichnet mit F1, F2, F3, ...). Wir erinnern nochmals daran, daß die Notierung einer formalisierten Definition den Übergang zu der durch die definierten Sprachbestandteile erweiterten Sprache bei gleichzeitiger Erweiterung des bisher benutzten Axiomensystems durch die entsprechenden Einführungsaxiome bedeutet und daß alle mit Hilfe der definierten Begriffe formulierten Sätze (Axiome und Folgerungen) prinzipiell in die anfangs benutzte Sprache $\boldsymbol{S}_m$ rückübersetzt werden können.

Zunächst verwandeln wir unsere oben diskutierten inhaltlichen Vorstellungen über die verschiedenen Arten von Dingen in formalisierte Definitionen:

D1 $\quad x$ Element $:\leftrightarrow \bigvee y\ x \in y$.

D2 $\quad x$ Klasse $:\leftrightarrow x = \emptyset \vee \bigvee y\ y \in x$.

D3 $\quad x$ Menge $:\leftrightarrow x$ Klasse $\wedge$ x Element.

D4 $\quad x$ Unmenge $:\leftrightarrow \neg\, x$ Element.

D5 $\quad x$ Urelement $:\leftrightarrow \neg\, x$ Klasse.

Wie man durch entsprechende Modellbetrachtungen sofort sieht, folgt aus D1 bis D5 noch nicht, daß jedes Urelement ein Element ist, daß jede Unmenge eine Klasse ist, daß die leere Menge kein Element besitzt und eine Menge ist und daß es überhaupt Dinge aller definierten Arten gibt.

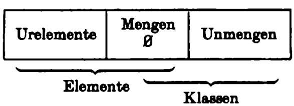
A1 $\quad \emptyset$ Element.

F1 $\quad \emptyset$ Menge.

A2 $\quad u$ Urelement $\rightarrow u$ Element.

F2 $\quad x$ Unmenge $\rightarrow x$ Klasse.

(Eine Unmenge, die keine Klasse ist, wäre nach D5 ein Urelement, folglich nach A2 ein Element, im Widerspruch zu D4. Somit folgt F2.) Der Bereich aller Dinge ist demnach wie folgt aufgeteilt:



A 3 (Axiom der leeren Menge) $\neg \bigvee x\, x \in \emptyset$.

A 4 (Extensionalitätsaxiom)
x Klasse $\wedge$ y Klasse $\wedge \bigwedge z(z \in x \leftrightarrow z \in y) \rightarrow x = y$.

(Wollten wir das Extensionalitätsaxiom für beliebige Dinge x, y statt nur für Klassen formulieren, so hätte dies sofort zur Folge, daß es keine Urelemente gibt, da dann jedes Urelement gleich $\emptyset$, andererseits wegen D 2 und D 5 ungleich $\emptyset$ sein müßte.)

F 3 $\neg \bigvee x\, x \in y \leftrightarrow y$ Urelement $\vee\, y = \emptyset$.

F 4 *Ist* $H(x) \in S_m$ *ein beliebiger Ausdruck, der* x *vollfrei enthält, so folgt aus einer Aussage der Form*

$\bigvee z\big(z \text{ Klasse} \wedge \bigwedge x\big(x \in z \leftrightarrow H(x)\big)\big)$

stets sofort die Verschärfung

$\bigvee !!z\big(z \text{ Klasse} \wedge \bigwedge x\big(x \in z \leftrightarrow H(x)\big)\big)$.

D 6 $x \notin y :\leftrightarrow \neg\, x \in y$.

D 7 $x \neq y :\leftrightarrow \neg\, x = y$.

D 8 $x \subseteq y :\leftrightarrow x$ Klasse $\wedge$ y Klasse $\wedge \bigwedge z(z \in x \rightarrow z \in y)$.

(Würden wir D 8 nicht auf Klassen x, y beschränken, so würde sich sofort $u \subseteq x$ für beliebige Urelemente u und Dinge x ergeben.)

D 9 $x \subset y :\leftrightarrow x \subseteq y \wedge x \neq y$.

F 5 x Klasse $\wedge$ y Klasse $\rightarrow (x = y \leftrightarrow x \subseteq y \wedge y \subseteq x)$.

A 5 (Paarmengenaxiom)
x Element $\wedge$ y Element $\rightarrow \bigvee z\big(z \text{ Menge} \wedge \bigwedge a(a \in z \leftrightarrow a = x \vee a = y)\big)$.

Unter Berücksichtigung von F 4 verschärft sich A 5 sofort zu

F 6 x Element $\wedge$ y Element $\rightarrow \bigvee !!z\big(z \text{ Menge} \wedge \bigwedge a(a \in z \leftrightarrow a = x \vee a = y)\big)$.

Unter Berufung auf F 6 definieren wir

D 10 $\{x, y\} := \iota z\big(z \text{ Menge} \wedge \bigwedge a(a \in z \leftrightarrow a = x \vee a = y)\big)$,
falls x Element $\wedge$ y Element.

D 11 $\{x\} := \{x, x\}$,

D 12 $(x, y) := \big\{\{x\}, \{x, y\}\big\}$.

Die so definierten Operationen Paarmenge, Einermenge und geordnetes Paar sind jeweils für Elemente definiert und liefern als Resultat Mengen.

F 7 $(x_1, y_1) = (x_2, y_2) \leftrightarrow x_1 = x_2 \wedge y_1 = y_2$.

Beweis. Offenbar ist nur „$\rightarrow$" zu zeigen. Ist $x_1 = y_1$, so ist $(x_1, y_1) = \{\{x_1\}\}$. Folglich kann auch $(x_2, y_2) = \big\{\{x_2\}, \{x_2, y_2\}\big\}$ nur ein Element enthalten, d. h., es muß auch $x_2 = y_2$ sein und dann wegen $\{\{x_1\}\} = \{\{x_2\}\}$ schließlich $x_1 = x_2$. Ist $x_1 \neq y_1$, so ist $\{x_1\} \neq \{x_1, y_1\}$, d. h., $(x_1, y_1) = \big\{\{x_1\}, \{x_1, y_1\}\big\}$ enthält zwei verschiedene Elemente. Dann muß auch (x_2, y_2) aus zwei Elementen bestehen, d. h., es muß $x_2 \neq y_2$ sein. Jetzt muß jedes der beiden Elemente $\{x_1\}, \{x_1, y_1\}$ gleich einem der beiden Elemente $\{x_2\}, \{x_2, y_2\}$ sein, was sich nur mit $x_1 = x_2$ und $y_1 = y_2$ vereinbaren läßt.

Aus D12 gewinnt man für jede natürliche Zahl $n \geqq 3$ den Begriff des n-Tupels von Elementen:

D13 $(x, y, z) := \big((x, y), z\big)$.

D14 $(x, y, z, w) := \big((x, y, z), w\big)$.

$\vdots$

Wir bemerken, daß die naheliegende Zusammenfassung unendlich vieler korrekter Einzeldefinitionen zur induktiven Definition

$$(x_1, \ldots, x_n, x_{n+1}) := \big((x_1, \ldots, x_n), x_{n+1}\big)$$

sich innerhalb der Sprache S_m oder einer elementaren Erweiterung nicht korrekt formulieren läßt (vgl. S. 81 f.).

Unter Benutzung der induktiven Definition kann man jedoch sehr leicht induktiv beweisen, wobei F7 Anfangs- und Induktionsschritt liefert:

M1 *Für jede natürliche Zahl $n \geqq 2$ gilt*

$$(x_1, \ldots, x_n) = (y_1, \ldots, y_n) \leftrightarrow x_1 = y_1 \wedge \ldots \wedge x_n = y_n. \qquad (1)$$

Natürlich ist M1 genausowenig wie die zugrunde liegende Definition in der Sprache S_m formulierbar, sondern eine Zusammenfassung von unendlich vielen in S_m formulierbaren Einzelsätzen. Sätze dieser Art werden wir im folgenden als *Metasätze* bezeichnen und mit M1, M2, M3, ... numerieren. Nachträglich stellen wir fest, daß F4 ebenfalls ein Metasatz ist. Für das Folgende vereinbaren wir

D15 $(x) := x$, *falls* x Element.

Dann gilt (1) auch für $n = 1$. Wir formulieren nun einige Klassenbildungsaxiome.

A6 $\bigvee a \bigwedge x\big(x \in a \leftrightarrow \bigvee yz(x = (y, z) \wedge y \in z)\big)$.

D16 $E := \iota a \bigwedge x\big(x \in a \leftrightarrow \bigvee yz(x = (y, z) \wedge y \in z)\big)$.

A7 $\bigvee a \bigwedge x\big(x \in a \leftrightarrow \bigvee y\, x = (y, y)\big)$.

D17 $I := \iota a \bigwedge x\big(x \in a \leftrightarrow \bigvee y\, x = (y, y)\big)$.

Die Axiome A6 bzw. A7 bedeuten: Die auf Elemente eingeschränkten Relationen des Elementseins und der Identität existieren als (wegen A4 eindeutig bestimmte) Klassen von geordneten Paaren. Die Symbole E bzw. I sind demnach definitorisch eingeführte Konstanten zur Bezeichnung spezieller Klassen.

A8 $\bigwedge b \bigvee a \bigwedge x(x \in a \leftrightarrow x \text{ Element} \wedge x \notin b)$.

1D8 $\bar{b} := \iota a \bigwedge x(x \in a \leftrightarrow x \text{ Element} \wedge x \notin b))$, *falls* b Klasse.

A9 $\bigwedge bc \bigvee a \bigwedge x(x \in a \leftrightarrow x \in b \wedge x \in c)$.

D19 $b \cap c := \iota a \bigwedge x(x \in a \leftrightarrow x \in b \wedge x \in c)$, *falls* b Klasse $\wedge$ c Klasse.

(Es ist nicht nötig, in der Formulierung von A8 und A9 b bzw. b und c auf Klassen einzuschränken, da die entsprechende Aussage für Urelemente durch die axiomatische Forderung einer leeren Menge jeweils automatisch erfüllt ist. Die zugehörigen Operationen der Komplementbildung und des Durchschnitts sollen jedoch nur für Klassen definiert sein.)

D20 $\quad b \setminus c := b \cap \bar{c}, \quad b \cup c := \overline{\bar{b} \cap \bar{c}}.$

D21 $\quad A := \bar{\emptyset}.$

(D21 ist ein Beispiel für die Definition einer Konstanten durch einen variablenfreien Term; vgl. S. 85. Die so definierte Klasse A aller Elemente wird als *Allklasse* bezeichnet.)

F8 $\quad x$ Element $\leftrightarrow x \in A.$

F9 $\quad x$ Klasse $\leftrightarrow x \subseteqq A.$

F10 $\quad x$ Unmenge $\leftrightarrow x \notin A.$

A10 $\quad \bigwedge bc \bigvee a \bigwedge x\big(x \in a \leftrightarrow \bigvee yz(x = (y, z) \wedge y \in b \wedge z \in c)\big).$

D22 $\quad b \times c := \iota a \bigwedge x\big(x \in a \leftrightarrow \bigvee yz(x = (y, z) \wedge y \in b \wedge z \in c)\big),$
falls b Klasse $\wedge$ c Klasse.

Metadefinitionen:

$$b_1 \times b_2 \times \cdots \times b_n := \big(\ldots (b_1 \times b_2) \times \cdots \times b_n\big),$$
$$b^0 := \emptyset, \quad b^1 := b, \quad b^{n+1} := b^n \times b.$$

Vorsicht! Es ist im allgemeinen $a \times (b \times c) \neq a \times b \times c$ usw., insbesondere ist $b^n \times b^m \neq b^{n+m}$.

A11 $\quad \bigwedge b \bigvee a \bigwedge x\big(x \in a \leftrightarrow \bigvee yz(x = (y, z) \wedge (z, y) \in b)\big).$

D23 $\quad b^{-1} := \iota a \bigwedge x\big(x \in a \leftrightarrow \bigvee yz(x = (y, z) \wedge (z, y) \in b)\big),$
falls b Klasse.

(b^{-1} soll von jeder Klasse b gebildet werden können. Enthält insbesondere b keine geordneten Paare, so ist $b^{-1} = \emptyset$. Allgemein ist $(b^{-1})^{-1} \subseteqq b$ und gerade die Klasse aller in b enthaltenen geordneten Paare.)

A12 $\quad \bigwedge b \bigvee a \bigwedge x\big(x \in a \leftrightarrow \bigvee yzw(x = (y, z, w) \wedge (w, y, z) \in b)\big).$

D24 $\quad b^{-2} := \iota a \bigwedge x\big(x \in a \leftrightarrow \bigvee yzw(x = (y, z, w) \wedge (w, y, z) \in b)\big),$
falls b Klasse.

A13 $\quad \bigwedge b \bigvee a \bigwedge x\big(x \in a \leftrightarrow \bigvee y(x, y) \in b\big).$

D25 $\quad V(b) := \iota a \bigwedge x(x \in a \leftrightarrow \bigvee y(x, y) \in b)$, *falls* b Klasse.

D26 $\quad N(b) := V(b^{-1}).$

Die Klassen V(b) bzw. N(b) werden als *Vorbereich* bzw. *Nachbereich* von b bezeichnet.

F11 $\quad z \in N(b) \leftrightarrow \bigvee y(y, z) \in b.$

Ein Ausdruck $H \in \mathbf{S}_m$ heiße *primitiv*, wenn alle in ihm vorkommenden Quantifizierungen auf Elemente relativiert sind, d. h., wenn in jedem Teilausdruck der Form

$\bigvee \mathrm{x}H'(\mathrm{x})$ der Teilausdruck $H'(\mathrm{x})$ die Form $(\mathrm{x}\ \text{Element} \wedge H''(\mathrm{x}))$ und in jedem Teilausdruck der Form $\bigwedge \mathrm{x}H'(\mathrm{x})$ der Teilausdruck $H'(\mathrm{x})$ die Form $(\mathrm{x}\ \text{Element} \rightarrow H''(\mathrm{x}))$ hat. Wir bemerken, daß man zu jedem primitiven Ausdruck einen semantisch äquivalenten primitiven Ausdruck finden kann, in dem die logischen Funktoren $\vee$, $\rightarrow$, $\leftrightarrow$ und $\bigwedge$ nicht vorkommen.

M2 (Satz von Gödel) *Ist* $n \geqq 1$, *sind* $\mathrm{x}_1, \ldots, \mathrm{x}_n$ *beliebige Variablen und ist* H *ein primitiver Ausdruck, der keine der Variablen* $\mathrm{x}_1, \ldots, \mathrm{x}_n$ *gebunden enthält, so folgt aus* A1 bis A13

$$\bigvee \mathrm{a} \bigwedge \mathrm{x}\big(\mathrm{x} \in \mathrm{a} \leftrightarrow \bigvee \mathrm{x}_1 \ldots \mathrm{x}_n(\mathrm{x} = (\mathrm{x}_1, \ldots, \mathrm{x}_n) \wedge H)\big). \tag{2}$$

Das heißt, jede n-stellige, durch einen primitiven Ausdruck definierbare Relation im Bereich der Elemente existiert als Klasse von n-Tupeln. Ist H ein derartiger Ausdruck, so werden im allgemeinen nicht alle Variablen $\mathrm{x}_1, \ldots, \mathrm{x}_n$ in H frei vorkommen. Von den nicht frei vorkommenden Variablen hängt die durch (2) geforderte Klasse a bzw. die durch sie dargestellte Relation dann nur fiktiv ab. Andererseits können in H auch gewisse Variablen $\mathrm{y}_1, \mathrm{y}_2, \ldots$ frei vorkommen, die von $\mathrm{x}_1, \ldots, \mathrm{x}_n$ verschieden sind. Diese Variablen spielen dann die Rolle von Parametern, von denen die Relation a noch abhängt. Sind $\mathrm{y}_1, \ldots, \mathrm{y}_m$ diese Parameter in einem primitiven Ausdruck H und ist $n = 1$, so ergibt sich als Spezialfall von (2)

$$\bigwedge \mathrm{y}_1 \ldots \mathrm{y}_m \bigvee \mathrm{a} \bigwedge \mathrm{x}\big(\mathrm{x} \in \mathrm{a} \leftrightarrow \mathrm{x}\ \text{Element} \wedge H(\mathrm{x}, \mathrm{y}_1, \ldots, \mathrm{y}_m)\big), \tag{3}$$

d. h. eine unendliche Schar von Klassenbildungsaussagen, die zu den Definitionen

$$F_H(\mathrm{y}_1, \ldots, \mathrm{y}_m) := \iota\mathrm{a} \bigwedge \mathrm{x}\big(\mathrm{x} \in \mathrm{a} \leftrightarrow \mathrm{x}\ \text{Element} \wedge H(\mathrm{x}, \mathrm{y}_1, \ldots, \mathrm{y}_m)\big)$$

berechtigen. Es ist das Verdienst von Gödel, erkannt zu haben, daß man die unendlich vielen Klassenbildungsaxiome (3) und allgemeiner (2) durch ein endliches Axiomensystem (im wesentlichen A6 bis A13) ersetzen kann. Nachträglich erkennt man, daß die Klassenbildungsaxiome A6 bis A13 Spezialfälle von (2) sind. Wir skizzieren einen Beweis des Gödelschen Satzes, dessen Originalfassung natürlich gemäß der dort benutzten Formalisierungsvariante etwas anders lautet.

Hilfssatz 1. *Zu jeder Klasse* a *und beliebigen natürlichen Zahlen* i, n *mit* $i \leqq n$ *existiert eine Klasse* $\mathrm{b} \subseteqq A^n$, *so daß*

$$\mathrm{x}_i \in \mathrm{a} \leftrightarrow (\mathrm{x}_1, \ldots, \mathrm{x}_i, \ldots, \mathrm{x}_n) \in \mathrm{b}.$$

Offenbar leistet $\mathrm{b} = \underbrace{A \times \cdots \times A}_{i-1} \times \mathrm{a} \times \underbrace{A \times \cdots \times A}_{n-i}$ das Verlangte.

Hilfssatz 2. *Zu jeder Klasse* a *und beliebigen natürlichen Zahlen* i, j, n *mit* $i < j \leqq n$ *existiert eine Klasse* $\mathrm{b} \subseteqq A^n$, *so daß* $(\mathrm{x}_i, \mathrm{x}_j) \in \mathrm{a} \leftrightarrow (\mathrm{x}_1, \ldots, \mathrm{x}_i, \ldots, \mathrm{x}_j, \ldots, \mathrm{x}_n) \in \mathrm{b}$.

Zum Beweis können wir o. B. d. A. annehmen, daß a eine Klasse von geordneten Paaren ist, indem wir andernfalls a durch $\mathrm{a} \cap A^2$ ersetzen. Man beweist nun zunächst die drei Spezialfälle $n = 3$ und (a) $i = 1$, $j = 2$, (b) $i = 1$, $j = 3$, (c) $i = 2$, $j = 3$. Im Fall (a) haben wir

$$(\mathfrak{x}_1, \mathfrak{x}_2) \in \mathfrak{a} \leftrightarrow (\mathfrak{x}_1, \mathfrak{x}_2, \mathfrak{x}_3) \in \mathfrak{a} \times A.$$

Im Fall (b) ergibt sich

$$(\mathfrak{x}_1, \mathfrak{x}_3) \in \mathfrak{a} \leftrightarrow (\mathfrak{x}_3, \mathfrak{x}_1) \in \mathfrak{a}^{-1} \leftrightarrow (\mathfrak{x}_3, \mathfrak{x}_1, \mathfrak{x}_2) \in \mathfrak{a}^{-1} \times A$$
$$\leftrightarrow (\mathfrak{x}_1, \mathfrak{x}_2, \mathfrak{x}_3) \in (\mathfrak{a}^{-1} \times A)^{-2}.$$

Im Fall (c) ist

$$(\mathfrak{x}_2, \mathfrak{x}_3) \in \mathfrak{a} \leftrightarrow (\mathfrak{x}_2, \mathfrak{x}_3, \mathfrak{x}_1) \in \mathfrak{a} \times A \leftrightarrow (\mathfrak{x}_3, \mathfrak{x}_1, \mathfrak{x}_2) \in (\mathfrak{a} \times A)^{-2}$$
$$\leftrightarrow (\mathfrak{x}_1, \mathfrak{x}_2, \mathfrak{x}_3) \in \left((\mathfrak{a} \times A)^{-2}\right)^{-2}.$$

Im allgemeinen Fall erhält man zunächst analog zu (c)

$$(\mathfrak{x}_i, \mathfrak{x}_j) \in \mathfrak{a} \leftrightarrow \left((\mathfrak{x}_i, \mathfrak{x}_j), (\mathfrak{x}_1, \ldots, \mathfrak{x}_{i-1})\right) \in \mathfrak{a} \times A^{i-1}$$
$$\leftrightarrow (\mathfrak{x}_1, \ldots, \mathfrak{x}_{i-1}, \mathfrak{x}_i, \mathfrak{x}_j) \in \left((\mathfrak{a} \times A^{i-1})^{-2}\right)^{-2} \qquad (:= \mathfrak{b}).$$

Durch wiederholte Anwendung von (b) ergibt sich nun die Existenz von Klassen $\mathfrak{c}_1, \mathfrak{c}_2, \ldots, \mathfrak{c}_{j-i-1}$, so daß

$$\left((\mathfrak{x}_1, \ldots, \mathfrak{x}_i), \mathfrak{x}_j\right) \in \mathfrak{b} \leftrightarrow \left((\mathfrak{x}_1, \ldots, \mathfrak{x}_i), \mathfrak{x}_{i+1}, \mathfrak{x}_j\right) \in \mathfrak{c}_1,$$
$$\left((\mathfrak{x}_1, \ldots, \mathfrak{x}_{i+1}), \mathfrak{x}_j\right) \in \mathfrak{c}_1 \leftrightarrow \left((\mathfrak{x}_1, \ldots, \mathfrak{x}_{i+1}), \mathfrak{x}_{i+2}, \mathfrak{x}_j\right) \in \mathfrak{c}_2,$$
$$\vdots$$
$$\left((\mathfrak{x}_1, \ldots, \mathfrak{x}_{j-2}), \mathfrak{x}_j\right) \in \mathfrak{c}_{j-i-2} \leftrightarrow \left((\mathfrak{x}_1, \ldots, \mathfrak{x}_{j-2}), \mathfrak{x}_{j-1}, \mathfrak{x}_j\right) \in \mathfrak{c}_{j-i-1},$$

d. h.

$$(\mathfrak{x}_1, \ldots, \mathfrak{x}_j) \in \mathfrak{c}_{j-i-1} \qquad (:= \mathfrak{d}).$$

Durch wiederholte Anwendung von (a) ergibt sich schließlich

$$(\mathfrak{x}_1, \ldots, \mathfrak{x}_j) \in \mathfrak{d} \leftrightarrow (\mathfrak{x}_1, \ldots, \mathfrak{x}_j, \mathfrak{x}_{j+1}) \in \mathfrak{d} \times A$$
$$\leftrightarrow (\mathfrak{x}_1, \ldots, \mathfrak{x}_{j+2}) \in \mathfrak{d} \times A \times A \ldots$$
$$\leftrightarrow (\mathfrak{x}_1, \ldots, \mathfrak{x}_n) \in \mathfrak{d} \times \underbrace{A \times \cdots \times A}_{n-j}.$$

Wir beweisen nun den Satz M2 (Satz von Gödel) durch Induktion über die Kompliziertheit der Ausdrücke, indem wir zu gegebenem primitivem Ausdruck H eine Klasse $K_H^{\mathfrak{x}_1, \ldots, \mathfrak{x}_n}$ konstruieren, so daß

$$(\mathfrak{x}_1, \ldots, \mathfrak{x}_n) \in K_H^{\mathfrak{x}_1, \ldots, \mathfrak{x}_n} \leftrightarrow H$$

gilt. Zunächst betrachten wir prädikative Ausdrücke H der Form $\mathfrak{x} \in \mathfrak{y}$.

1. **Fall**: $\mathfrak{x} = \mathfrak{x}_i$, $\mathfrak{y} = \mathfrak{x}_j$, $i < j$. Dann folgt die Existenz der gesuchten Klasse durch Anwendung von Hilfssatz 2 auf die Klasse E.

2. **Fall:** wie oben, jedoch mit $j < i$. Dann ist Hilfssatz 2 auf E^{-1} anzuwenden.

3. **Fall:** $i = j$ (d. h., H hat die Form $x_i \in x_i$). Dann gilt $H \leftrightarrow x_i \in V(E \cap I)$. Demnach ist Hilfssatz 1 auf die Klasse $V(E \cap I)$ anzuwenden.

4. **Fall:** $x = x_i$, y kommt nicht unter $x_1, \ldots, x_n$ vor. Dann ist Hilfssatz 1 auf die Klasse y anzuwenden.

5. **Fall:** $y = x_i$, x kommt nicht unter $x_1, \ldots, x_n$ vor. Falls x Element ist, gilt

$$x \in x_i \leftrightarrow x_i \in N(E \cap (\{x\} \times A)),$$

demnach ist Hilfssatz 1 auf diese Klasse anzuwenden. Andernfalls gibt es kein x_i mit $x \in x_i$, d. h. $K_H{}^{x_1,\ldots,x_n} := \emptyset$.

6. **Fall:** Weder x noch y kommt unter den Variablen $x_1, \ldots, x_n$ vor. Die jetzt von zwei Parametern x, y abhängige Klasse $K^{x_1,\ldots,x_n}_{x \in y}$ ist durch Fallunterscheidung zu definieren:

$$K^{x_1,\ldots,x_n}_{x \in y}(x, y) := \begin{cases} A^n, & \textit{falls } x \in y, \\ \emptyset, & \textit{falls } x \notin y. \end{cases}$$

Analog (wobei im allgemeinen die Klasse I statt E zu benutzen ist) lassen sich die Fälle behandeln, in denen H die Form $x = y$ hat. Wir nehmen nun an, es seien schon für beliebige Variablensysteme $x_1, \ldots, x_n$ die Klasse $K_H{}^{x_1,\ldots,x_n}$ bzw. die Klassen $K^{x_1,\ldots,x_n}_{H_1}$ und $K^{x_1,\ldots,x_n}_{H_2}$ definiert. Dann ist offenbar

$$K^{x_1,\ldots,x_n}_{\neg H} := A^n \setminus K_H{}^{x_1,\ldots,x_n} \quad \text{(vgl. D20)},$$

$$K^{x_1,\ldots,x_n}_{H_1 \wedge H_2} := K^{x_1,\ldots,x_n}_{H_1} \cap K^{x_1,\ldots,x_n}_{H_2},$$

$$K^{x_1,\ldots,x_n}_{\bigvee x(x \text{ Element} \wedge H)} := V(K_H{}^{x_1,\ldots,x_n,x})$$

zu setzen. (Im letzten Fall muß x eine Variable sein, die in H nicht gebunden vorkommt.) Damit ist der Satz M2 bewiesen.

Wir formulieren nun einige Mengenbildungsaxiome.

A 14 (Potenzmengenaxiom)

$$\bigwedge x(x \text{ Menge} \rightarrow \bigvee y(y \text{ Menge} \wedge \bigwedge z(z \in y \leftrightarrow z \subseteqq x))).$$

D27 $\mathfrak{P}(x) := \iota y(y \text{ Menge} \wedge \bigwedge z(z \in y \leftrightarrow z \subseteqq x))$, *falls* x Menge.

F12 $x \text{ Menge} \wedge z \subseteqq x \rightarrow z \text{ Menge}$.

D28 $\text{Un } f :\leftrightarrow \bigwedge xyz((x, y) \in f \wedge (x, z) \in f \rightarrow y = z)$.

(Un f bedeutet demnach: Die Klasse $(f^{-1})^{-1}$ derjenigen Elemente von f, die geordnete Paare sind, ist eindeutig, d. h. eine Abbildung. Dieser Begriff wird der Bequemlichkeit halber auf beliebige Dinge f übertragen.)

A 15 (Ersetzbarkeitsaxiom)

$$\text{Un } f \wedge V(f) \text{ Menge} \rightarrow N(f) \text{ Menge},$$

(d. h., das eindeutige Bild einer Menge bei einer durch eine Klasse von geordneten Paaren gegebenen Abbildung ist eine Menge. Dieses Axiom drückt die anschauliche Vorstellung aus, daß jede Klasse, die der Mächtigkeit nach nicht größer als eine gewisse Menge ist, selbst eine Menge ist.)

D29 $\bigcup a := K^{x}_{\bigvee y(y \text{ Element} \wedge x \in y \wedge y \in a)}$.

(„y Element" sichert, daß der als unterer Index benutzte Ausdruck primitiv ist und daß damit nach M2 die Klasse $\bigcup a$ existiert. Es kann jedoch semantisch äquivalent fortgelassen werden, da es aus $y \in a$ folgt. Die Klasse $\bigcup a$ heißt *Vereinigung* der Klasse a. Aus D29 folgt, daß $\bigcup a = \emptyset$ ist, falls $a = \emptyset$ oder falls a nur Urelemente und eventuell die leere Menge als Elemente enthält.)

A16 (Axiom der Vereinigungsmenge)

$$a \text{ Menge} \rightarrow \bigcup a \text{ Menge}.$$

Wir formulieren nun eine Art Induktionsaxiom (vgl. das Induktionsaxiom von Peano), das zum Ausdruck bringen soll, daß es außer den Urelementen, der leeren Menge und allen Mengen, die man durch fortlaufende Mengenbildung aus bereits bekannten Elementen gewinnen kann, keine weiteren Elemente gibt.

A17 (Fundierungsaxiom)

$$\bigwedge a\big(\emptyset \in a \wedge \bigwedge u(u \text{ Urelement} \rightarrow u \in a) \wedge \bigwedge x(x \text{ Menge} \wedge x \subseteqq a \rightarrow x \in a) \rightarrow a = A\big).$$

F13 $x \notin x$.

Beweis. Für Unmengen x gilt F13 nach D1 und D4. Es sei a die nach M2 existierende Klasse $K^{x}_{\neg x \in x}$ aller Elemente x mit $x \notin x$. Wir zeigen, daß diese Klasse a die Voraussetzungen von A17 erfüllt. $\emptyset \in a$ ist klar, ebenso $u \in a$ für alle Urelemente u. Es sei $x \subseteqq a$ eine Menge. Wäre $x \notin a$, so wäre $x \in x$. Demnach würde x wenigstens ein nicht zu a gehörendes Element (nämlich x) enthalten, im Widerspruch zu $x \subseteqq a$. Nach A17 folgt nun $a = A$, d. h., F 13 gilt auch für alle Elemente x.

Analog zu F13 beweist man

M3 *Zu keiner natürlichen Zahl* $n \geqq 2$ *existieren Dinge* $x_1, \ldots, x_n$, *so daß*

$$x_1 \in x_2 \wedge x_2 \in x_3 \wedge \ldots \wedge x_{n-1} \in x_n \wedge x_n \in x_1$$

gilt.

F14 A Unmenge.

Beweis. Wäre A keine Unmenge, so wäre A Element, demnach wäre $A \in A$, im Widerspruch zu F13.

Einige Mathematiker haben in Erwägung gezogen, auf A17 zu verzichten und die dann eventuell existierenden Dinge x mit $x = \{x\}$ als Urelemente anzusehen. Dann lassen sich die Urelemente natürlich ohne Benutzung eines zusätzlichen Grundbegriffs (wie der Konstanten $\emptyset$ in unserer Formalisierung) von der leeren Menge unterscheiden. Mit A 17 sind wir bereits in den Bereich solcher Axiome gelangt, über deren Aufnahme in ein Axiomensystem der Mengenlehre nicht alle Mathematiker der gleichen Meinung sind. Weitere umstrittene Axiome sind u. a. das Auswahlaxiom, die Formulierung des Unendlichkeitsaxioms und die verallgemeinerte Kontinuumhypothese.

A 18 (Auswahlaxiom für Klassen)

$$\bigvee a \bigwedge x(x \text{ Menge} \wedge x \neq \emptyset \rightarrow \bigvee!!y((x, y) \in a \wedge y \in x)),$$

d. h., es gibt eine als Klasse von geordneten Paaren existierende *universelle Auswahlfunktion* a, die jeder nichtleeren Menge eines ihrer Elemente zuordnet. Ist a eine universelle Auswahlfunktion und b eine beliebige Klasse von nichtleeren Mengen, so ist $a \cap (b \times A)$ eine *Auswahlfunktion für* b. Umgekehrt schließt die Forderung, daß jede solche Klasse b eine Auswahlfunktion besitzen soll, die Existenz einer Auswahlfunktion für die Klasse A, d. h. einer universellen Auswahlfunktion ein. Ist b eine nichtleere Klasse von nichtleeren und paarweise disjunkten Mengen und a eine Auswahlfunktion für b, so ist N(a) eine *Auswahlklasse für* b, d. h. eine Klasse, die mit jeder Menge aus b genau ein Element gemeinsam hat. Eine echte Abschwächung von A 18, die für viele Anwendungen ausreicht, ist

A 18′ (Auswahlaxiom für Mengen)

$$\bigwedge b(b \text{ Menge} \rightarrow \bigvee a \bigwedge x(x \text{ Menge} \wedge x \neq \emptyset \wedge x \in b \rightarrow \bigvee!!y((x, y) \in a \wedge y \in x))).$$

Während die Mengenbildungsaxiome A 5, A 14 und A 16 und die Klassenbildungsaxiome A 6 bis A 13 auf Grund des Extensionalitätsaxioms A 4 zu solchen Aussagen verschärft werden können, die die Definition von Konstanten (E, I, A usw.) und partiellen Operationen wie Komplement, Durchschnitt, kartesisches Produkt usw. ermöglichen, sind die auf Grund von A 18 existierenden Auswahlmengen bzw. -funktionen im allgemeinen nicht eindeutig bestimmt, vielmehr gibt es in denjenigen Fällen, in denen die Existenz einer solchen Auswahlmenge bzw. -funktion tatsächlich nur auf Grund von A 18 erschlossen werden kann, sogar unendlich viele solche, von denen keine irgendwie vor der anderen ausgezeichnet ist. Daher wird A 18 im Gegensatz zu den übrigen Existenzaussagen des Axiomensystems als nichtkonstruktiv bezeichnet. Hinzu kommt, daß mit Hilfe von A 18 einige paradox erscheinende Sätze bewiesen werden können. Andererseits lassen sich viele wichtige Sätze ohne Benutzung von A 18(′) nicht beweisen bzw. sind zu diesem Axiom äquivalent. Diese Umstände führten dazu, daß die Benutzung des Auswahlaxioms von manchen Mathematikern abgelehnt wird und daß metamathematische Untersuchungen über die Stellung dieses Axioms im Gesamtsystem der axiomatischen Mengenlehre in den Grundlagen der Mengenlehre eine ähnliche hervorragende Rolle gespielt haben wie die Untersuchungen über das Parallelenpostulat in den Grundlagen der Geometrie. Eine der wichtigsten zu A 18 äquivalenten Aussagen ist der Vergleichbarkeitssatz für Mächtigkeiten A 18″. Dazu erst einige Definitionen.

D29 $\quad x \sim y :\leftrightarrow x \text{ Klasse} \wedge y \text{ Klasse} \wedge \bigvee f(\text{Un } f \wedge \text{Un } f^{-1} \wedge V(f) = x \wedge N(f) = y).$

D30 $\quad x \lesssim y :\leftrightarrow \bigvee z(x \sim z \wedge z \subseteqq y).$

A 18″ $\quad x \text{ Klasse} \wedge y \text{ Klasse} \rightarrow x \lesssim y \vee y \lesssim x.$

Wir wenden uns nun dem grundlegenden Begriff der Ordinalzahl zu.

D31 $\quad x \text{ Ordinal} :\leftrightarrow x \text{ Klasse} \wedge \bigwedge y(y \in x \rightarrow y \subseteqq x)$

$$\wedge \bigwedge yz(y \in x \wedge z \in x \rightarrow y \in z \vee z \in y \vee y = z).$$

D32 $\quad x \text{ Ordinalzahl} :\leftrightarrow x \text{ Ordinal} \wedge x \text{ Menge}.$

F15 $\quad \emptyset \text{ Ordinalzahl}.$

F16 $\quad x \text{ Ordinalzahl} \rightarrow x \cup \{x\} \text{ Ordinalzahl}.$

F17 $\quad \bigvee a \bigwedge x(x \in a \leftrightarrow x \text{ Ordinalzahl}).$

D33 $\quad \text{On} := \iota a \bigwedge x(x \in a \leftrightarrow x \text{ Ordinalzahl}).$

F18 $\quad x \text{ Ordinal} \leftrightarrow x \in \text{On} \vee x = \text{On}.$

F19 $\quad \text{On Unmenge}.$

F20 $\quad x \in \text{On} \wedge y \in \text{On} \wedge z \in \text{On} \wedge x \in y \wedge y \in z \rightarrow x \in z.$

F21 $\quad y \subseteqq \text{On} \wedge y \neq \emptyset \rightarrow \bigvee!!x\big(x \in y \wedge \bigwedge z(z \in y \wedge z \neq x \rightarrow x \in z)\big).$

Die Sätze F15 bis F21 sind nicht schwer zu beweisen und seien dem Leser als Übungsaufgaben empfohlen. F20 und F21 bedeuten zusammen:

Die Klasse On *ist durch die zweistellige Relation* $\in$ *wohlgeordnet.*

Dabei ist $\emptyset$ offenbar das kleinste Element von On. Weiter ergibt sich, daß $x \cup \{x\}$ der unmittelbare Nachfolger einer Ordinalzahl x ist, daß jede Ordinalzahl gleich der Menge der ihr vorangehenden Ordinalzahlen ist und daß $\bigcup y$ die erste Ordinalzahl ist, die auf alle Elemente von y folgt, falls y eine Menge von Ordinalzahlen ohne größtes Element ist. F21 rechtfertigt die Definition

D34 $\quad \min y := \iota x\big(x \in y \wedge \bigwedge z(z \in y \wedge z \neq x \rightarrow x \in z)\big), \quad$ *falls* $y \subseteqq \text{On} \wedge y \neq \emptyset.$

Es ist üblich, die natürlichen Zahlen mit dem Anfang der wohlgeordneten Ordinalzahlreihe zu identifizieren, d. h.

D35 $\quad 0 := \emptyset,$

$$1 := \emptyset \cup \{\emptyset\} = \{\emptyset\},$$

$$2 := \{\emptyset\} \cup \{\{\emptyset\}\} = \big\{\emptyset, \{\emptyset\}\big\},$$

$$\vdots$$

$$n + 1 := n \cup \{n\}.$$

D36 $\quad x \text{ natürliche Zahl} :\leftrightarrow x \in \text{On} \wedge \bigwedge y(y \in x \rightarrow y = \emptyset \vee \bigvee z\, z \cup \{z\} = y),$

d. h., natürliche Zahlen sind solche Ordinalzahlen, denen außer $\emptyset$ keine Ordinalzahl ohne unmittelbaren Vorgänger vorangeht.

Unter Benutzung von A 17 kann man jedem Ding x ein Ordinal st(x) (Stufe von x) so zuordnen, daß

$$\mathrm{st}(x) = 0, \quad \textit{falls } x \text{ Urelement}, \ \mathrm{st}(\emptyset) = 1,$$

$$\mathrm{st}(x) = \min\left(\mathrm{On} \setminus \{\mathrm{st}(y) : y \in x\}\right), \quad \textit{falls } x \text{ nichtleere Menge},$$

$$\mathrm{st}(x) = \mathrm{On}, \quad \textit{falls } x \text{ Unmenge}$$

gilt. Demnach ist also die Stufe einer nichtleeren Klasse das kleinste Ordinal, das größer ist als die Stufen aller ihrer Elemente. Im axiomatischen Aufbau der Mengenlehre nach KLAUA [20] wird die hieraus ableitbare Relation $x \sqsubseteq y$ (x ist stufenkleinergleich y) neben $\in$ als Grundbegriff benutzt. Diese Formalisierung erweist sich hinsichtlich ihrer Ausdrucksfähigkeit als gleichwertig zur vorliegenden, da man unter Benutzung von $\sqsubseteq$

$$x \text{ Urelement} :\leftrightarrow \bigwedge y\, x \sqsubseteq y$$

$$\emptyset := \iota x(\neg\, x \text{ Urelement} \wedge \neg \bigvee y\, y \in x)$$

definieren kann.

Die Axiome A 1 bis A 18 besitzen ein Modell, das aus (im anschaulichen Sinne) abzählbar vielen endlichen Mengen und Urelementen und abzählbar vielen Unmengen besteht. Gibt man sich etwa zwei beliebige Dinge a, b als Urelemente vor, so erhält man

Mengen 1. Stufe: $\emptyset$, $\{a\}$, $\{b\}$, $\{a, b\}$,
Mengen 2. Stufe: $\{\emptyset\}$, $\{\{a\}\}$, $\{\{b\}\}$, $\{\{a, b\}\}$, $\{\emptyset, a\}$, $\{\emptyset, b\}$, $\{\emptyset, \{a\}\}$, $\{\emptyset, \{b\}\}$, $\{\emptyset, \{a, b\}\}$, $\{\{a\}, a\}$, $\{\{a\}, b\}$, $\{\{a\},\{b\}\}$, $\{\{a\},\{a, b\}\}$, $\{\{b\}, a\}$, $\{\{b\}, b\}$, $\{\{b\}, \{a, b\}\}$, $\{\{a, b\}, a\}$, $\{\{a, b\}, b\}$, $\{\emptyset, a, b\}$, $\{\emptyset, a, \{a\}\}$ usw.

In diesem Bereich ist die Klasse aller Elemente offenbar abzählbar. Es genügt, als Unmengen des Modells alle abzählbaren Teilmengen dieses Bereichs hinzuzunehmen, die die Form $K_H{}^x$ haben. Wegen der Abzählbarkeit der Menge der Ausdrücke von S_m ist der Bereich dieser Unmengen ebenfalls abzählbar. Das skizzierte Modell läßt sich unter der Voraussetzung der Widerspruchsfreiheit der Peanoschen Axiome exakt definieren. Analoge Modelle erhält man bei Vorgabe einer beliebigen endlichen Anzahl von Urelementen, insbesondere auch, wenn man von der Nichtexistenz von Urelementen ausgeht. Demnach gilt:

Aus A 1 bis A 18 *folgt nicht die Existenz von Urelementen und nicht die Existenz einer unendlichen Menge.*

Insbesondere ist in den genannten Modellen die Klasse der natürlichen Zahlen gleich On, also bereits eine Unmenge. Daher können die weiteren Zahlbereiche auf der Grundlage von A 1 bis A 18 nicht in der üblichen Weise konstruiert werden. Wir beschäftigen uns nun mit einigen Varianten von Unendlichkeitsaxiomen, d. h.

Axiomen, die die Existenz einer unendlichen Menge aussagen. Diesen Zweck erfüllt

A 19 $\quad \bigvee a\big(a \text{ Menge} \wedge \emptyset \in a \wedge \bigwedge x(x \in a \rightarrow \{x\} \in a)\big).$

Dieses Axiom bedeutet: Es gibt eine (im anschaulichen Sinne) unendliche Menge, deren Stufe keine natürliche Zahl ist. Hieraus folgt, daß es wenigstens eine Ordinalzahl gibt, die keine natürliche Zahl ist, und folglich (etwa nach F 16) unendlich viele solche Ordinalzahlen.

Will man die Existenz einer unendlichen Menge schon im Bereich der Mengen endlicher Stufe fordern, so ist dies offenbar gleichbedeutend mit der Forderung der Existenz einer unendlichen Menge von Urelementen. Man kann in Anlehnung an A 19 etwa formulieren:

A 19′ $\quad \bigvee f\big(\mathrm{Un}\, f \wedge \mathrm{Un}\, f^{-1} \wedge V(f) \text{ Menge} \wedge \emptyset \in V(f) \wedge \bigwedge x\big(x \in V(f) \rightarrow \{x\} \in V(f)\big)$

$\wedge \bigwedge y\big(y \in N(f) \rightarrow y \text{ Urelement}\big)\big).$

Offenbar folgt A 19 aus A 19′, da $V(f) = a$ der Forderung von A 19 genügt, falls f der Forderung von A 19′ genügt. Es sei bemerkt, daß offen bleibt, ob die Klasse aller Urelemente eine Menge oder eine Unmenge ist, und daß sich aus der scheinbar stärkeren Forderung, daß die Klasse aller Urelemente eine Unmenge ist, nicht die Existenz einer unendlichen Menge von Urelementen oder überhaupt einer unendlichen Menge ergibt. Will man die Existenz einer unendlichen Menge endlicher Stufe fordern, ohne gleichzeitig die Existenz von Mengen unendlicher Stufe vorauszusetzen, so benötigt man eine explizite Endlichkeitsdefinition.

Eine *explizite Endlichkeitsdefinition* ist eine Definition der Form

$$x \text{ endlich} :\leftrightarrow H(x),$$

wobei $H(x) \in S_m$ keine freie Variable außer x enthält. Dabei soll in jedem Modell ω der Axiome A 1 bis A 17 folgendes gelten:

(∗) *Wert* $\big(H(x), \omega, f\big) = W \leftrightarrow$ *Es gibt nur endlich viele Dinge* ξ *im Grundbereich von* ω*, für die Wert* $\left(z \in x, \omega, f\left\langle \begin{matrix} z \\ \xi \end{matrix} \right\rangle\right) = W.$

(Die von uns gewählte Formulierung schließt ein, daß die Eigenschaft „endlich“ auf alle Urelemente zutreffen soll. Die einstelligen Relationen „x endlich“ und „x endliche Menge“ sind aber offenbar gegenseitig definierbar.) Es wird sich zeigen (vgl. 6.3.), daß eine Endlichkeitsdefinition in diesem strengen Sinne unmöglich ist: Jede in einer elementaren Sprache der Mengenlehre definierbare einstellige Relation H(x), die in jedem Modell der Mengenlehre auf alle (im anschaulichen bzw. metatheoretischen Sinne) endlichen Objekte zutrifft, trifft in wenigstens einem Modell auch auf eine im anschaulichen Sinne unendliche Gesamtheit zu. Demnach muß die Definition des Begriffs Endlichkeitsdefinition dahin abgeschwächt werden, daß in (∗) „↔“ durch „←“ ersetzt wird. Weiterhin muß man aber noch fordern, daß aus einer Endlichkeitsdefinition alle häufig benutzten Eigenschaften der endlichen

Mengen gefolgert werden können, wie z. B.

$\emptyset$ endlich,

$\{x\}$ endlich,

x endliche Menge $\wedge$ y endliche Menge $\rightarrow x \cup y$ endlich $\wedge\ x \times y$ endlich,

x endliche Menge $\rightarrow \mathfrak{P}(x)$ endlich,

x endliche Menge $\wedge\ y \subseteqq x \rightarrow y$ endlich.

In der Literatur sind zahlreiche Endlichkeitsdefinitionen verbreitet, die allen diesen Forderungen genügen. Hinzu kommt, daß alle bisher vorgeschlagenen Endlichkeitsdefinitionen sich unter der Voraussetzung des Auswahlaxioms als äquivalent erweisen. Manche dieser Äquivalenzen gelten jedoch nachweislich ohne Voraussetzung des Auswahlaxioms nicht allgemein.[1] Wir stellen zwei der verbreitetsten Endlichkeitsdefinitionen vor, die von DEDEKIND bzw. RUSSELL eingeführt wurden:

D36 x endlich (D) $:\leftrightarrow \neg \bigvee y(y \subset x \wedge y \sim x)$.

Hiernach sind Urelemente x automatisch endlich, da sie nach D8 bzw. D9 nicht in der Relation $\subset$ stehen können. Eine Klasse x ist nach DEDEKIND endlich, wenn sie keiner echten Teilklasse gleichmächtig ist. Es folgt z. B. sofort: $\emptyset$ endlich (D), da $\emptyset$ keine echte Teilklasse besitzt.

D37 x endlich (R) $:\leftrightarrow \bigwedge a(\emptyset \in a \wedge \bigwedge u(u$ Urelement $\rightarrow u \in a)$

$\wedge \bigwedge yz(y \in a \wedge y$ Klasse $\wedge\ z$ Element $\rightarrow y \cup \{z\} \in a) \rightarrow x \in a)$,

d. h., ein Ding x ist endlich im Sinne von RUSSELL, wenn es Element jeder Klasse a ist, die $\emptyset$ und alle Urelemente enthält und mit einer beliebigen Menge y stets auch die (eventuell) um ein Element größere Menge $y \cup \{z\}$ bei beliebigem z enthält. Hiernach ist z. B. sofort klar, daß $\emptyset$ endlich (R), u endlich (R) für alle Urelemente u und $\{z\}$ ($= \emptyset \cup \{z\}$) endlich (R) für jedes Element z ist. (Beide Definitionen sind gegenüber ihrer ursprünglichen Formulierung geringfügig modifiziert, um sie unserer Formalisierungsvariante mit Urelementen anzupassen.) Es gilt ohne Voraussetzung des Auswahlaxioms

F22 x endlich (R) $\rightarrow$ x endlich (D).

Unter Benutzung von A18′ folgt

F23 x endlich (D) $\rightarrow$ x endlich (R).

Weiterhin gilt jedoch:

F23 ist unabhängig von A1 bis A17 und A19′.

[1] Siehe hierzu etwa A. LEVY, The independence of various definitions of finiteness, Fund. Math. 46 (1958), 1–13.

Unter Benutzung einer beliebigen Endlichkeitsdefinition kann man verschiedene weitere Varianten des Unendlichkeitsaxioms formulieren, z. B.

A 19'' $\quad \bigvee a(\neg \text{ endlich } a \wedge a \text{ Menge})$

oder schärfer

A 19''' $\quad \bigvee a(a \text{ Menge} \wedge \neg \text{ endlich } a \wedge \bigwedge x(x \in a \rightarrow x \text{ Urelement}))$.

Offenbar ist A 19' die stärkste und A 19'' die schwächste Forderung, während A 19 und A 19''' voneinander unabhängig sind.

Wir kommen nun zur Diskussion der verallgemeinerten Kontinuumhypothese. Aus A 1 bis A 17 folgt

$$x \text{ endlich (D)} \wedge x \neq \emptyset \wedge \neg \bigvee y\, x = \{y\} \rightarrow \bigvee y(y \subset \mathfrak{P}(x) \wedge \neg\, y \lesssim x$$
$$\wedge \neg\, y \sim \mathfrak{P}(x)).$$

(Dies entspricht der Tatsache, daß für natürliche Zahlen $n \geqq 2$ (hier als Mächtigkeiten endlicher Mengen aufgefaßt) stets ein m mit $n < m < 2^n$ (der Elementanzahl der Potenzmenge) existiert.) Schon CANTOR scheiterte an dem Versuch, für unendliche Mengen (insbesondere abzählbare) entweder die Existenz oder die Unmöglichkeit der Existenz einer Menge nachzuweisen, die der Mächtigkeit nach echt zwischen der gegebenen Menge und ihrer Potenzmenge liegt. Für verschiedene Anwendungen innerhalb der Mathematik benötigt man aber die im folgenden als Axiom A 20 notierte Voraussetzung

A 20 (Verallgemeinerte Kontinuumhypothese)

$$\neg\, x \text{ endlich (D)} \rightarrow \bigwedge y(y \subseteqq \mathfrak{P}(x) \wedge x \lesssim y \rightarrow y \sim \mathfrak{P}(x)).$$

Aus dem Axiom A 20 *folgt das Auswahlaxiom.*[1])

Wir beschließen dieses Kapitel mit der Mitteilung einiger wichtiger Resultate aus den Grundlagen der Mengenlehre. Hierzu sei im voraus bemerkt, daß diese Resultate in bezug auf unterschiedliche und nicht immer voll formalisierte Axiomensysteme erhalten wurden und daher nicht ohne weiteres auf andere Formalisierungsvarianten (insbesondere die hier benutzte) übertragen werden können. Weiter soll nicht unerwähnt bleiben, daß die zur Erzielung derartiger metamathematischer Sätze im Bereich der Grundlagen der Mengenlehre erforderlichen Modellkonstruktionen bzw. sonstigen Methoden häufig sehr kompliziert und jeder Anschauung entzogen sind und zum Teil zu den hervorragendsten Leistungen der Mathematik in unserem Jahrhundert gehören.[2])

---

[1]) W. SIERPIŃSKI, Fund. Math. 34 (1947), 1—5.

[2]) Ein gut zugänglicher und relativ aktueller Bericht über den Stand der Grundlagen der Mengenlehre ist der Kommentar zum ersten Hilbertschen Problem von A. S. ESENIN-VOL'PIN in: Die Hilbertschen Probleme, Akad. Verlagsgesellschaft Leipzig 1971 (russ. Original: Nauka, Moskau 1969). Außerdem sei auf die umfangreiche Bibliographie in [9] verwiesen.

Das System von v. NEUMANN-BERNAYS-GÖDEL (mit Unmengen) ist relativ widerspruchsfrei in bezug auf das System von ZERMELO-FRAENKEL (ohne Unmengen) und umgekehrt (ROSSER, HAO WANG 1950, u. a.).

Die Annahme der Existenz unendlich vieler Urelemente ist relativ widerspruchsfrei in bezug auf ein System ohne Urelemente und umgekehrt (MOSTOWSKI 1939).

Sind die Axiome A 1 bis A 17 und A 19 widerspruchsfrei, so ist auch das System der Axiome A 1 bis A 20 widerspruchsfrei (GÖDEL 1938).

Das Auswahlaxiom ist unabhängig von den Axiomen A 1 bis A 17 und A 19 bzw. A 19' (FRAENKEL 1922, MOSTOWSKI 1939, COHEN 1963).

Das Axiom A 20 ist unabhängig von den übrigen Axiomen (einschließlich Auswahlaxiom) (COHEN 1963).

Die beiden letztgenannten Sätze bedeuten insbesondere, daß sowohl das Auswahlaxiom als auch (bei Forderung des Auswahlaxioms) die verallgemeinerte Kontinuumhypothese mit gleicher Berechtigung als gültig oder ungültig angesehen werden können. Die auf den ersten Blick verwirrende Erkenntnis, daß sich die Verhältnisse im Bereich unendlicher Mengen nicht eindeutig festlegen lassen, hat bei vielen Mathematikern Zweifel an der Brauchbarkeit der axiomatischen Methode ausgelöst. Auf die hiermit zusammenhängenden Fragen kommen wir im letzten Kapitel zurück.

6. Syntaktische Grundbegriffe der Metamathematik

6.1. Schlußregeln und Beweiskalküle

Definition 1. Es sei S eine formalisierte Sprache, $m, k \geqq 0$ seien beliebige natürliche Zahlen. Eine *k-gliedrige m-parametrige Schlußregel für S* ist eine $(k + m)$-stellige Wortfunktion φ mit folgenden Eigenschaften:

(a) *Ist für gewisse Wörter $W_1, \dots, W_{k+m}$ die Anwendung von φ definiert, so ist $W_i \in S$ für $i = 1, \dots, k$ und*

$$\varphi(W_1, \dots, W_{k+m}) \in Fl_S(\{W_1, \dots, W_k\}). \tag{1}$$

(b) *In einem später zu präzisierenden Sinne ist für beliebige $(k + m)$-Tupel von Wörtern effektiv entscheidbar, ob φ anwendbar ist, und im Fall der Anwendbarkeit läßt sich $\varphi(W_1, \dots, W_{k+m})$ aus den Wörtern $W_1, \dots, W_{k+m}$ durch ein algorithmisches Verfahren gewinnen.*

Unter Berücksichtigung der Bedingung (a) schreiben wir die Argumente einer k-gliedrigen m-parametrigen Schlußregel φ im folgenden in der Form $H_1, \dots, H_k; P_1, \dots, P_m$ und bezeichnen $H_1, \dots, H_k$ als *Prämissen* (Voraussetzungen), $P_1, \dots, P_m$ als *Parameter* und das Resultat $\varphi(H_1, \dots, H_k; P_1, \dots, P_m)$ der Anwendung von φ auf $H_1, \dots, H_k; P_1, \dots, P_m$ als *Conclusio* (Folgerung) der Schlußregel φ. Die Bedingung (1) bezeichnen wir als *Zulässigkeit* der Regel φ. Auf Grund der Hülleneigenschaften des Folgerungsoperators Fl_S ist sie der folgenden Bedingung gleichwertig:

Für beliebige $X \subseteqq S$ gilt: Ist $H_i \in Fl_S(X)$ für $i = 1, \dots, k$ und existiert das Resultat $\varphi(H_1, \dots, H_k; P_1, \dots, P_m)$, so ist $\varphi(H_1, \dots, H_k; P_1, \dots, P_m) \in Fl_S(X)$.

0-parametrige Schlußregeln bezeichnen wir auch als *parameterfrei*. Ist umgekehrt φ eine 0-gliedrige m-parametrige Schlußregel ($m > 0$), so gilt für alle Parametertupel $P_1, \dots, P_m$ aus dem Definitionsbereich von φ nach (1) die Beziehung

$$\varphi(P_1, \dots, P_m) \in Fl_S(\emptyset).$$

Daher nennen wir die Menge aller Ausdrücke der Form $\varphi(P_1, \ldots, P_m)$ und auch die Abbildung φ, die diese Menge „aufzählt" bzw. als Wertebereich hat, ein *Schema logischer Axiome*. Ist $k = m = 0$, so ist φ eine Konstante und daher ein einzelnes *logisches Axiom*. Um 0-parametrige und 0-gliedrige Schlußregeln auch bezeichnungsmäßig möglichst deutlich voneinander zu unterscheiden, schreiben wir im ersten Fall $\varphi(H_1, \ldots, H_k)$ und im zweiten Fall $\varphi(;P_1, \ldots, P_m)$.

Beispiele.

1. *Regel der Konstanteneinsetzung.*

$\varphi_0(H(x); x, c) := Sub\,(H(x), x, c)$, *falls* x *in* H(x) *vollfrei vorkommt und* c *ein zu* x *sortengleiches Konstantensymbol ist.*

Die so definierte Regel φ_0 ist offenbar eingliedrig und zweiparametrig. Ihre Effektivität ist anschaulich klar. Ihre Zulässigkeit ergibt sich aus $H(c) \in Fl_S(\{H(x)\})$.

2. *Abtrennungsregel.*

$$\varphi_1(H_1 \to H_2, H_1) := H_2 \qquad (H_1, H_2 \in S \textit{ beliebig}).$$

Diese Regel wird in der klassischen Logik auch als *modus ponens* bezeichnet. Sie ist zweigliedrig und parameterfrei.

3. Für $m \geqq 1$ sei $\varphi_1{}^m(;A, H_1, \ldots, H_m)$ definiert, falls $H_\mu \in S$ $(\mu = 1, \ldots, m)$ und A ein allgemeingültiger Ausdruck der Aussagenlogik ist, der höchstens die Variablen $p_1, \ldots, p_m$ enthält. Es sei dann $\varphi_1{}^m(;A, H_1, \ldots, H_m)$ derjenige Ausdruck der Sprache S, den man erhält, wenn man in A jede der Variablen p_μ an allen Stellen ihres Vorkommens durch den Ausdruck H_μ ersetzt. Die so definierten Schlußregeln sind offenbar $(m + 1)$-parametrige Axiomenschemata, die in ihrer Gesamtheit genau die Menge aller aussagenlogisch allgemeingültigen Ausdrücke der Sprache S umfassen. Die Effektivität der Schlußregeln $\varphi_1{}^m$ ergibt sich aus der Entscheidbarkeit der aussagenlogischen Allgemeingültigkeit (etwa durch Wertetabelle; vgl. 2.2.). Es sei noch bemerkt, daß der in diesen Regeln bzw. Axiomenschemata vorkommende erste Parameter A zwar auch ein Wort in einem gewissen Alphabet, jedoch im Unterschied zu den sonstigen in den bisher betrachteten Beispielen vorkommenden Parametern weder ein Ausdruck noch ein Term der betrachteten Sprache S ist.

4. Wir wollen nun die bereits in 4.2. behandelte *Regel der gebundenen Umbenennung* als formale Schlußregel im hier definierten Sinne präzisieren. Es sei zunächst daran erinnert, daß H_2 durch gebundene Umbenennung aus H_1 entsteht, indem man in einem gewissen Vorkommen eines gewissen Teilausdrucks von H_1, der die Form $\bigwedge xH(x)$ oder $\bigvee xH(x)$ hat, überall x durch eine Variable y gleicher Sorte ersetzt, die folgende Bedingungen erfüllt:

a) (Vermeidung von *Variablenkonfusion*) y kommt im Ausdruck H(x) nicht frei vor (da solche Vorkommen durch die Umbenennung inhaltlich unrichtig in die Quantifizierung einbezogen würden).

b) (Vermeidung von *Variablenkollision*) y steht im betrachteten Vorkommen des Teilausdrucks $H(x)$ in H_1 nirgends im Wirkungsbereich einer in H_1 vorkommenden Quantifizierung von y (da solche Vorkommen von y durch die Umbenennung zweimal quantifiziert würden).

Eine Variable y, die die Bedingungen a) und b) erfüllt, nennen wir (für die beabsichtigte Umbenennung) *verfügbar*. Sehen wir nun H_1 als Prämisse und H_2 als conclusio der zu definierenden Schlußregel an, so müssen die Parameter dieser Schlußregel angeben, in welchem Teilausdruck $H(x)$ von H_1 und welchem Vorkommen dieses Teilausdrucks die Umbenennung ausgeführt werden soll und welche Variable y benutzt werden soll. Demnach können wir etwa definieren:

$\varphi_2(H_1; P, y) := H_2$, *falls* $H_1 \in S$, P *ein auf* $\wedge$ *oder* $\vee$ *endendes Anfangsstück von* H_1 *ist (worauf dann ein eindeutig bestimmtes Teilstück der Form* $xH(x)$ *folgen muß), die Variable* y *von gleicher Sorte wie* x *und für die Umbenennung an dieser Stelle verfügbar ist und* H_2 *dadurch aus* H_1 *entsteht, daß man in dem auf* P *folgenden Teilstück* $xH(x)$ *überall* x *durch* y *ersetzt.*

5. Wir formulieren nun vier Schlußregeln, die in diesem Buch keine weitere Rolle spielen werden, die jedoch zusammen mit φ_1, φ_2 und einer noch zu behandelnden Verallgemeinerung φ_7 von φ_0 einen in der traditionellen Logik häufig benutzten Kalkül bilden:

Regel der vorderen Generalisierung:

$\varphi_3\big(H_1(x) \rightarrow H_2; x\big) := \wedge\, xH_1(x) \rightarrow H_2$, *falls* x *vollfrei in* $H_1(x)$ *vorkommt.*

(Die Angabe des Parameters x ist nötig, da $H_1(x)$ ungeachtet der von uns bereits gewählten Schreibweise außer x weitere vollfreie Variablen enthalten kann. Analoges gilt für die folgenden Regeln.)

Regel der hinteren Partikularisierung:

$\varphi_4\big(H_1 \rightarrow H_2(x); x\big) := H_1 \rightarrow \vee\, xH_2(x)$, *falls* x *vollfrei in* $H_2(x)$ *vorkommt.*

Regel der hinteren Generalisierung:

$\varphi_5(H_1 \rightarrow H_2(x); x) := H_1 \rightarrow \wedge\, xH_2(x)$, *falls* x *in* H_2 *vollfrei und in* H_1 *nicht frei vorkommt.*

Regel der vorderen Partikularisierung:

$\varphi_6(H_1(x) \rightarrow H_2; x) := \vee\, xH_1(x) \rightarrow H_2$, *falls* x *in* H_1 *vollfrei und in* H_2 *nicht frei vorkommt.*

Der Leser überprüfe die Zulässigkeit der vier Regeln und mache sich klar, daß die Regeln φ_5, φ_6 ohne die genannten einschränkenden Bedingungen nicht zulässig wären.

Der Begriff der Schlußregel läßt sich leicht auf beliebige nichtelementare Sprachen (S, σ) übertragen. Dazu hat man lediglich in Definition 1 die Bedingung (1) in

$$\varphi(W_1, \ldots, W_{k+m}) \in Fl_S^\sigma(\{W_1, \ldots, W_k\}) \tag{1'}$$

abzuändern (*Zulässigkeit der Schlußregel* φ *für* (S, σ)). Sind σ_1, σ_2 Klassen von Interpretationen der Sprache S mit $\sigma_1 \subseteq \sigma_2$, so ist wegen

$$Fl_S^{\sigma_2}(X) \subseteq Fl_S^{\sigma_1}(X) \quad \text{für } X \subseteq S$$

jede für (S, σ_2) zulässige Schlußregel erst recht für (S, σ_1) zulässig, insbesondere ist also jede für (S, σ_0) zulässige Schlußregel (solche Schlußregeln werden wir als *elementar* bezeichnen) für jede nichtelementare Sprache (S, σ) zulässig. Als Beispiel einer *nichtelementaren* (d. h. für (S, σ_0) nicht zulässigen) Schlußregel betrachten wir die

Termeinsetzungsregel:

$\varphi_7(\mathrm{H}(\mathrm{x}); \mathrm{x}, \mathrm{t}) := Sub\,(\mathrm{H}(\mathrm{x}), \mathrm{x}, \mathrm{t})$, *falls* x *vollfrei in* H(x) *vorkommt und* t *ein zu* x *sortengleicher Term ist.*

Die eingliedrige einparametrige Schlußregel φ_7 ist offenbar eine naheliegende Verallgemeinerung der Regel φ_0 der Konstanteneinsetzung und enthält diese Regel als Spezialfall. Sie ist jedoch nicht zulässig für (S, σ_0), falls S eine Sprache mit Operationssymbolen ist, da sie z. B. den allgemeingültigen Ausdruck $\bigvee \mathrm{y}\, \mathrm{y} = \mathrm{x}$ in den nicht allgemeingültigen Ausdruck $\bigvee \mathrm{y}\, \mathrm{y} = \mathrm{F}(\mathrm{x}_1, \ldots, \mathrm{x}_n)$ überführt (F ein n-stelliges Operationssymbol von S). Wie man sieht, ist die Regel φ_7 gerade zulässig für nichtelementare Sprachen $(S, \sigma_{(1)})$. (Wir erinnern daran, daß $\sigma_{(1)}$ stets die Klasse aller Interpretationen bezeichnet, die jedem Operationssymbol eine volle Operation zuordnen; vgl. 4.5.)

Definition 2. Es sei S eine formalisierte Sprache. Eine nichtleere (endliche oder wenigstens in einem später zu präzisierenden Sinne effektiv aufzählbare) Menge $\mathcal{K}$ von Schlußregeln definiert einen *Beweiskalkül* für S. (Analog definiert eine Menge $\mathcal{K}$ von für (S, σ) zulässigen Schlußregeln einen Beweiskalkül für die nichtelementare Sprache (S, σ).) Ist $\mathcal{K}$ ein Beweiskalkül für S, so heißt eine Menge $X \subseteq S$ von Ausdrücken *$\mathcal{K}$-abgeschlossen*, wenn für $\varphi \in \mathcal{K}$ gilt:

Wenn $\varphi(\mathrm{H}_1, \ldots, \mathrm{H}_k; \mathrm{P}_1, \ldots, \mathrm{P}_m) = \mathrm{H}$ existiert und $\mathrm{H}_\varkappa \in X$ für $\varkappa = 1, \ldots, k$ gilt, so ist $\mathrm{H} \in X$.

Für $X \subseteq S$ sei

$$Bew_{\mathcal{K}}(X) := \bigcap_{\substack{X \subseteq Y \subseteq S \\ Y\ \mathcal{K}\text{-abge-} \\ \text{schlossen.}}} Y$$

$Bew_{\mathcal{K}}(X)$ heißt die *Beweishülle* von X (bezüglich $\mathcal{K}$).

Zunächst erläutern wir die Definition 2: $\mathcal{K}$-Abgeschlossenheit einer Menge $\boldsymbol{X}$ von Ausdrücken schließt insbesondere ein, daß $\boldsymbol{X}$ jede durch eine etwa zu $\mathcal{K}$ gehörige 0-gliedrige Schlußregel aufgezählte Menge von logischen Axiomen enthält. Zerlegt man insbesondere einen Beweiskalkül, wie es traditionell üblich ist, in eine Menge $\boldsymbol{A}$ von logischen Axiomen und eine Menge $\mathcal{K}'$ von *echten*, d. h. nicht 0-gliedrigen Schlußregeln, so ist eine Menge $\boldsymbol{X} \subseteq \boldsymbol{S}$ genau dann $\mathcal{K}$-abgeschlossen, wenn sie $\boldsymbol{A}$ umfaßt und $\mathcal{K}'$-abgeschlossen ist. Die Beweishülle $Bew_{\mathcal{K}}(\boldsymbol{X})$ läßt sich dann auch in der Form $Bew_{\mathcal{K}'}(\boldsymbol{X} \cup \boldsymbol{A})$ darstellen. Allgemein gilt: Die gesamte Ausdrucksmenge $\boldsymbol{S}$ ist stets $\mathcal{K}$-abgeschlossen, und der Durchschnitt eines beliebigen nichtleeren Systems von $\mathcal{K}$-abgeschlossenen Teilmengen von $\boldsymbol{S}$ ist $\mathcal{K}$-abgeschlossen. Demnach ist $Bew_{\mathcal{K}}(\boldsymbol{X})$ gerade die kleinste $\mathcal{K}$-abgeschlossene Obermenge von $\boldsymbol{X}$.

Satz 1. *Ist $\mathcal{K}$ ein beliebiger Beweiskalkül für $\boldsymbol{S}$ (bzw. für $(\boldsymbol{S}, \sigma)$), so ist $Bew_{\mathcal{K}}$ ein Hüllenoperator.*

Der Beweis von Satz 1 sei dem Leser überlassen.

Definition 3. Es sei $\boldsymbol{S}$ eine formalisierte Sprache, $\mathcal{K}$ ein Beweiskalkül für $\boldsymbol{S}$, $\boldsymbol{X} \subseteq \boldsymbol{S}$ ein Axiomensystem und $\mathrm{H} \in \boldsymbol{S}$ ein beliebiger Ausdruck. Ein *$\mathcal{K}$-Beweis für* H (*aus* $\boldsymbol{X}$) ist eine endliche Folge $\mathrm{H}_1, \mathrm{H}_2, \ldots, \mathrm{H}_n$ von Ausdrücken aus $\boldsymbol{S}$, so daß gilt: $\mathrm{H}_n = \mathrm{H}$ und für $1 \leqq i \leqq n$ ist $\mathrm{H}_i \in \boldsymbol{X}$, oder es existieren eine k-gliedrige m-parametrige Schlußregel $\varphi \in \mathcal{K}$, Parameter $\mathrm{P}_1, \ldots, \mathrm{P}_m$ und (nicht notwendig verschiedene) Indizes $i_1, \ldots, i_k < i$, so daß

$$\mathrm{H}_i = \varphi(\mathrm{H}_{i_1}, \ldots, \mathrm{H}_{i_k}; \mathrm{P}_1, \ldots, \mathrm{P}_m)$$

gilt. (Ist insbesondere hierbei $k = 0$, so ist H_i demnach ein logisches Axiom des Kalküls $\mathcal{K}$.)

Definition 3 bedeutet also, daß ein Beweis (bezüglich $\mathcal{K}$) eine endliche Folge von Ausdrücken ist, von denen jeder ein logisches Axiom des Kalküls oder ein (inhaltliches) Axiom des benutzten Axiomensystems ist oder durch Anwendung einer echten Schlußregel auf gewisse frühere Folgenglieder entsteht. Hieraus folgt sofort, daß jedes nichtleere Anfangsstück eines Beweises ebenfalls ein Beweis ist.

Satz 2. $\mathrm{H} \in Bew_{\mathcal{K}}(\boldsymbol{X})$ *gilt genau dann, wenn ein $\mathcal{K}$-Beweis für* H *aus* $\boldsymbol{X}$ *existiert*

Beweis. Es bezeichne $B(\boldsymbol{X})$ die Menge aller Ausdrücke, für die ein $\mathcal{K}$-Beweis aus $\boldsymbol{X}$ existiert. Für $\mathrm{H} \in \boldsymbol{X}$ ist die nur aus H bestehende Folge ein Beweis für H aus $\boldsymbol{X}$; folglich ist $\boldsymbol{X} \subseteq B(\boldsymbol{X})$. Ist $\mathrm{H}_\varkappa \in B(\boldsymbol{X})$ für $\varkappa = 1, \ldots, k$, φ eine k-gliedrige m-parametrige Schlußregel aus $\mathcal{K}$ und $\mathrm{H} = \varphi(\mathrm{H}_1, \ldots, \mathrm{H}_k; \mathrm{P}_1, \ldots, \mathrm{P}_m)$ für gewisse Parameter $\mathrm{P}_1, \ldots, \mathrm{P}_m$, so sei für $\varkappa = 1, \ldots, k$ die Folge $\mathrm{H}_1{}^\varkappa, \ldots, \mathrm{H}^\varkappa_{n_\varkappa}$ ein Beweis für $\mathrm{H}_\varkappa$ aus $\boldsymbol{X}$. Dann ist die Folge $\mathrm{H}_1{}^1, \ldots, \mathrm{H}^1_{n_1}, \mathrm{H}_1{}^2, \ldots, \mathrm{H}^2_{n_2}, \ldots, \mathrm{H}_1{}^k, \ldots, \mathrm{H}^k_{n_k}, \mathrm{H}$ ein Beweis für H aus $\boldsymbol{X}$. Folglich ist $B(\boldsymbol{X})$ $\mathcal{K}$-abgeschlossen, also $Bew_{\mathcal{K}}(\boldsymbol{X}) \subseteq B(\boldsymbol{X})$.

Zum Beweis der umgekehrten Inklusion zeigen wir durch Induktion über die Beweislänge: Das Endglied jedes Beweises aus $\boldsymbol{X}$ (d. h. jedes Element von $B(\boldsymbol{X})$)

liegt in jeder $\mathcal{K}$-abgeschlossenen Obermenge Y von X, folglich auch im Durchschnitt aller dieser Mengen, demnach in $Bew_{\mathcal{K}}(X)$. Es sei also Y eine beliebige $\mathcal{K}$-abgeschlossene Obermenge von X. Ist H ein $\mathcal{K}$-Beweis der Länge 1 aus X, so muß sein Endglied H ein logisches Axiom oder Element von X sein. Im ersten Fall ist $H \in Y$ wegen der $\mathcal{K}$-Abgeschlossenheit von Y, im zweiten Fall wegen $X \subseteqq Y$. Wir nehmen nun an, daß die Behauptung für die Endglieder aller $\mathcal{K}$-Beweise mit einer Länge $\leqq n$ gilt, und es sei $H_1, \ldots, H_n, H_{n+1}$ ein $\mathcal{K}$-Beweis der Länge $n+1$. Falls $H_{n+1} \in X$ ist, so ist $H_{n+1} \in Y$, andernfalls existieren eine Schlußregel $\varphi \in \mathcal{K}$, Parameter $P_1, \ldots, P_m$ und Indizes $i_1, \ldots, i_k \leqq n$, so daß

$$H_{n+1} = \varphi(H_{i_1}, \ldots, H_{i_k}; P_1, \ldots, P_m)$$

gilt. Für jeden der Ausdrücke $H_{i_\varkappa}$ $(\varkappa = 1, \ldots, k)$ ist dasjenige Anfangsstück des Beweises $H_1, \ldots, H_{n+1}$, welches mit $H_{i_\varkappa}$ endet, ein $\mathcal{K}$-Beweis für diesen Ausdruck, dessen Länge $\leqq n$ ist. Folglich ist nach Induktionsannahme $H_{i_\varkappa} \in Y$ für $\varkappa = 1, \ldots, k$. Wegen der $\mathcal{K}$-Abgeschlossenheit von Y ist daher auch $H_{n+1} \in Y$. Damit ist Satz 2 bewiesen.

Wir empfehlen dem Leser, mit Hilfe von Satz 2 den Satz 1 neu zu beweisen.

Satz 3. *Für $X \subseteqq S$ gilt $Bew_{\mathcal{K}}(X) \subseteqq Fl_S(X)$.* (Völlig analog gilt $Bew_{\mathcal{K}}(X) \subseteqq Fl_S^{\sigma}(X)$, falls $\mathcal{K}$ ein Beweiskalkül für die nichtelementare Sprache (S, σ) ist.)

Beweis. Wir stützen uns auf Satz 2 und zeigen durch Induktion über die Beweislänge: Für das Endglied H eines beliebigen Beweises aus X gilt $H \in Fl_S(X)$. Ist H ein Beweis der Länge 1, so ist $H \in X$ oder H ein logisches Axiom des Kalküls $\mathcal{K}$, in beiden Fällen ist daher $H \in Fl_S(X)$. Die Behauptung sei für alle Beweise der Länge $\leqq n$ gültig, und $H_1, \ldots, H_{n+1}$ sei ein Beweis aus X. Falls $H_{n+1} \in X$ oder H_{n+1} ein logisches Axiom ist, schließt man wie oben. Andernfalls existiert eine echte Schlußregel $\varphi \in \mathcal{K}$, so daß $H_{n+1} = \varphi(H_{i_1}, \ldots, H_{i_k}; P_1, \ldots, P_m)$ für gewisse Parameter $P_1, \ldots, P_m$ und gewisse im Beweis vorangehende Ausdrücke H_i gilt. Für diese gilt nach Induktionsannahme $H_i \in Fl_S(X)$. Daher folgt $H_{n+1} \in Fl_S(X)$ aus der Zulässigkeit von φ.

Bei der zum Beweis von Satz 3 angewandten Induktion über die Beweislänge wird man immer wieder darauf stoßen, daß die beim Induktionsanfang vorkommenden Fälle H logisches Axiom oder $H \in X$ beim Induktionsschritt noch einmal auftreten. Die Charakterisierung der Beweishülle $Bew_{\mathcal{K}}(X)$ als kleinste bezüglich aller Schlußregeln $\varphi \in \mathcal{K}$ abgeschlossene Obermenge von X liefert die theoretische Rechtfertigung für eine weniger schwerfällige Methode, Sätze der Form „*Für alle* H $\in Bew_{\mathcal{K}}(X)$ *gilt* …" zu beweisen. Wir zeigen:

a) Ist $H \in X$, so gilt …

b) Ist H logisches Axiom von $\mathcal{K}$, so gilt …

c) Entsteht H durch Anwendung einer echten Schlußregel $\varphi \in \mathcal{K}$ auf Ausdrücke $H_1, \ldots, H_k$, für die bereits … gilt, so gilt … auch für H.

Damit ist gezeigt: Die Menge $M \subseteq S$ aller Ausdrücke, für die ... gilt, ist eine bezüglich $\mathcal{K}$ abgeschlossene Obermenge von X. Demnach ist $Bew_{\mathcal{K}}(X) \subseteq M$, d. h., ... gilt für alle $\mathrm{H} \in Bew_{\mathcal{K}}(X)$.

Diese Beweismethode werden wir als *Induktion über die Beweiskompliziertheit* bezeichnen. Daß Definition 3 und Satz 2 dennoch für gewisse Eigenschaften der Beweishülle den einfachsten Beweis ermöglichen, zeigt z. B.

Satz 4 (Endlichkeitssatz für $Bew_{\mathcal{K}}$). *Ist* $\mathrm{H} \in Bew_{\mathcal{K}}(X)$, *so existiert eine endliche Teilmenge* E *von* X, *so daß* $\mathrm{H} \in Bew_{\mathcal{K}}(E)$ *gilt.*

Beweis. Ist $\mathrm{H} \in Bew_{\mathcal{K}}(X)$, so existiert nach Satz 2 ein Beweis von H aus X. Ist E die Menge aller in diesem Beweis vorkommenden Elemente von X, so ist derselbe Beweis offenbar ein Beweis von H aus E, folglich ist $\mathrm{H} \in Bew_{\mathcal{K}}(E)$ nach Satz 2.

Neben den hier betrachteten durch Systeme von Schlußregeln definierten Beweiskalkülen werden in der Literatur häufig Beweishüllen *Bew* betrachtet, die induktiv durch Festlegungen etwa folgender Art definiert sind:

Wenn $\mathrm{H}_1 \in Bew(X)$ und $\mathrm{H}_2 \in Bew(X)$, so sei auch $\mathrm{H}_1 \wedge \mathrm{H}_2 \in Bew(X)$. (3)

Wenn $\mathrm{H}_2 \in Bew(X \cup \{\mathrm{H}_1\})$ und H_1 abgeschlossen, so sei $\mathrm{H}_1 \rightarrow \mathrm{H}_2 \in Bew(X)$. (4)

Regeln der ersten Art lassen sich ohne weiteres in Schlußregeln umformen. Zum Beispiel kann die Eigenschaft (3) für einen durch einen Beweiskalkül $\mathcal{K}$ definierten Hüllenoperator $Bew_{\mathcal{K}}$ dadurch erreicht werden, daß man in $\mathcal{K}$ die Schlußregel $\varphi_8(\mathrm{H}_1, \mathrm{H}_2) := \mathrm{H}_1 \wedge \mathrm{H}_2$ aufnimmt. Dagegen ist die Umwandlung von Definitionsregeln der Form (4) in Schlußregeln nicht ohne weiteres möglich. Entscheidend ist jedoch, daß auch derartige induktiv definierte Beweishüllenoperatoren die in den Sätzen 1, 3 und 4 formulierten Eigenschaften haben. Bei einer abstrakteren Auffassung des Beweisbarkeitsbegriffs könnte man daher diese Eigenschaften als axiomatische Charakterisierung des Begriffs Beweiskalkül benutzen.

Ist *Bew* ein beliebiger Beweishüllenoperator für die Sprache S (bzw. die nichtelementare Sprache (S, σ)), so kann man wissenschaftstheoretische Begriffe wie Widerspruchsfreiheit, Unabhängigkeit und Vollständigkeit auf syntaktische Weise definieren, indem man in den Definitionen der entsprechenden semantischen Begriffe (vgl. 4.3.) den Folgerungsoperator Fl (bzw. Fl^{σ}) durch den Beweishüllenoperator *Bew* ersetzt. Dies sei zunächst am Beispiel der Widerspruchsfreiheit erläutert.

Ein Axiomensystem X in einer formalisierten Sprache S ist nach 4.3. genau dann semantisch widerspruchsfrei, wenn eine der drei folgenden äquivalenten Bedingungen erfüllt ist:

X besitzt ein Modell. (5)

$Fl(X) \subset S$. (6)

Für $\mathrm{H} \in S$ gilt: $\mathrm{H} \notin Fl(X)$ oder $\neg\, \mathrm{H} \notin Fl(X)$. (7)

Offenbar besitzt (5) kein unmittelbares syntaktisches Analogon, während (6) und (7) zu folgenden Definitionen führen:

Definition 4.

a) $\boldsymbol{X} \subseteqq \boldsymbol{S}$ heißt *syntaktisch widerspruchsfrei bezüglich Bew*, wenn $Bew(\boldsymbol{X}) \subset \boldsymbol{S}$ gilt.

b) $\boldsymbol{X} \subseteqq \boldsymbol{S}$ heißt *formal widerspruchsfrei bezüglich Bew*, wenn für $\mathrm{H} \in \boldsymbol{S}$ gilt: $\mathrm{H} \notin Bew(\boldsymbol{X})$ oder $\neg\, \mathrm{H} \notin Bew(\boldsymbol{X})$.

Satz 5. *Für beliebige Beweiskalküle Bew für* $\boldsymbol{S}$ *gilt:*

a) *Wenn* $\boldsymbol{X}$ *semantisch widerspruchsfrei ist, so ist* $\boldsymbol{X}$ *formal widerspruchsfrei.*

b) *Wenn* $\boldsymbol{X}$ *formal widerspruchsfrei ist, so ist* $\boldsymbol{X}$ *syntaktisch widerspruchsfrei.*

Die Umkehrungen von a) und b) gelten für beliebige Beweiskalküle im allgemeinen nicht. Jedoch läßt sich die Umkehrung von b) leicht durch Aufnahme einer geeigneten Schlußregel erreichen.

Beweis. a) Da stets $Bew(\boldsymbol{X}) \subseteqq Fl(\boldsymbol{X})$ sein muß, folgt: Wenn für einen gewissen Ausdruck $\mathrm{H} \in \boldsymbol{S}$ gleichzeitig $\mathrm{H} \in Bew(\boldsymbol{X})$ und $\neg\, \mathrm{H} \in Bew(\boldsymbol{X})$ gilt, so ist auch $\mathrm{H}, \neg\, \mathrm{H} \in Fl(\boldsymbol{X})$, d. h., eine nicht formal widerspruchsfreie Menge $\boldsymbol{X}$ kann erst recht nicht semantisch widerspruchsfrei sein.

b) Ist $\boldsymbol{X}$ formal widerspruchsfrei und $\mathrm{H} \in \boldsymbol{S}$ ein beliebiger Ausdruck, dann gehört wenigstens einer der beiden Ausdrücke $\mathrm{H}, \neg\, \mathrm{H}$ nicht zu $Bew(\boldsymbol{X})$; folglich ist $Bew(\boldsymbol{X}) \subset \boldsymbol{S}$. Ist der Beweishüllenoperator *Bew* z. B. mit Hilfe eines Kalküls definiert, der die Schlußregel $\varphi_9(\mathrm{H}, \neg\, \mathrm{H}; \mathrm{H}_0) := \mathrm{H}_0$ enthält, die es gestattet, aus zwei Prämissen der Form $\mathrm{H}, \neg\, \mathrm{H}$ auf jeden beliebigen (hier als Parameter auftretenden) Ausdruck $\mathrm{H}_0 \in \boldsymbol{S}$ zu schließen, so gilt für *Bew* offenbar die Umkehrung von b).

Zwei einfache Beispiele mögen zeigen, daß die Umkehrungen von a) und b) für beliebige Beweishüllenoperatoren nicht zu erwarten sind. Ist *Bew* durch einen Kalkül definiert, der nur die bereits erwähnte Schlußregel φ_8 $\big(\varphi_8(\mathrm{H}_1, \mathrm{H}_2) = \mathrm{H}_1 \wedge \mathrm{H}_2\big)$ enthält, so kann man auch aus einer semantisch widerspruchsvollen Menge $\boldsymbol{X}$ nicht auf einen formalen Widerspruch $\mathrm{H}, \neg\, \mathrm{H}$ schließen, falls $\boldsymbol{X}$ selbst keine mit dem Negationszeichen beginnenden Ausdrücke enthält. Enthält andererseits $\boldsymbol{X}$ genau zwei Ausdrücke der Form $\mathrm{H}_0, \neg\, \mathrm{H}_0$, so ist $\boldsymbol{X}$ schon formal widerspruchsvoll, jedoch bezüglich des genannten Kalküls syntaktisch widerspruchsfrei, da man mittels φ_8 aus $\boldsymbol{X}$ niemals den Ausdruck $\neg\, (\mathrm{H}_0 \wedge \neg\, \mathrm{H}_0)$ ableiten kann.

Analog zu Definition 4 erhält man aus den entsprechenden semantischen Definitionen:

$\mathrm{H} \in \boldsymbol{S}$ heißt *syntaktisch unabhängig* (*bezüglich Bew*) *von* $\boldsymbol{X} \subseteqq \boldsymbol{S}$, wenn $\mathrm{H} \notin Bew(\boldsymbol{X})$ gilt.

$\boldsymbol{X} \subseteqq \boldsymbol{S}$ heißt *syntaktisch vollständig* (*bezüglich Bew*), wenn für $\mathrm{H} \in \overline{\boldsymbol{S}}$ stets $\mathrm{H} \in Bew(\boldsymbol{X})$ oder $\neg\, \mathrm{H} \in Bew(\boldsymbol{X})$ gilt.

Der Leser sei angeregt, analog zu Satz 5 die allgemein gültigen Beziehungen zwischen diesen und den entsprechenden semantischen Begriffen und auch zwischen sämtlichen syntaktischen Begriffen untereinander zu ermitteln und die nicht allgemein gültigen Beziehungen durch Gegenbeispiele zu belegen.

Satz 6. *Zu jedem für eine (eventuell nichtelementare) Sprache* S *(bzw.* (S, σ)*) zulässigen Kalkül* $\mathcal{K}$ *kann man einen Kalkül* $\mathcal{K}'$ *konstruieren, dessen einzige echte Schlußregeln die Abtrennungsregel* φ_1, *die Regel* φ_2 *der gebundenen Umbenennung und die durch*

$$\varphi_{10}\big(\mathrm{H}(\mathrm{x}); \mathrm{x}\big) := \bigwedge \mathrm{x}\mathrm{H}(\mathrm{x}), \textit{ falls } \mathrm{x} \textit{ in } \mathrm{H}(\mathrm{x}) \textit{ vollfrei vorkommt,}$$

definierte Regel φ_{10} *sind, so daß für alle* $X \subseteq S$

$$Bew_{\mathcal{K}}(X) \subseteq Bew_{\mathcal{K}'}(X) = Bew_{\mathcal{K}''}(X \cup A) \subseteq Fl^{\sigma}(X)$$

gilt, wobei $\mathcal{K}'' = \{\varphi_1, \varphi_2, \varphi_{10}\}$ *und* A *das System aller logischen Axiome von* $\mathcal{K}'$ *ist. Insbesondere kann man, da* $\varphi_1, \varphi_2, \varphi_{10}$ *elementare Schlußregeln sind, den Gebrauch nichtelementarer echter Schlußregeln völlig vermeiden bzw. sie durch nichtelementare logische Axiome ersetzen.*

Beweis. Für jede echte Schlußregel $\varphi \in \mathcal{K}$ (k-gliedrig, m-parametrig) sei

$$\varphi'(;\mathrm{H}_1, \ldots, \mathrm{H}_k, \mathrm{P}_1, \ldots, \mathrm{P}_m) := \mathrm{H}_1' \to \big(\mathrm{H}_2' \to \ldots \to (\mathrm{H}_k' \to \mathrm{H}_0) \ldots\big), \tag{8}$$

falls $\mathrm{H}_0 = \varphi(\mathrm{H}_1, \ldots, \mathrm{H}_k; \mathrm{P}_1, \ldots, \mathrm{P}_m)$ existiert, wobei für $i = 1, \ldots, k$ der Ausdruck H_i' aus H_i dadurch entsteht, daß man durch gebundene Umbenennungen alle freien Variablen vollfrei macht und anschließend zur Generalisierten übergeht. Es ist also insbesondere $\mathrm{H}_i' = \mathrm{H}_i$, falls H_i abgeschlossen ist. Somit ist jeder k-gliedrigen m-parametrigen Schlußregel $\varphi \in \mathcal{K}$ ein $(k + m)$-parametriges Axiomenschema φ' zugeordnet. Für 0-gliedrige Schlußregeln $\varphi \in \mathcal{K}$ sei $\varphi' = \varphi$, und es sei

$$\mathcal{K}' = \{\varphi' : \varphi \in \mathcal{K} \wedge \varphi \neq \varphi_1, \varphi_2, \varphi_{10}\} \cup \{\varphi_1, \varphi_2, \varphi_{10}\}.$$

Die Voraussetzung, daß alle Regeln $\varphi \in \mathcal{K}$ zulässig für (S, σ) sind, bedeutet

$$\mathrm{H}_0 \in Fl^{\sigma}(\{\mathrm{H}_1, \ldots, \mathrm{H}_k\}).$$

Wir haben daraus zunächst zu folgern, daß die Axiomenschemata φ' Ausdrücke aus $\mathrm{Fl}^{\sigma}(\emptyset)$ ergeben, daß $\mathcal{K}'$ mithin ebenfalls für (S, σ) zulässig ist. Es sei $\omega \in \sigma$ eine zulässige Interpretation von S. Falls ω Modell für $\mathrm{H}_1, \ldots, \mathrm{H}_k$ ist, ist ω auch Modell für H_0 und daher der Ausdruck (8) bei ω allgemeingültig. Ist dagegen einer der Ausdrücke H_i ($i = 1, \ldots, k$) bei ω nicht allgemeingültig, so ist der entsprechende Ausdruck H_i' (da abgeschlossen) bei jeder Belegung f bezüglich ω falsch, mithin ist (8) wiederum bei ω allgemeingültig. Wie man sieht, wäre die Zulässigkeit der Axiomenschemata φ' nicht allgemein gegeben, wenn man in (8) statt H_i' die Ausdrücke H_i einsetzen würde. Zum Beispiel würde sich aus der zulässigen elementaren Schlußregel φ_{10} das nicht allgemeingültige Axiomenschema $\mathrm{H}(\mathrm{x}) \to \bigwedge \mathrm{x}\mathrm{H}(\mathrm{x})$ ergeben.

Wir zeigen nun durch Induktion über die Beweiskompliziertheit: Ist $H_0 \in Bew_{\mathcal{K}}(X)$, so ist $H_0 \in Bew_{\mathcal{K}'}(X)$. Dabei ist nur der Fall interessant, daß H_0 durch Anwendung einer echten Schlußregel $\varphi \in \mathcal{K}$ aus Ausdrücken $H_1, \ldots, H_k$ erhalten wird, für die schon $H_i \in Bew_{\mathcal{K}'}(X)$ gilt. Die Herleitung von H_0 bezüglich $\mathcal{K}'$ kann offenbar dadurch geleistet werden, daß man zunächst mittels φ_2 und φ_{10} aus den Prämissen $H_1, \ldots, H_k$ die entsprechenden Ausdrücke $H_1', \ldots, H_k'$ gewinnt und dann auf diese und das Axiom (8) k-mal die Abtrennungsregel anwendet.

6.2. Aussagenlogisches Schließen

Eine Abbildung f, die jeder Aussagenvariablen p_i ($i = 0, 1, 2, \ldots$) einen Ausdruck $H = f(p_i)$ einer gewissen formalisierten Sprache S zuordnet, heißt eine *Belegung bezüglich* S. Ist A ein aussagenlogischer Ausdruck, so bezeichne $f(A)$ denjenigen Ausdruck der Sprache S, den man erhält, wenn man jede in A vorkommende Variable p_i durch den Ausdruck $f(p_i)$ ersetzt. Offenbar kann man $f(A)$ auch induktiv durch die Regeln

$$f(\neg A) = \neg f(A),$$

$$f(A_1 \circ A_2) = f(A_1) \circ f(A_2) \quad \textit{für } \circ = \wedge, \vee, \rightarrow, \leftrightarrow$$

definieren und berechnen, wenn $f(p_i)$ für die vorkommende Variablen vorgegeben ist.

Eine für eine Sprache S (bzw. (S, σ)) zulässige k-gliedrige und m-parametrige Schlußregel φ heißt eine *aussagenlogische Schlußregel*, wenn zu beliebigen $H_0, H_1, \ldots, H_k \in S$ und $P_1, \ldots, P_m$ mit

$$\varphi(H_1, \ldots, H_k; P_1, \ldots, P_m) = H_0$$

aussagenlogische Ausdrücke $A_0, A_1, \ldots, A_k$ und eine Belegung f bezüglich S existieren, so daß der aussagenlogische Ausdruck $A_1 \rightarrow (A_2 \rightarrow \cdots \rightarrow (A_k \rightarrow A_0) \ldots)$ allgemeingültig und

$$f(A_i) = H_i \quad \text{für } i = 0, 1, \ldots, k$$

ist, mit anderen Worten, wenn H_0 aus $H_1, \ldots, H_k$ bereits auf Grund der aussagenlogischen Struktur der beteiligten Ausdrücke folgt.

Von den im vorigen Abschnitt definierten Schlußregeln sind φ_1, φ_1^m ($m \geq 1$), φ_8 und φ_9 aussagenlogische Schlußregeln, die übrigen nicht. Für den Spezialfall 0-gliedriger Schlußregeln ergibt sich: Ein Schema logischer Axiome (bzw. ein einzelnes logisches Axiom) heißt *aussagenlogisch*, wenn jedes Axiom H dieses Schemas die Form $f(A_0)$ hat, wobei A_0 ein allgemeingültiger Ausdruck der Aussagenlogik und f eine Belegung der Aussagenvariablen bezüglich S ist.

Zur Motivation dieser Definition sei bemerkt, daß zum Beispiel die Axiomenschemata φ_1^m $(m \geqq 1)$ nicht unter den Begriff aussagenlogischer Schlußregel fallen würden, wenn man, wie die übrigen Beispiele nahelegen, eine aussagenlogische Schlußregel φ als eine solche definiert, zu der feste Ausdrücke $A_0, A_1, \ldots, A_k$ existieren, so daß das oben Verlangte immer dann für eine geeignete Belegung f gilt, wenn $\varphi(H_1, \ldots, H_k; P_1, \ldots, P_m) = H_0$ ist.

Ein Kalkül *Bew* (speziell $Bew_{\mathcal{K}}$) heißt *aussagenlogisch vollständig*, wenn für jeden aussagenlogisch allgemeingültigen Ausdruck der Form $A_1 \to \big(A_2 \to \cdots \to (A_k \to A_0) \cdots\big)$ und jede Belegung f

$$f(A_0) \in Bew\big(\{f(A_1), \ldots, f(A_k)\}\big)$$

gilt (also insbesondere $f(A_0) \in Bew(\emptyset)$ für jeden aussagenlogisch allgemeingültigen Ausdruck A_0 und jede Belegung f).

Satz 1. *Der Kalkül* $\mathcal{K}_0 := \{\varphi_1^m\colon\ m \geqq 1\} \cup \{\varphi_1\}$ *ist aussagenlogisch vollständig.*

Beweis. Ist $A = A_1 \to \big(A_2 \to \cdots \to (A_k \to A_0) \cdots\big)$ aussagenlogisch allgemeingültig und m der größte Index darin vorkommender Variablen p_i, so wird für jede Belegung f der Ausdruck $f(A)$ in der Form

$$\varphi_1^m\big(;A, f(p_1), \ldots, f(p_m)\big)$$

als logisches Axiom erhalten. Daher ist $f(A) \in Bew_{\mathcal{K}_0}\big(\{f(A_1), \ldots, f(A_k)\}\big)$. Wendet man auf $f(A)$ und die Ausdrücke $f(A_i)$ $(i = 1, \ldots, k)$ k-mal die Abtrennungsregel φ_1 an, so erhält man

$$f(A_0) \in Bew_{\mathcal{K}_0}\big(\{f(A_1), \ldots, f(A_k)\}\big),$$

was zu beweisen war.

Im Kalkül $\mathcal{K}_0$ ist die Abtrennungsregel die einzige echte Schlußregel, während die Gesamtheit aller aussagenlogisch allgemeingültigen Ausdrücke als Axiomensystem dient. Das letzte ist wegen der Entscheidbarkeit (im anschaulichen Sinne) der aussagenlogischen Allgemeingültigkeit zulässig und für die praktische Handhabung des aussagenlogischen Schließens relativ unproblematisch. In der Aussagenlogik beschäftigt man sich jedoch ausführlich mit den Möglichkeiten, aussagenlogisch vollständige endliche Kalküle aufzusuchen, d. h. solche, die aus nur endlich vielen Schlußregeln und jeweils durch einen einzelnen aussagenlogisch allgemeingültigen Ausdruck A gegebenen Axiomenschemata bestehen, wobei meist die Abtrennungsregel als einzige echte Schlußregel beibehalten wird. Einzelheiten hierüber findet man in weiterführenden Lehrbüchern (siehe insbesondere [2]). Wir teilen hier nur ohne Beweis ein gebräuchliches System mit, wobei wir lediglich die den einzelnen Schemata zugrunde liegenden allgemeingültigen aussagenlogischen Ausdrücke angeben:

1. $p \to (q \to p)$
2. $\big((p \to q) \to p\big) \to p$

3. $(p \to q) \to \big((q \to r) \to (p \to r)\big)$
4. $p \wedge q \to p$
5. $p \wedge q \to q$
6. $(p \to q) \to \big((p \to r) \to (p \to q \wedge r)\big)$
7. $p \to p \vee q$
8. $q \to p \vee q$
9. $(p \to r) \to \big((q \to r) \to (p \vee q \to r)\big)$
10. $(p \leftrightarrow q) \to (p \to q)$
11. $(p \leftrightarrow q) \to (q \to p)$
12. $(p \to q) \to \big((q \to p) \to (p \leftrightarrow q)\big)$
13. $(p \to q) \to (\neg q \to \neg p)$
14. $p \to \neg\neg p$
15. $\neg\neg p \to p$

Dabei entspricht z. B. dem Ausdruck 1. die 0-gliedrige 2-parametrige Schlußregel (bzw. das 2-parametrige Axiomenschema)

$$\varphi^{(1)}(;H_1, H_2) := H_1 \to (H_2 \to H_1) \qquad (H_1, H_2 \in S \text{ beliebig}).$$

Zu den übrigen Ausdrücken 2. bis 15. gehören analog definierte Axiomenschemata $\varphi^{(2)}$ bis $\varphi^{(15)}$. Als einzige echte Schlußregel dient hier wieder die Abtrennungsregel φ_1.

Als Beispiel eines korrekten Beweises bezüglich des aussagenlogisch vollständigen Kalküls $\mathcal{K} = \{\varphi_1, \varphi^{(1)}, \ldots, \varphi^{(15)}\}$ zeigen wir:

$$\big(H' \to (H' \to H'')\big) \to (H' \to H'') \in Bew_{\mathcal{K}}(\emptyset) \qquad \text{für beliebige } H', H''.$$

Gemäß der allgemeinen Definition des Beweisbegriffs ist hierzu eine Folge $H_1, H_2, \ldots, H_n$ von Ausdrücken so anzugeben, daß H_n der zu beweisende Ausdruck und jedes Glied der Folge ein logisches Axiom ist oder durch Anwendung von φ_1 auf vorangehende Folgenglieder erhalten werden kann. Diese Bedingungen erfüllt z. B. die Folge

$$\begin{aligned} H_1 &= \varphi^{(3)}(;H', H' \to H'', H'') \\ &= \underbrace{\big(H' \to (H' \to H'')\big)}_{H^{(1)}} \to \underbrace{\Big(\big((H' \to H'') \to H''\big) \to (H' \to H'')\Big)}_{H^{(2)}}, \end{aligned}$$

$$\begin{aligned} H_2 &= \varphi^{(2)}(;H' \to H'', H'') \\ &= \underbrace{\Big(\big((H' \to H'') \to H''\big) \to (H' \to H'')\Big)}_{H^{(2)}} \to \underbrace{(H' \to H'')}_{H^{(3)}}, \end{aligned}$$

$$
\begin{aligned}
H_3 &= \varphi^{(3)}(H^{(1)}, H^{(2)}, H^{(3)}) \\
&= (H^{(1)} \to H^{(2)}) \to \big((H^{(2)} \to H^{(3)}) \to (H^{(1)} \to H^{(3)})\big), \\
H_4 &= \varphi_1(H_3, H_1) = H_2 \to (H^{(1)} \to H^{(3)}), \\
H_5 &= \varphi_1(H_4, H_2) = H^{(1)} \to H^{(3)} \quad \text{ist der zu beweisende Ausdruck.}
\end{aligned}
$$

Offenbar erhält man durch Vertauschung einiger Folgeglieder weitere Beweise für H_5, z. B. H_2, H_1, H_3, H_4, H_5 oder H_3, H_1, H_4, H_2, H_5. Ferner sei bemerkt, daß wir, genau betrachtet, nicht einen einzelnen Beweis für einen bestimmten Ausdruck, sondern ein Schema von Beweisen für alle Ausdrücke der Form

$$\big(H' \to (H' \to H'')\big) \to (H' \to H'')$$

aufgeschrieben haben. Ersetzt man in diesem Schema die Variablen H', H'' für beliebige Ausdrücke einer beliebigen prädikatenlogischen Sprache wieder durch Variablen der Form p_i bzw. $p, q, r, \ldots$, so erhält man einen Beweis für die Ableitbarkeit des Ausdrucks

$$\big(p \to (p \to q)\big) \to (p \to q)$$

aus den Axiomen 1. bis 15. mittels Abtrennungsregel und Einsetzung beliebiger aussagenlogischer Ausdrücke für die Variablen in den Axiomen. Demnach sind solche aussagenlogischen Ableitungen im Prinzip nichts anderes als Schemata von Beweisen für die aussagenlogische Allgemeingültigkeit ganzer Klassen von Ausdrücken beliebiger prädikatenlogischer Sprachen. Abschließend wollen wir $H \to H \in Bew_{\mathcal{X}}(\emptyset)$ für beliebige Ausdrücke H zeigen:

Als Spezialfall des oben Bewiesenen ergibt sich für $H' = H'' = H$

$$\big(H \to (H \to H)\big) \to (H \to H) \in Bew_{\mathcal{X}}(\emptyset),$$

ferner

$$\varphi^{(1)}(;H, H) = H \to (H \to H) \in Bew_{\mathcal{X}}(\emptyset),$$

demnach die Behauptung durch Anwendung der Abtrennungsregel. Um aus diesem „Metabeweis" einen Beweis im Sinne der Definition 3 aus 6.1. zu erhalten, müßte man zunächst die bereits erhaltene Folge H_1, H_2, H_3, H_4, H_5 für den Fall $H' = H'' = H$ neu aufschreiben und dann durch

$$H_6 = \varphi^{(1)}(;H, H), \qquad H_7 = \varphi_1(H_5, H_6)$$

fortsetzen. Auch hier bleibt die Beweiseigenschaft bei gewissen Vertauschungen der Beweisglieder wieder erhalten. Der Leser ermittle alle Permutationen der Folge $H_1, \ldots, H_7$, die Beweise für H_7 sind.

6.3. Der Hauptsatz der mathematischen Logik und seine Folgerungen

Definition. Ein Beweiskalkül $\mathcal{K}$ für eine Sprache S (bzw. (S, σ)) heißt *vollständig*, wenn

$$Bew_{\mathcal{K}}(X) = Fl^{(\sigma)}(X) \quad \text{für } X \subseteq S$$

gilt.

Da die Zulässigkeit eines Kalküls für eine Sprache (S, σ) durch $Bew_{\mathcal{K}}(X) \subseteq Fl^{\sigma}(X)$ charakterisiert ist, ist nur die umgekehrte Inklusion von Bedeutung. Demnach ist ein Kalkül genau dann vollständig, wenn für jeden Ausdruck $\mathrm{H} \in Fl^{\sigma}(X)$ ein $\mathcal{K}$-Beweis für H aus X existiert, mit anderen Worten, wenn das inhaltliche σ-Folgern in S gleichwertig durch ein formales Beweisen mit Hilfe der Schlußregeln und logischen Axiome von $\mathcal{K}$ ersetzt werden kann. Es gilt der

Hauptsatz der mathematischen Logik (Satz von Gödel-Mal'cev). *Zu jeder elementaren Sprache S kann man einen endlichen vollständigen Kalkül $\mathcal{K}$ angeben. (Da ein solcher Kalkül im wesentlichen für alle elementaren Sprachen auf einheitliche Weise definierbar ist, kann man auch sagen: Man kann einen endlichen Kalkül $\mathcal{K}$ so angeben, daß $Bew_{\mathcal{K}}(X) = Fl(X)$ für jede in einer elementaren Sprache formulierte Satzmenge X ist.)*

Der Beweis dieses Satzes, der seinem wesentlichen Inhalt nach zuerst von Gödel ausgesprochen und bewiesen wurde,[1] ist trotz vieler im Laufe der letzten Jahrzehnte daran vorgenommener Vereinfachungen bzw. neuer Beweiswege immer noch recht kompliziert.[2] Einen Beweisweg werden wir im nächsten Abschnitt skizzieren und uns in diesem Abschnitt ausschließlich mit verschiedenen Folgerungen und Anwendungen des Hauptsatzes und deren gegenseitigen Beziehungen beschäftigen.

1. Als unmittelbarer Spezialfall des Hauptsatzes ergibt sich für $X = \emptyset$ der

Vollständigkeitssatz von Gödel. *Man kann einen endlichen Kalkül $\mathcal{K}$ so konstruieren, daß*

$$Bew_{\mathcal{K}}(\emptyset) = Fl_S(\emptyset)$$

für jede elementare Sprache S gilt, d. h., die Menge ag_S der allgemeingültigen Ausdrücke einer elementaren Sprache S ist durch fortlaufende Anwendung der Schlußregeln von $\mathcal{K}$ auf die logischen Axiome von $\mathcal{K}$ „aufzählbar".

---

[1]) K. Gödel, Die Vollständigkeit der Axiome des logischen Funktionenkalküls, Monatsh. Math. Phys. 37 (1930), 349–360; nachgedruckt in [4].

[2]) Siehe hierzu N. J. Zygmunt, A survey of the methods of proof of the Gödel-Malcev's completeness theorem, in [38].

2. *Enthält $\mathcal{K}$ die Abtrennungsregel, so folgt aus* $Bew_{\mathcal{K}}(\emptyset) = Fl(\emptyset)$ *für endliche Mengen* X *von abgeschlossenen Ausdrücken bereits* $Bew_{\mathcal{K}}(X) = Fl(X)$. *Enthält* $\mathcal{K}$ *außerdem* φ_2 *und* φ_{10} (vgl. 6.1., Satz 6), *so gilt dies für beliebige endliche Mengen* X.

Beweis. Es sei $X = \{H_1, \ldots, H_n\}$, $\overline{H}_i$ ein durch Anwendungen von φ_2 und φ_{10} aus H_i entstehender abgeschlossener Ausdruck. Ist nun $H \in Fl(X)$, so ist auch $H \in Fl(\{\overline{H}_1, \ldots, \overline{H}_n\})$, folglich gilt (wegen der Gültigkeit des Deduktionstheorems für das Folgern, vgl. S. 133)

$$\overline{H}_1 \to \big(\overline{H}_2 \to (\ldots \to (\overline{H}_n \to H) \ldots)\big) \in Fl(\emptyset) = Bew_{\mathcal{K}}(\emptyset) \subseteq Bew_{\mathcal{K}}(X).$$

Falls nicht alle H_i selbst abgeschlossen sind, ergibt sich $\overline{H}_1, \ldots, \overline{H}_n \in Bew_{\mathcal{K}}(X)$ mittels φ_2 und φ_{10}, woraus $H \in Bew_{\mathcal{K}}(X)$ durch n-malige Abtrennung beweisbar ist.

3. Da für jeden Beweisoperator $Bew_{\mathcal{K}}$ trivialerweise der Endlichkeitssatz (6.1., Satz 4) gilt, folgt aus dem Hauptsatz sofort der

Endlichkeitssatz für das Folgern. *Ist* $H \in Fl(X)$, *so existiert eine endliche Teilmenge* E *von* X *mit* $H \in Fl(E)$.

4. Dem Endlichkeitssatz für das Folgern ist äquivalent der

Satz von Mal'cev[1] (Endlichkeitssatz für Modelle). *Besitzt jede endliche Teilmenge eines Axiomensystems* X *ein Modell, so besitzt auch* X *ein Modell.*

Wir beweisen die Äquivalenz der beiden Endlichkeitssätze. Zunächst sei der Endlichkeitssatz für das Folgern vorausgesetzt, und es sei X eine solche Teilmenge einer Sprache S, daß jede endliche Teilmenge von X ein Modell besitzt. Aus der Annahme, X besitzt kein Modell, d. h. ist semantisch widerspruchsvoll, ergibt sich für einen beliebigen Ausdruck $H \in S$ sowohl $H \in Fl(X)$ als auch $\neg H \in Fl(X)$. Nach Voraussetzung gibt es daher endliche Teilmengen E_1, E_2 von X mit $H \in Fl(E_1)$, $\neg H \in Fl(E_2)$, daher ist $H, \neg H \in Fl(E_1 \cup E_2)$. Dies ist aber ein Widerspruch zur Voraussetzung, da $E_1 \cup E_2$ als endliche Teilmenge von X ein Modell besitzen muß. Wir setzen nun die Gültigkeit des Satzes von Mal'cev voraus und betrachten eine Satzmenge X und einen Ausdruck $H \in Fl(X)$. Die Annahme, daß $H \notin Fl(E)$ für alle endlichen Teilmengen E von X gilt, ist gleichbedeutend damit, daß für alle derartigen Mengen E gilt: Es gibt ein Modell für $E \cup \{H'\}$, wobei H' dadurch aus H entsteht, daß man zunächst alle freien Variablen vollfrei macht, dann zur Generalisierten übergeht und diese negiert. Da somit jede endliche Teilmenge von $X \cup \{H'\}$ ein Modell besitzt, besitzt nach dem Satz von Mal'cev die Menge $X \cup \{H'\}$ ein Modell, d. h., X besitzt ein Modell, in dem H nicht allgemeingültig ist, im Widerspruch zur Voraussetzung $H \in Fl(X)$.

[1]) Мальцев, А. И., Untersuchungen aus dem Gebiete der mathematischen Logik, Матем. сб. 1 (43) (1936), 323–336.

5. *Enthält $\mathcal{K}$ die Regeln $\varphi_1, \varphi_2, \varphi_{10}$, so folgt aus dem Satz von* MAL'CEV *und der Voraussetzung $Bew_{\mathcal{K}}(\emptyset) = Fl(\emptyset)$ die Vollständigkeit von $\mathcal{K}$.*

Beweis. Zu $\mathrm{H} \in Fl(\boldsymbol{X})$ existiert nach dem Satz von MAL'CEV eine endliche Teilmenge $\boldsymbol{E} = \{\mathrm{H}_1, \ldots, \mathrm{H}_n\}$ von $\boldsymbol{X}$ mit $\mathrm{H} \in Fl(\{\mathrm{H}_1, \ldots, \mathrm{H}_n\})$, woraus nach 2.

$$\mathrm{H} \in Bew_{\mathcal{K}}(\{\mathrm{H}_1, \ldots, \mathrm{H}_n\}) \subseteq Bew_{\mathcal{K}}(\boldsymbol{X})$$

folgt.

Der Satz von MAL'CEV, von diesem bei algebraischen Untersuchungen entdeckt und unabhängig vom Gödelschen Vollständigkeitssatz mit algebraischen Methoden bewiesen, hat in den letzten Jahrzehnten viele Anwendungen in der Mathematik, insbesondere in der Algebra gefunden. Wir geben nachfolgend einige einfache Beispiele für solche Anwendungen an.

6. *Zu jeder Primzahl p gibt es einen unendlichen Körper der Charakteristik p.*

Aus der Theorie der endlichen Körper (Galois-Felder) ist bekannt, daß es bei vorgegebener Primzahlcharakteristik p genau dann einen Körper der Charakteristik p mit n Elementen gibt, wenn n eine Potenz p^m von p ist, und daß es zu vorgegebenem p und m bis auf Isomorphie genau einen Körper $GF(p^m)$ der Charakteristik p mit p^m Elementen gibt. Die Theorie der Galois-Felder sagt jedoch nichts über die Existenz unendlicher Körper mit Primzahlcharakteristik aus, und alle „bekannten" unendlichen Körper haben die Charakteristik 0.

Es sei $\boldsymbol{S}$ eine elementare einsortige Sprache und $\boldsymbol{X}_p$ ein in dieser Sprache formuliertes endliches Axiomensystem der Körpertheorie, dem wir noch als Axiom über die Charakteristik den Ausdruck $\underbrace{\mathrm{x} + \mathrm{x} + \cdots + \mathrm{x}}_{p} = 0$ hinzufügen. Ferner sei H_n für $n = 1, 2, 3, \ldots$ der Ausdruck $\bigvee \mathrm{x}\, \mathrm{x} \underset{n}{=} \mathrm{x}$ (d. h., es gibt mindestens n Dinge) und $\boldsymbol{X} = \boldsymbol{X}_p \cup \{H_n : n = 1, 2, 3, \ldots\}$. Da jede endliche Teilmenge von $\boldsymbol{X}$ höchstens endlich viele der Aussagen H_n über die Mindestzahl enthält, besitzt sie ein Modell in Gestalt eines Galois-Feldes $GF(p^m)$ mit hinreichend großem m. Daher besitzt nach dem Satz von MAL'CEV auch die gesamte Menge $\boldsymbol{X}$ ein Modell. Dieses muß ein (wegen der Gültigkeit aller H_n in ihm) unendlicher Körper der Charakteristik p sein.

7. *Es gibt einen nichtarchimedisch geordneten Körper.*

Wir erweitern die in 6. betrachtete Sprache der Körpertheorie durch ein zweistelliges Relationssymbol „$<$" und ein weiteres Konstantensymbol c und formulieren in dieser Sprache $\boldsymbol{S}'$ ein endliches Axiomensystem $\boldsymbol{X}_0$ für geordnete Körper, in dessen Axiomen die Konstante c zunächst nicht vorkommt, so daß eine Interpretation der Sprache $\boldsymbol{S}'$ genau dann ein Modell von $\boldsymbol{X}_0$ ist, wenn sie einen geordneten Körper definiert, in dem die Konstante c durch ein beliebiges Körperelement interpretiert ist. Für $n = 1, 2, 3, \ldots$ sei nun H_n der Ausdruck $\underbrace{1 + 1 + 1 + \cdots + 1}_{n} < \mathrm{c}$, und es sei

$$X = X_0 \cup \{\mathrm{H}_n : n = 1, 2, 3, \ldots\}.$$

Jede endliche Teilmenge von X besitzt ein Modell in Gestalt etwa des Körpers der rationalen Zahlen, wobei die Konstante c durch eine hinreichend große natürliche Zahl zu interpretieren ist, damit die jeweils endlich vielen Ausdrücke H_n erfüllt werden. Folglich besitzt die Gesamtmenge X ein Modell, und dies muß ein geordneter Körper sein, in dem die Konstante c durch ein Element interpretiert ist, das größer als alle endlichen Vielfachen des Einselements ist. Mithin muß dieser Körper nichtarchimedisch geordnet sein.

8. Mit einer ähnlichen Überlegung wie in 7. können wir zeigen, daß das archimedische Axiom für geordnete Körper in keiner elementaren Sprache der Körpertheorie formulierbar ist. Dazu nehmen wir an, S' wäre eine zur Formulierung des archimedischen Axioms geeignete elementare Sprache (die eventuell außer den Grundbegriffen „+", „·", „<", 0, 1, c noch beliebig viele weitere Relations-, Operations- und Konstantensymbole enthalten kann), und X_0 wäre ein in S' formuliertes endliches Axiomensystem für geordnete Körper einschließlich des archimedischen Axioms. Es sei dann für $n = 1, 2, 3, \ldots$ der Ausdruck H_n wie in 7. gebildet und wieder

$$X = X_0 \cup \{\mathrm{H}_n : n = 1, 2, 3, \ldots\}.$$

Falls die Sprache S', das Axiomensystem X_0 und insbesondere die darin enthaltene Formulierung des archimedischen Axioms „vernünftig" sind, muß sich S' so im Körper der rationalen Zahlen interpretieren lassen, daß ein Modell von X_0 entsteht, wobei wiederum die Konstante c, die in X_0 nicht vorkommen soll, durch eine beliebige rationale Zahl interpretiert werden kann. Demnach ist jede endliche Teilmenge von X, und damit auch X selbst, semantisch widerspruchsfrei, d. h., es muß einen im Sinne des Axiomensystems X_0 „archimedisch" geordneten Körper geben, in dem zugleich ein Element existiert, das größer als alle endlichen Vielfachen des Einselements ist.

Aus 8. folgt sofort: Ist umgekehrt in einer nichtelementaren Sprache ein inhaltlich zutreffendes archimedisches Axiom formulierbar, so kann für diese nichtelementare Sprache (S, σ) der Satz von Mal'cev nicht gelten und daher auch kein für (S, σ) vollständiger Kalkül existieren. 8. und die eben formulierte Folgerung bilden das Musterbeispiel für eine Reihe von Resultaten folgender Art: Gewisse Begriffe können in elementaren Sprachen nicht definierbar bzw. gewisse Sachverhalte nicht formulierbar sein. Sind diese Begriffe bzw. Sachverhalte andererseits in gewissen nichtelementaren Sprachen nachweislich doch definier- bzw. formulierbar, so kann für diese Sprachen ein zum Hauptsatz analoger Satz nicht gelten. Als wesentliche weitere Resultate dieser Art behandeln wir noch die Nichtexistenz eines vollständigen

Kalküls für die durch ein einstelliges Relationssymbol erweiterte nichtelementare Sprache der Peano-Axiome (vgl. 4.5., Beispiel 5) und die Unmöglichkeit einer Endlichkeitsdefinition in elementaren Formalisierungen der Mengenlehre (vgl. Kapitel 5).

9. *Es sei S die um ein einstelliges Relationssymbol* R *bereicherte Sprache zur Formulierung der Peano-Axiome und $\sigma_{(2)}$ die dort betrachtete Interpretationseinschränkung. Dann gibt es keinen vollständigen Kalkül für* $(S, \sigma_{(2)})$.

Beweis. Wir nehmen an, daß ein vollständiger Kalkül für $(S, \sigma_{(2)})$ existiert. Dann gilt für das $\sigma_{(2)}$-Folgern der Endlichkeitssatz und folglich auch ein Endlichkeitssatz für $\sigma_{(2)}$-Modelle: Besitzt jede endliche Teilmenge von $X \subseteq S$ ein $\sigma_{(2)}$-Modell, so besitzt ebenfalls X ein $\sigma_{(2)}$-Modell. Es sei nun X_0 das in S formulierte endliche System der Peano-Axiome (worin das Symbol R nicht vorkommt) und H_0 der Ausdruck $\bigvee \mathrm{n} \neg \mathrm{R}(\mathrm{n})$ sowie H_m der Ausdruck $\mathrm{R}(\mathrm{o}\overbrace{'\cdots'}^{m})$ für $m \geqq 1$ (d. h., H_m hat die Bedeutung: Die durch R bezeichnete einstellige Relation trifft auf die Zahl m zu). Jede endliche Teilmenge von $X = X_0 \cup \{\mathrm{H}_m : m = 0, 1, 2, \ldots\}$ besitzt ein $\sigma_{(2)}$-Modell in Gestalt der natürlichen Zahlen, wobei R durch eine geeignete einstellige Relation im Bereich der natürlichen Zahlen zu interpretieren ist. Ist m_0 der größte Index der in einer solchen Teilmenge von X vorkommenden Ausdrücke H_m, so braucht man für R nur diejenigen Relationen zu wählen, die auf die Zahlen $0, 1, \ldots, m_0$ zutrifft und auf die übrigen nicht (damit gegebenenfalls H_0 erfüllt wird). Auf Grund des Endlichkeitssatzes für $\sigma_{(2)}$-Modelle müßte daher auch die gesamte Menge X ein $\sigma_{(2)}$-Modell besitzen. Da X_0 bereits $\sigma_{(2)}$-kategorisch ist, kann es sich hierbei nur um „die" natürlichen Zahlen handeln, in denen zusätzlich eine Relation existiert, die auf $0, 1, 2, 3, \ldots$ zutrifft, jedoch nicht auf alle natürlichen Zahlen. Ein herrlicher Widerspruch!

10. *In einer elementar formalisierten Mengenlehre gibt es keinen Ausdruck* H(x) *mit genau einer vollfreien Variablen* x, *der in jedem Modell des betrachteten Axiomensystems genau auf die endlichen Mengen dieses Modells zutrifft.*

Beweis. Es sei X_0 ein endliches elementares widerspruchsfreies Axiomensystem der Mengenlehre, S die zu seiner Formulierung benutzte elementare Sprache, vermehrt um eine Konstante M zur Bezeichnung einer beliebigen festen Menge. Ist H(x) eine „Endlichkeitsdefinition" in der Sprache S, so sei H_0 der Ausdruck H(M) und, für $m \geqq 1$ sei der Ausdruck H_m durch $\bigvee_{m} \mathrm{x}\, \mathrm{x} \in \mathrm{M}$ gegeben. Jede endliche Teilmenge von

$$X_0 \cup \{\mathrm{H}_m : m = 0, 1, 2, 3, \ldots\}$$

besitzt als Modell zumindest das vorausgesetzte Modell von X_0, worin die zusätzliche Konstante M durch eine genügend große endliche Menge interpretiert ist. Daher muß nach dem Satz von Mal'cev ein Modell von X_0 existieren, in dem es eine im

Sinne der Definition H(x) endliche Menge gibt, die jedoch im anschaulichen Sinn unendlich viele Elemente enthält. Für viele konkrete Axiomensysteme und Endlichkeitsdefinitionen konnte man solche „Nichtstandardmodelle“ explizit angeben. Wir verweisen nochmals auf Kapitel 5 und die dort zum Thema Endlichkeitsdefinitionen angegebene Literatur (vgl. S. 112).

11. Eine nichtelementare Sprache (S, σ) war als unwesentlich nichtelementar bezeichnet worden (vgl. 4.5., Definition 5), wenn ein Axiomensystem $Y \subset S$ existiert, so daß

$$Fl_S{}^\sigma(X) = Fl_S(X \cup Y) \qquad \text{für } X \subseteqq S$$

gilt. Präzisieren wir dies nachträglich dahingehend, daß Y durch ein endliches oder wenigstens überschaubares System von Axiomenschemata (im Sinne 0-gliedriger Schlußregeln) erzeugbar sein soll, so können wir einen beliebigen vollständigen Kalkül $\mathcal{K}$ für die elementare Sprache S so zu einem Kalkül $\mathcal{K}'$ erweitern, daß für $X \subseteqq S$ gilt:

$$Bew_{\mathcal{K}'}(X) = Bew_{\mathcal{K}}(X \cup Y) = Fl(X \cup Y) = Fl^\sigma(X),$$

d. h., *zu jeder unwesentlich nichtelementaren Sprache* (S, σ) *gibt es einen vollständigen Kalkül, und umgekehrt kann eine solche nichtelementare Sprache, für die kein vollständiger Kalkül existiert* (vgl. die in Kapitel 8 und 9 betrachteten Sprachen) *nicht unwesentlich nichtelementar sein.*

6.4. Beweis des Hauptsatzes

Für den detaillierten Nachweis, daß bestimmte Kalküle vollständig sind, verweisen wir auf weiterführende Lehrbücher. In diesem Abschnitt geht es darum, einen der heute bekannten Beweiswege, der auf L. Henkin[1]) zurückgeht, in seinen Grundzügen, ohne Bezugnahme auf einen bestimmten Kalkül, darzulegen und dabei herauszuarbeiten, wie man anhand der Beweisidee die Axiome und Schlußregeln, die ein vollständiger Kalkül enthalten müßte, nach und nach zusammenstellen kann.

Für das Folgende wird sich eine mögliche Eigenschaft von Beweiskalkülen als äußerst wichtig erweisen, die im Fall ihrer Gültigkeit als *Deduktionstheorem* für den entsprechenden Kalkül bezeichnet wird:

Ist $\mathrm{H}_2 \in Bew(X \cup \{\mathrm{H}_1\})$ *und* H_1 *abgeschlossen, so ist* $\mathrm{H}_1 \rightarrow \mathrm{H}_2 \in Bew(X)$.

Der Leser überprüfe zunächst, daß dieses Deduktionstheorem für beliebige Folgerungsoperatoren Fl^σ (für Bew eingesetzt) erfüllt und daß hierbei die Voraussetzung

[1]) L. Henkin, The completeness of the first order functional calculus, J. of Symb. Logic **14** (1949), 159–166.

der Abgeschlossenheit von H_1 wesentlich ist. Die Gültigkeit des Deduktionstheorems für das Folgern bedeutet insbesondere, daß es für jeden vollständigen Kalkül ebenfalls gelten muß, mit anderen Worten, daß es eine notwendige Bedingung für die Vollständigkeit eines Kalküls ist. Im folgenden Satz formulieren wir ein System von hinreichenden (jedoch keineswegs notwendigen) Bedingungen dafür, daß für einen Beweiskalkül $\mathcal{K}$ das Deduktionstheorem gilt.

Satz 1. *Ist $\mathcal{K}$ ein aussagenlogisch vollständiger Kalkül mit $\varphi_1, \varphi_2, \varphi_{10}$ als einzigen echten Schlußregeln* (vgl. 6.1., Satz 6) *und*

a) $(H' \to H''(x)) \to (H' \to \wedge x H''(x)) \in Bew_{\mathcal{K}}(\emptyset)$, *falls* H' *abgeschlossen,*

b) $(H' \to H'') \to (H' \to H^*) \in Bew_{\mathcal{K}}(\emptyset)$, *falls* H' *abgeschlossen und* H^* *durch einmalige Anwendung von* φ_2 *auf* H'' *entsteht,*

(die Voraussetzungen a), b) lassen sich z. B. durch Aufnahme der entsprechenden Ausdrücke unter die logischen Axiome von $\mathcal{K}$ erfüllen), *so gilt für den Kalkül $\mathcal{K}$ das Deduktionstheorem.*

Zum Beweis von Satz 1 denken wir uns X und den abgeschlossenen Ausdruck H_1 beliebig gewählt und zeigen die Behauptung durch Induktion bezüglich des induktiven Aufbaus von $Bew_{\mathcal{K}}(X \cup \{H_1\})$.

1. Fall: $H_2 \in X$ oder H_2 ist logisches Axiom von $\mathcal{K}$. Dann ist $H_2 \in Bew_{\mathcal{K}}(X)$, ferner ist $H_2 \to (H_1 \to H_2)$ aussagenlogisch allgemeingültig, folglich Element von $Bew_{\mathcal{K}}(X)$. Anwendung von φ_1 auf diese beiden Prämissen ergibt die Behauptung.

2. Fall: $H_2 = H_1$. Dann ist $H_1 \to H_2 \in Bew_{\mathcal{K}}(X)$ wegen seiner aussagenlogischen Allgemeingültigkeit.

3. Fall: H_2 entsteht durch Anwendung von φ_1 auf Ausdrücke $H_3 \to H_2, H_3 \in Bew_{\mathcal{K}}(X \cup \{H_1\})$, für die schon $H_1 \to (H_3 \to H_2), H_1 \to H_3 \in Bew_{\mathcal{K}}(X)$ gilt. Aussagenlogisch allgemeingültig ist aber $(H_1 \to (H_3 \to H_2)) \to ((H_1 \to H_3) \to (H_1 \to H_2))$. Daher gehört dieser Ausdruck zu $Bew_{\mathcal{K}}(X)$, und man erhält die Behauptung durch zweimalige Anwendung der Abtrennungsregel.

4. Fall: H_2 entsteht durch Anwendung der Regel φ_2 der gebundenen Umbenennung aus einem Ausdruck $H_3 \in Bew_{\mathcal{K}}(X \cup \{H_1\})$, für den schon $H_1 \to H_3 \in Bew_{\mathcal{K}}(X)$ gilt. Dann erhält man $H_1 \to H_2 \in Bew_{\mathcal{K}}(X)$ durch Anwendung der Abtrennungsregel auf das logische Axiom $(H_1 \to H_3) \to (H_1 \to H_2) \in Bew_{\mathcal{K}}(X)$, das unter den gegebenen Voraussetzungen in das Schema von Axiom b) gehört.

5. Fall: $H_2 = \varphi_{10}(H_3(x); x) = \wedge x H_3(x)$ mit $H_3(x) \in Bew_{\mathcal{K}}(X \cup \{H_1\})$, wobei schon $H_1 \to H_3(x) \in Bew_{\mathcal{K}}(X)$ gilt. In diesem Fall erhält man die Behauptung analog zum 4. Fall unter Benutzung von Axiom a).

Satz 2. *Gilt für einen Kalkül $\mathcal{K}$:*

a) *$\mathcal{K}$ ist aussagenlogisch vollständig,*

b) *das Deduktionstheorem,*

c) *entsteht der abgeschlossene Ausdruck* $\overline{\mathrm{H}}$ *aus* H *durch gebundene Umbenennungen und Generalisierungen, so ist* $\overline{\mathrm{H}} \to \mathrm{H} \in Bew_{\mathcal{K}}(\emptyset)$,

d) *jede syntaktisch widerspruchsfreie Menge besitzt ein Modell,*

so ist $\mathcal{K}$ *vollständig.*

Alle vier Voraussetzungen sind offenbar notwendige Bedingungen für die Vollständigkeit. Dabei lassen sich a) und b) gemäß 6.2. und Satz 1 leicht verwirklichen und c) gegebenenfalls durch Aufnahme eines entsprechenden Schemas logischer Axiome, ohne daß dadurch die Voraussetzungen a), b) wieder gefährdet werden. Demnach liefert Voraussetzung d) eine handliche Methode für den Nachweis, daß gewisse Kalküle vollständig sind. Wir beweisen zunächst Satz 2:

Es ist nur zu zeigen, daß für beliebige Ausdrücke H und Ausdrucksmengen X aus $\mathrm{H} \notin Bew_{\mathcal{K}}(X)$ stets $\mathrm{H} \notin Fl(X)$ folgt. Es sei $\mathrm{H} \notin Bew_{\mathcal{K}}(X)$, $\overline{H}$ ein abgeschlossener Ausdruck, der dadurch aus H entsteht, daß man alle eventuell in H frei vorkommenden Variablen durch gebundene Umbenennungen vollfrei macht und dann zur Generalisierten übergeht. Wäre $X \cup \{\neg \overline{H}\}$ syntaktisch widerspruchsvoll, so wäre insbesondere $\overline{H} \in Bew_{\mathcal{K}}(X \cup \{\neg \overline{H}\})$, folglich nach dem Deduktionstheorem $\neg \overline{H} \to \overline{H} \in Bew_{\mathcal{K}}(X)$, woraus wegen der aussagenlogischen Vollständigkeit von $\mathcal{K}$ nun $\overline{H} \in Bew_{\mathcal{K}}(X)$ und daher wegen Voraussetzung c) und a) auch $\mathrm{H} \in Bew_{\mathcal{K}}(X)$ folgt, im Widerspruch zum Beweisanfang. Demnach ist $X \cup \{\neg \overline{H}\}$ syntaktisch widerspruchsfrei, besitzt also nach d) ein Modell. Dies ist ein Modell von X, in dem H nicht allgemeingültig ist, d. h. $\mathrm{H} \notin Fl(X)$, was zu zeigen war.

Die Idee von Henkin, für geeignete Kalküle die Voraussetzung d) von Satz 2 zu beweisen, besteht darin, im wesentlichen die Menge der variablenfreien Terme der betrachteten Sprache als Grundbereich eines zu konstruierenden Modells ω für eine als syntaktisch widerspruchsfrei vorausgesetzte Menge X zu benutzen. In diesem Grundbereich soll ω wie folgt definiert werden:

a) Für Konstantensymbole c der Sprache sei $\omega(\mathrm{c}) = \mathrm{c}$.

b) Ist F ein n-stelliges Operationssymbol und sind $\mathrm{t}_1, \ldots, \mathrm{t}_n$ variablenfreie Terme entsprechender Sorte, so sei $\omega(\mathrm{F})\,(\mathrm{t}_1, \ldots, \mathrm{t}_n) := \mathrm{F}(\mathrm{t}_1, \ldots, \mathrm{t}_n)$. Auf diese Weise ordnet in der Tat die Operation $\omega(\mathrm{F})$ variablenfreien Termen $\mathrm{t}_1, \ldots, \mathrm{t}_n$ wieder einen variablenfreien Term zu.

c) Ist R ein n-stelliges Relationssymbol und sind $\mathrm{t}_1, \ldots, \mathrm{t}_n$ variablenfreie Terme entsprechender Sorte, so sei

$$(\mathrm{t}_1, \ldots, \mathrm{t}_n) \in \omega\,(\mathrm{R}) :\leftrightarrow \mathrm{R}(\mathrm{t}_1, \ldots, \mathrm{t}_n) \in Bew_{\mathcal{K}}(X).$$

Auf diese Weise ist in der Tat eine n-stellige Relation $\omega(\mathrm{R})$ im Bereich der variablenfreien Terme definiert.

Aus der Anfangsfestsetzung c) soll sich durch Induktion bezüglich der Kompliziertheit der Ausdrücke ergeben: Kommen in H genau die Variablen $\mathrm{x}_1, \ldots, \mathrm{x}_n$

frei vor und ordnet eine Belegung f bezüglich der betrachteten Interpretation ω diesen Variablen die Terme $t_1, \ldots, t_n$ zu, so gilt

$$Wert(H, \omega, f) = W \textit{ genau dann, wenn } Sub(H; x_1, t_1; \ldots; x_n, t_n) \in Bew_{\mathcal{K}}(X). \tag{1}$$

Ist $\mathcal{K}$ nun vollständig, so muß durch geeignete Schlußregeln bzw. logische Axiome gesichert sein, daß aus H jeder Ausdruck der Form $Sub(H; x_1, t_1; \ldots; x_n, t_n)$ beweisbar ist. Daher ergibt sich aus (1): Ist $H \in Bew_{\mathcal{K}}(X)$, so ist H bei der Interpretation ω allgemeingültig, d. h., ω ist ein Modell für $Bew_{\mathcal{K}}(X)$ und damit erst recht für X.

Die Ausarbeitung dieser Idee zu einem Beweis erfordert allerdings einige Korrekturen des Ansatzes:

1. Da die eventuell vorhandenen Operationssymbole auch durch partielle Operationen interpretiert werden dürfen, bezeichnet nicht jeder variablenfreie Term bei jeder Interpretation ein Ding. Vielmehr folgt die Existenz einer Bedeutung des variablenfreien Terms t genau dann aus X, wenn $t = t \in Fl(X)$, d. h. (unter Voraussetzung der Vollständigkeit des betrachteten Kalküls) wenn $t = t \in Bew_{\mathcal{K}}(X)$ gilt. Die Menge dieser variablenfreien Terme bezeichnen wir mit $T(X)$. (Im Fall des in der Literatur meist betrachteten Nachweises der Vollständigkeit von Kalkülen für Sprachen $(S, \sigma_{(1)})$ (vgl. 4.5.) entfällt diese Betrachtung, ebenso, falls die betrachtete Sprache keine Operationssymbole enthält. In beiden Fällen ist also $T(X)$ unabhängig von X die Menge aller variablenfreien Terme.)

2. Ist $t_1 = t_2 \in Bew_{\mathcal{K}}(X)$ für variablenfreie Terme t_1, t_2 (dann ist schon $t_1, t_2 \in T(X)$), so bezeichnen t_1 und t_2 in jedem Modell von X das gleiche Ding. Demnach darf man nicht $T(X)$ selbst, sondern die Äquivalenzklassen bezüglich der durch

$$t_1 \cong t_2 :\leftrightarrow t_1 = t_2 \in Bew_{\mathcal{K}}(X)$$

definierten Relation als Grundbereich des zu konstruierenden Modells nehmen. Der Nachweis, daß „$\cong$" eine Äquivalenzrelation in $T(X)$ ist und daß die sinngemäß auf die Äquivalenzklassen übertragenen Definitionen von $\omega(F)$ bzw. $\omega(R)$ repräsentantenunabhängig sind, führt auf ein System von identitätstheoretischen Ausdrücken, die für jeden vollständigen Kalkül (analog zu einem aussagenlogisch vollständigen Axiomensystem) in $Bew_{\mathcal{K}}(\emptyset)$ enthalten, also entweder logische Axiome von $\mathcal{K}$ oder aus solchen mittels der Schlußregeln von $\mathcal{K}$ beweisbar sein müssen. Zum Beispiel führt die Forderung der Symmetrie von „$\cong$" auf

$$t_1 = t_2 \rightarrow t_2 = t_1 \in Bew_{\mathcal{K}}(\emptyset),$$

so daß wegen der aussagenlogischen Vollständigkeit von $\mathcal{K}$ aus $t_1 = t_2 \in Bew_{\mathcal{K}}(X)$ wie gewünscht $t_2 = t_1 \in Bew_{\mathcal{K}}(X)$ folgt. Analog führt die Forderung der Repräsentantenunabhängigkeit der Definition von $\omega(R)$ auf

$$t_i = t_i' \wedge R(t_1, \ldots, t_{i-1}, t_i, \ldots, t_n) \rightarrow R(t_1, \ldots, t_{i-1}, t_i', \ldots, t_n) \in Bew_{\mathcal{K}}(\emptyset).$$

Reduziert man diese Ausdrücke unter Berücksichtigung der Schlußregeln eines speziellen Kalküls noch so weit wie möglich, so gelangt man auf diesem Wege fast zwangsläufig zu einem System identitätstheoretischer Axiome als notwendigem Bestandteil eines vollständigen Kalküls.

3. Im allgemeinen ist der Übergang von der zunächst betrachteten Sprache S zu einer durch hinreichend viele Konstantensymbole erweiterten Sprache S' nötig (in S braucht es überhaupt keine variablenfreien Terme zu geben!), denn zur Realisierung des angestrebten Modells muß es zu jedem Ausdruck der Form $\bigvee \mathrm{xH(x)}$ aus $Bew_{\mathcal{K}}(X)$ einen variablenfreien Term t geben, so daß $\mathrm{H(t)} \in Bew_{\mathcal{K}}(X)$ gilt.

4. Als Spezialfall ergibt sich aus (1), daß für abgeschlossene Ausdrücke H

$$\mathrm{H} \in Bew_{\mathcal{K}}(X) \Leftrightarrow \neg\, \mathrm{H} \notin Bew_{\mathcal{K}}(X)$$

gelten muß. Hierin folgt die Richtung „$\Rightarrow$" aus der vorausgesetzten syntaktischen Widerspruchsfreiheit von X (die wegen der aussagenlogischen Vollständigkeit von $\mathcal{K}$ mit der formalen Widerspruchsfreiheit übereinstimmt). Die Gegenrichtung „$\Leftarrow$" gilt aber nur dann, wenn $Bew_{\mathcal{K}}(X)$ sogar eine maximale widerspruchsfreie Menge ist. Ist andererseits ω ein Modell für eine maximale widerspruchsfreie Menge Y mit $Y \supseteqq X$, so ist ω auch ein Modell für X. Daher benötigen wir:

Satz 3. *Zu jeder syntaktisch (und formal) widerspruchsfreien Menge X gibt es eine maximale syntaktisch widerspruchsfreie Menge $Y \supseteqq X$.*

Beweis. Unter den von uns angenommenen Voraussetzungen ist jede formalisierte Sprache S als unendliche Teilmenge einer Wortmenge $W(A)$ mit abzählbarem Alphabet selbst abzählbar. Wir gehen von einer Abzählung $S = \{\mathrm{H}_0, \mathrm{H}_1, \mathrm{H}_2, \ldots\}$ von S aus und definieren induktiv Obermengen X_n von X:

$$X_0 := X,$$

$$X_{n+1} := \begin{cases} X_n \cup \{\mathrm{H}_n\}, & \textit{falls } X_n \cup \{\mathrm{H}_n\} \textit{ widerspruchsfrei}, \\ X_n & \textit{sonst}. \end{cases}$$

Dann gilt $X_0 \subseteqq X_1 \subseteqq X_2 \subseteqq \cdots$, und alle Mengen X_n sind nach Konstruktion widerspruchsfrei. Wir behaupten nun, daß $Y := \bigcup\limits_{n=0}^{\infty} X_n$ eine maximale widerspruchsfreie Obermenge von X ist. Klar ist $X \subseteqq Y$. Wir zeigen, daß Y widerspruchsfrei ist: Falls nicht, gibt es einen Ausdruck H, so daß $\mathrm{H}, \neg\, \mathrm{H} \in Y$ ist. Nach Definition von Y existieren daher natürliche Zahlen m, n mit $\mathrm{H} \in X_n$ und $\neg\, \mathrm{H} \in X_m$. Ist dabei $m \leqq n$, so ist $\mathrm{H}, \neg\, \mathrm{H} \in X_n$. Für $n < m$ ist $\mathrm{H}, \neg\, \mathrm{H} \in X_m$. Alle Mengen X_n sind aber widerspruchsfrei. Wir zeigen weiter, daß Y eine maximale widerspruchsfreie Menge ist, mit anderen Worten, daß $Y \cup \{\mathrm{H}\}$ widerspruchsvoll für $\mathrm{H} \notin Y$ wird. Jeder Ausdruck H hat eine gewisse Nummer n bezüglich der betrachteten Abzählung von S. Ist nun $X_n \cup \{\mathrm{H}_n\}$ widerspruchsfrei, so ist $\mathrm{H}_n \in X_{n+1} \subseteqq Y$. Falls also $\mathrm{H}_n \notin Y$, so ist $X_n \cup \{\mathrm{H}_n\}$ widerspruchsvoll, daher ist $Y \cup \{\mathrm{H}_n\}$ erst recht widerspruchsvoll.

Wir geben nun in Einzelschritten die endgültige Konstruktion eines Modells für eine syntaktisch widerspruchsfreie Menge an, wobei wir solche Beweise auslassen, die nur unter Bezug auf den jeweils betrachteten konkreten Kalkül geführt werden können.

1. Schritt. Gegeben sei eine syntaktisch widerspruchsfreie Teilmenge X einer elementaren Sprache S. Die Sprache S' entstehe aus S durch Hinzunahme abzählbar vieler neuer Konstantensymbole $c_0, c_1, c_2, \ldots$ (für jede Sorte; wir beschränken uns hier und im folgenden zwecks Vereinfachung der Formulierungen auf den Fall einsortiger Sprachen). Ferner sei $H_0, H_1, H_2, \ldots$ eine Abzählung aller abgeschlossenen Ausdrücke der Form $\bigvee x H(x)$ in S', so daß H_n die Form $\bigvee x H_{(n)}(x)$ mit einer gewissen Variablen x hat. Wir definieren induktiv in S' abzählbar viele neue Axiome A_n:

A_0 sei der Ausdruck $H_0 \rightarrow H_{(0)}(c_{k_0})$, wobei c_{k_0} die bezüglich der Numerierung dieser Konstanten erste neue Konstante sei, die in H_0 nicht vorkommt.

Ist schon $A_0, \ldots, A_n$ definiert, so sei A_{n+1} der Ausdruck $H_{n+1} \rightarrow H_{(n+1)}(c_{k_{n+1}})$, wobei $c_{k_{n+1}}$ die erste neue Konstante ist, die in $H_0, \ldots, H_{n+1}$ nicht vorkommt und von $c_{k_0}, \ldots, c_{k_n}$ verschieden ist.

Der Sinn dieser Axiome besteht darin, festzulegen, daß c_{k_n} ein gewisses Ding mit der Eigenschaft $H_{(n)}(x)$ bezeichnen soll, falls die Existenz eines solchen Dinges x beweisbar ist.

2. Schritt. Unter Benutzung des konkreten Kalküls $\mathcal{K}$ zeigt man: *Ist X syntaktisch widerspruchsfrei in S, so ist die Menge $X \cup \{A_0, A_1, \ldots, A_n\}$ für jede natürliche Zahl n syntaktisch widerspruchsfrei in S'.*

Nach dem für $Bew_{\mathcal{K}}$ geltenden Endlichkeitssatz folgt hieraus sofort: *Die Menge $X \cup \{A_n : n = 0, 1, 2, \ldots\}$ ist syntaktisch widerspruchsfrei in S'.*

3. Schritt. Es sei Y eine nach Satz 3 existierende maximale syntaktisch (und formal) widerspruchsfreie Obermenge von $X \cup \{A_n : n = 0, 1, 2, \ldots\}$. Dann ist übrigens $Bew_{\mathcal{K}}(Y) = Y$, da mit Y auch $Bew_{\mathcal{K}}(Y)$ syntaktisch widerspruchsfrei ist und wegen der Maximalität von Y keine echte Obermenge von Y sein kann. Für $t \in T(Y)$ sei $\bar{t} := \{t' : t' = t \in Y\}$, und es sei $\overline{T(Y)} := \{\bar{t} : t \in T(Y)\}$. Wir definieren nun in der bereits besprochenen Weise eine Interpretation ω von S' im Grundbereich $\overline{T(Y)}$. Aus dieser Definition folgt zunächst für prädikative abgeschlossene Ausdrücke $H \in S'$ (einschließlich Termgleichungen)

$$Wert(H, \omega) = W \text{ genau dann, wenn } H \in Y. \tag{2}$$

4. Schritt. Zur Rechtfertigung des folgenden Beweises ist zunächst zu zeigen, daß jeder abgeschlossene Ausdruck $H \in S'$ aus abgeschlossenen prädikativen Ausdrücken durch aussagenlogische Verknüpfungen und Übergänge von Ausdrücken der Form H(t) (t variablenfreier Term) zu Ausdrücken $\bigwedge x H(x)$ bzw. $\bigvee x H(x)$ erhalten werden kann. Für beliebige Sprachen ist dies schon deshalb nicht richtig, weil es gar keine abgeschlossenen prädikativen Ausdrücke gibt, wenn es keine Konstantensymbole (und damit auch keine variablenfreien Terme) gibt. In S' kann man jedoch

in einem abgeschlossenen Ausdruck H alle in seinen prädikativen Teilausdrücken eventuell vorkommenden Variablen dort durch paarweise verschiedene Konstanten c_k ersetzen, die in H sonst nicht vorkommen. Beim induktiven Aufbau von H aus den so abgeänderten prädikativen Teilausdrücken werden die ursprünglichen Variablen durch Übergänge von $H'(t)$ zu $\bigwedge xH'(x)$ bzw. $\bigvee xH'(x)$ nach und nach wieder eingeführt. Zum Beispiel entsteht der abgeschlossene Ausdruck

$$\bigvee x \bigwedge y\big(R_1(x) \to R_2(x, y)\big) \tag{3}$$

wie folgt induktiv aus abgeschlossenen Ausdrücken: $R_1(c_1)$ und $R_2(c_1, c_2)$ sind abgeschlossene Ausdrücke. Folglich ist auch $(R_1(c_1) \to R_2(c_1, c_2))$ ein abgeschlossener Ausdruck. Da in ihm die Variable y noch nicht gebunden vorkommt, ist auch $\bigvee y(R_1(c_1) \to R_2(c_1, y))$ und aus dem gleichen Grund auch (3) ein abgeschlossener Ausdruck.

5. Schritt. Wir zeigen durch das im 4. Schritt gerechtfertigte Induktionsprinzip, daß (2) für alle abgeschlossenen Ausdrücke $H \in \boldsymbol{S}'$ gilt. Dabei kann die Gültigkeit für den Fall, daß H prädikativ ist, schon vorausgesetzt werden. Setzen wir voraus, daß (2) für H schon gilt, so ist $\mathit{Wert}(\neg H, \omega) = \boldsymbol{W}$ genau dann, wenn $\mathit{Wert}(H, \omega) = \boldsymbol{F}$ gilt. Das ist aber nach Induktionsannahme genau dann der Fall, wenn $H \notin \boldsymbol{Y}$ ist. Da $\boldsymbol{Y}$ widerspruchsfrei ist, folgt $H \notin \boldsymbol{Y}$ aus $\neg H \in \boldsymbol{Y}$, und da $\boldsymbol{Y}$ sogar maximal widerspruchsfrei ist, folgt umgekehrt $\neg H \in \boldsymbol{Y}$ aus $H \notin \boldsymbol{Y}$. Daher ist $\mathit{Wert}(\neg H, \omega) = \boldsymbol{W}$ genau dann, wenn $\neg H \in \boldsymbol{Y}$. Wir setzen nun voraus, daß (2) bereits für abgeschlossene Ausdrücke H_1, H_2 gilt. Es ist $\mathit{Wert}(H_1 \wedge H_2, \omega) = \boldsymbol{W}$ genau dann, wenn $\mathit{Wert}(H_1, \omega) = \mathit{Wert}(H_2, \omega) = \boldsymbol{W}$, also nach Induktionsannahme genau dann, wenn $H_1 \in \boldsymbol{Y}$ und $H_2 \in \boldsymbol{Y}$ gilt. Wegen der aussagenlogischen Vollständigkeit von $\mathcal{K}$ (und $\boldsymbol{Y} = \mathit{Bew}_{\mathcal{K}}(\boldsymbol{Y})$) ist dies genau dann der Fall, wenn $H_1 \wedge H_2 \in \boldsymbol{Y}$ ist. Die restlichen Fälle aussagenlogischer Verknüpfung von H_1, H_2 sind nach folgendem Muster zu behandeln: $\mathit{Wert}(H_1 \vee H_2, \omega) = \boldsymbol{W}$ genau dann, wenn $\mathit{Wert}(\neg(\neg H_1 \wedge \neg H_2), \omega) = \boldsymbol{W}$, d. h. auf Grund des bereits Gezeigten genau dann, wenn

$$\neg\neg(H_1 \wedge \neg H_2) \in \boldsymbol{Y}(= \mathit{Bew}_{\mathcal{K}}(\boldsymbol{Y}))$$

gilt. Wegen der aussagenlogischen Vollständigkeit des Kalküls $\mathcal{K}$ ist dies genau dann der Fall, wenn $H_1 \vee H_2 \in \boldsymbol{Y}$ ist. Wir setzen nun voraus, daß (2) bei festem H(x) für alle H(t) mit $t \in \boldsymbol{T}(\boldsymbol{Y})$ gilt. Es ist $\mathit{Wert}(\bigvee xH(x), \omega) = \boldsymbol{W}$ genau dann, wenn es einen Term $t \in \boldsymbol{T}(\boldsymbol{Y})$ gibt, so daß $\mathit{Wert}(H(x), \omega, f) = \boldsymbol{W}$ für alle Belegungen f mit $f(x) = \bar{t}$, d. h., wenn $\mathit{Wert}(H(t), \omega) = \boldsymbol{W}$ ist. Nach Induktionsannahme gilt das letzte genau dann für einen Term $t \in \boldsymbol{T}(\boldsymbol{Y})$, wenn $H(t) \in \boldsymbol{Y}$ ist. Durch geeignete Schlußregeln in $\mathcal{K}$ läßt sich nun leicht erreichen, daß mit $H(t) \in \mathit{Bew}_{\mathcal{K}}(\boldsymbol{Z})$ stets auch $\bigvee xH(x) \in \mathit{Bew}_{\mathcal{K}}(\boldsymbol{Z})$ für Variablen entsprechender Sorte gilt, die in H(t) nicht vorkommen. Die Umkehrung, die ja in beliebigen Sprachen und Theorien gar nicht richtig ist, wird für $\boldsymbol{Y}$ gerade durch die Axiome A_n erzwungen. Der betrachtete abgeschlossene Ausdruck $\bigvee xH(x)$ ist gleich einem gewissen H_n, so daß das entsprechende Axiom $\bigvee xH(x) \to H(c_{k_n})$ zu $\boldsymbol{Y}$ gehört. Ist nun noch $\bigvee xH(x) \in \boldsymbol{Y}$, so folgt aus der aussagenlogischen Vollständigkeit von $\mathcal{K}$ sofort $\mathrm{H}(c_{k_n}) \in \boldsymbol{Y}$, d. h. $\mathrm{H}(t) \in \boldsymbol{Y}$ für einen Term $t \in \boldsymbol{T}(\boldsymbol{Y})$. Insgesamt ergibt sich daher

$$Wert(\vee\, x\mathrm{H}(x), \omega) = \boldsymbol{W} \text{ genau dann, wenn } \vee\, x\mathrm{H}(x) \in \boldsymbol{Y}.$$

Wir setzen wie oben voraus, daß bei festem $\mathrm{H}(x)$ schon (2) für alle Ausdrücke der Form $\mathrm{H}(t)$ mit $t \in \boldsymbol{T}(\boldsymbol{Y})$ gilt. Dann ist

$$Wert(\wedge\, x\mathrm{H}(x), \omega) = \boldsymbol{W} \text{ genau dann, wenn } Wert(\neg \vee\, x \neg\, \mathrm{H}(x), \omega) = \boldsymbol{W},$$

und das ist auf Grund des schon Bewiesenen genau dann der Fall, wenn $\neg \vee\, x \neg\, \mathrm{H}(x) \in \boldsymbol{Y}$ ist. Durch ein geeignetes logisches Axiom bzw. geeignete Schlußregeln kann man nun leicht erreichen, daß aus $\neg \vee\, x \neg\, \mathrm{H}(x)$ der Ausdruck $\wedge\, x\mathrm{H}(x)$ beweisbar ist und umgekehrt.

6. Schritt. Ist $\mathrm{H} \in \boldsymbol{Y}$ ein beliebiger (nicht notwendig abgeschlossener) Ausdruck, so ist bei Vorhandensein geeigneter Schlußregeln (etwa φ_2 und φ_{10}) erreichbar, daß auch jeder abgeschlossene Ausdruck $\overline{\mathrm{H}}$, der aus H durch gebundene Umbenennungen und Generalisierungen entsteht, zu $\boldsymbol{Y}$ gehört. Nach dem Bewiesenen folgt aus $\overline{\mathrm{H}} \in \boldsymbol{Y}$ aber $Wert(\overline{\mathrm{H}}, \omega) = \boldsymbol{W}$, d. h., H selbst ist in ω allgemeingültig bzw. ω ist ein Modell für $\boldsymbol{Y}$. Daher ist die Einschränkung der Interpretation ω der Sprache $\boldsymbol{S}'$ auf die Sprache $\boldsymbol{S}$ ein Modell der in $\boldsymbol{S}$ formulierten Teilmenge $\boldsymbol{X}$ von $\boldsymbol{Y}$.

Da der Grundbereich des so für $\boldsymbol{X}$ konstruierten Modells aus den Äquivalenzklassen einer in der Menge $\boldsymbol{T}(\boldsymbol{Y})$ definierten Äquivalenzrelation besteht und $\boldsymbol{T}(\boldsymbol{Y})$ als Teilmenge einer Wortmenge $\boldsymbol{W}(\boldsymbol{A})$ mit abzählbarem Alphabet gewiß abzählbar ist, haben wir zugleich mit dem Hauptsatz bewiesen:

Satz von Löwenheim-Skolem. *Jede semantisch (und daher erst recht syntaktisch) widerspruchsfreie elementare Theorie besitzt ein höchstens abzählbares Modell.*

Die häufig als „Skolemsches Paradoxon" bezeichnete Anwendung dieses Satzes auf elementar formalisierte Systeme der Mengenlehre, in denen ja axiomatisch die Existenz überabzählbarer Mengen gefordert werden kann, besagt in Wirklichkeit nur: Eine im Sinne eines Modells der Mengenlehre überabzählbare Menge M kann im metatheoretischen Sinne durchaus abzählbar sein. Ihre Überabzählbarkeit im Modell ist dann einfach dadurch verursacht, daß keine der metatheoretisch existierenden eineindeutigen Abbildungen von M auf eine im Sinne des Modells abzählbare Menge als Menge von geordneten Paaren im Modell vorhanden ist.

Durch eine ähnliche Konstruktion, wie sie hier zum Beweis des Hauptsatzes und des Satzes von Löwenheim-Skolem benutzt wurde, kann man auch den Satz von Tarski (vgl. 4.4.) beweisen. Dazu ist im wesentlichen die Mächtigkeit der zur ursprünglichen Sprache $\boldsymbol{S}$ hinzugefügten Menge von Konstantensymbolen zu vergrößern, was dann freilich ein Arbeiten mit überabzählbaren Alphabeten erfordert.

7. Algorithmen

7.1. Kodierungen

Effektive Umformungs-, Berechnungs- und Entscheidungsprozesse an mathematischen (und zuweilen auch an außermathematischen) Objekten werden meist nicht direkt an diesen Objekten, sondern an gewissen Bezeichnungen für diese Objekte durchgeführt. Unter mathematischen Objekten verstehen wir hier z. B. natürliche, ganze, rationale Zahlen, Folgen, Funktionen und Mengen, unter effektiven Prozessen z. B. die Ausführung der vier Grundrechenarten, die Bestimmung des größten gemeinsamen Teilers und des kleinsten gemeinsamen Vielfachen von natürlichen Zahlen, die Superposition von Funktionen, die Entscheidung, ob eine gegebene natürliche Zahl eine Primzahl ist, usw. Die verwendeten Bezeichnungen für Objekte eines bestimmten Bereichs haben dabei entweder von vornherein den Charakter von Wörtern über einem endlichen Alphabet, oder man kann sie durch geringfügige Änderungen auf diese Form bringen. Das erste trifft z. B. auf die Darstellung der natürlichen Zahlen im Dezimal- oder Dualsystem zu. (Das benutzte Alphabet besteht hier aus den „Buchstaben" 0, 1, ..., 9 bzw. aus den beiden Grundzeichen 0, 1 (oder 0, L)). Das letzte ist z. B. dann der Fall, wenn man Brüche $\frac{m}{n}$, die zunächst keine Wörter sind, in der gleichwertigen Form m:n als Wörter schreibt. (Dabei wird vorausgesetzt, daß m bzw. n wortartige Darstellungen von natürlichen Zahlen sind.) Indem man jeden Teilterm der Form $\frac{t_1}{t_2}$ durch die Zeichenkombination $(t_1:t_2)$ ersetzt, kann man jeden rationalen Term „linearisieren" und die so erhaltenen Wörter als Darstellungen der abstrakten mathematischen Objekte rationale Funktionen benutzen. Wenn wir voraussetzen, daß wir bereits wortartige Darstellungen für die Elemente eines Ringes besitzen, so kann man die Polynome mit Koeffizienten aus diesem Ring, die in der üblichen Schreibweise zunächst keine Wörter sind, ohne

Informationsverlust in der Form

$$\mathrm{a}_0 + \mathrm{a}_1\mathrm{x} + \mathrm{a}_2\mathrm{xx} + \mathrm{a}_3\mathrm{xxx} + \cdots + \mathrm{a}_n\underbrace{\mathrm{x} \ldots \mathrm{x}}_{n}$$

als Wörter über einem Alphabet $\mathrm{A} \cup \{+, \mathrm{x}\}$ darstellen, wobei A das zur Darstellung der Koeffizienten benutzte Alphabet ist, das natürlich die Zeichen $+$ und x nicht enthalten darf.

Die Untersuchung vieler Beispiele lehrt, daß die Zuordnung zwischen den bezeichneten (abstrakten) Objekten und den sie bezeichnenden Wörtern nicht eineindeutig zu sein braucht, daß es vielmehr genügt, wenn jedes Wort, das in dem betreffenden Zusammenhang überhaupt als Bezeichnung eines Objektes auftritt, eine eindeutig bestimmte Bedeutung hat.

Definition 1. Ist M eine beliebige nichtleere Menge und A ein höchstens abzählbares Alphabet, so heißt eine Abbildung k aus der Menge $W(A)$ aller Wörter über dem Alphabet A auf die Menge M eine *Kodierung der Menge M im Alphabet A*. Ist k eine solche Kodierung und gilt für ein $x \in M$ und ein $\mathrm{W} \in W(A)$ die Beziehung $k(\mathrm{W}) = x$, so heißt *x die Bedeutung von* W *(bezüglich k) und* W *ein Kodewort von x (bezüglich k)*.

Der durch Definition 1 erfaßte allgemeinste Kodierungsbegriff umfaßt insbesondere folgende Fälle:

1. k ist eine eineindeutige Abbildung von $W(A)$ auf M, d. h., jedes Wort des benutzten Alphabets hat eine Bedeutung, und die Zuordnung zwischen Kodewörtern und ihren Bedeutungen ist umkehrbar eindeutig. Wichtigstes Beispiel für diesen Fall ist die durch $k_0(\mathrm{W}) =$ *Länge von* W definierte eineindeutige Abbildung k_0 von der Menge $W(\{|\})$ aller Wörter über einem einelementigen Alphabet auf die Menge der natürlichen Zahlen. Hier ist also die Zahl n die Bedeutung des Wortes $\underbrace{||\ldots|}_{n}$ und dieses Wort das Kodewort der Zahl n. Insbesondere entsprechen einander die Zahl 0 und das leere Wort. Die so definierte Kodierung k_0 der natürlichen Zahlen werden wir im folgenden als *primitive Kodierung der natürlichen Zahlen* bezeichnen und unter Berufung auf sie häufig die (abstrakten) natürlichen Zahlen mit ihren (konkreten bzw. „handhabbaren“) Kodewörtern bezüglich k_0 identifizieren.

2. k ist noch eineindeutig, aber nicht alle Wörter über dem benutzten Alphabet haben eine Bedeutung, d. h., der Definitionsbereich von k ist eine echte Teilmenge einer Menge $W(A)$. Hierunter fällt z. B. die bereits erwähnte Kodierung der natürlichen Zahlen im Dual- oder Dezimal- (oder einem anderen Positions-) System unter der Bedingung, daß das Auftreten überzähliger Nullen am Anfang der Kodewörter ausgeschlossen wird. Die entsprechende Dualkodierung im Alphabet $\{0, 1\}$ bezeichnen wir mit k_2, die Dezimalkodierung im Alphabet $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ mit k_{10}. Es ist also $k_2(1101)$ die Zahl *dreizehn*, $k_{10}(1101)$ die Zahl *tausendeinhundertundeins*, $k_2(01)$ und $k_{10}(01)$ sind nicht definiert.

3. Alle Wörter des benutzten Alphabets haben eine Bedeutung, aber die Kodierung ist nicht eineindeutig. Dies trifft z. B. auf die Dual- und Dezimalkodierung (und allgemein auf beliebige Positionskodierungen) zu, falls das Auftreten überzähliger Nullen am Anfang der Kodewörter erlaubt wird. Die so modifizierte Dual- bzw. Dezimalkodierung bezeichnen wir im folgenden mit k_2' bzw. k_{10}'. Es ist also z. B. $k_2'(01) = k_2'(001) = k_2'(1)$ die Zahl *eins*.

4. (Allgemeiner Fall) k ist nicht eineindeutig, und nicht alle Wörter des benutzten Alphabets haben eine Bedeutung. Als Beispiel betrachten wir die durch

$$k(\mathrm{W}) := \begin{cases} k'_{10}(\mathrm{W}), \textit{ falls } k'_{10}(\mathrm{W}) \textit{ definiert ist,} \\ \textit{die zu } k'_{10}(\mathrm{V}) \textit{ entgegengesetzte Zahl, falls } \mathrm{W} = -\mathrm{V} \textit{ und } k'_{10}(\mathrm{V}) \textit{ definiert ist} \end{cases}$$

definierte Kodierung der ganzen Zahlen im Alphabet $\{-, 0, 1, 2, \ldots, 9\}$. Kodewörter sind hier genau diejenigen Wörter des benutzten Alphabets, die das Zeichen „—" höchstens an erster Stelle enthalten.

Der Leser wird vielleicht schon selbst bemerkt haben, daß es recht schwierig ist, ohne Benutzung irgendeiner Kodierung über so abstrakte Objekte wie z. B. natürliche Zahlen zu sprechen. Die bei der Erläuterung der Kodierung k_2 und k_{10} benutzten Wörter *dreizehn* und *tausendeinhundertundeins* sind natürlich auch Kodewörter, allerdings in bezug auf eine kaum für alle natürlichen Zahlen exakt definierbare „Kodierung". Relativ häufig tritt jedoch in der Mathematik der Fall auf, daß eine neue Kodierung unter Benutzung einer bereits gegebenen Kodierung definiert wird. Ist allgemein k_1 eine Kodierung einer Menge M im Alphabet $\boldsymbol{A}$ und k_2 eine Kodierung der Menge $W(\boldsymbol{A})$ im Alphabet $\boldsymbol{B}$, so ist die Verkettung $k_1 \circ k_2$ der beiden Abbildungen offenbar eine Kodierung der Menge M im Alphabet $\boldsymbol{B}$. Diesen Prozeß der Gewinnung einer neuen Kodierung aus einer bereits vorhandenen unter Benutzung einer Kodierung der die ursprünglichen Kodewörter umfassenden Wortmenge bezeichnet man als *Umkodierung*. Unter den hierzu benötigten Kodierungen von Wortmengen $W(\boldsymbol{A})$ spielen die „buchstabenweisen" Kodierungen eine besondere Rolle, die sich durch verkettungshomomorphe Fortsetzung einer Kodierung des Alphabets $\boldsymbol{A}$ erzeugen lassen. Es sei $\boldsymbol{A} = \{\mathrm{a}_1, \ldots, \mathrm{a}_m\}$ ein beliebiges Alphabet. (Aus dem folgenden geht hervor, daß $\boldsymbol{A}$ auch abzählbar sein kann.) Ferner sei $\mathrm{W}_1, \ldots, \mathrm{W}_m \in W(\boldsymbol{B})$ ein System von *Pseudoatomen* in $W(\boldsymbol{B})$, d. h., jedes überhaupt als Verkettung der Wörter $\mathrm{W}_1, \ldots,$ W_m darstellbare Wort $\mathrm{W} \in W(\boldsymbol{B})$ läßt sich auf genau eine Weise so erzeugen. Dann ist

$$k(\mathrm{W}_\mu) := \mathrm{a}_\mu \qquad (\mu = 1, \ldots, m)$$

eine Kodierung von $\boldsymbol{A}$ im Alphabet $\boldsymbol{B}$ und die durch

$$k'(\Lambda) := \Lambda,$$

$$k'(\mathrm{W}_{i_1} \circ \cdots \circ \mathrm{W}_{i_k}) := k(\mathrm{W}_{i_1}) \circ \cdots \circ k(\mathrm{W}_{i_k}) = \mathrm{a}_{i_1} \ldots \mathrm{a}_{i_k}$$

definierte Abbildung k' aus $W(\boldsymbol{B})$ auf $W(\boldsymbol{A})$ eine eineindeutige Kodierung von $W(\boldsymbol{A})$ im Alphabet $\boldsymbol{B}$. Da offenbar in einem Alphabet $\boldsymbol{B}$, das wenigstens die beiden Buchstaben a, b enthält, die Wörter der Form $\mathrm{b}\underbrace{\mathrm{a}\ldots\mathrm{a}}_{n}\mathrm{b}$ ($n \geqq 0$) ein abzählbares System von Pseudoatomen bilden, kann man jede Wortmenge $W(A)$ über einem höchstens abzählbaren Alphabet A in einem zweielementigen Alphabet $\{\mathrm{a}, \mathrm{b}\}$ kodieren. Dabei ist dann $\mathrm{b}\underbrace{\mathrm{a}\ldots\mathrm{a}}_{i_1}\mathrm{bb}\underbrace{\mathrm{a}\ldots\mathrm{a}}_{i_2}\mathrm{b}\ldots\mathrm{b}\underbrace{\mathrm{a}\ldots\mathrm{a}}_{i_k}\mathrm{b}$ das eindeutig bestimmte Kodewort von $\mathrm{a}_{i_1}\mathrm{a}_{i_2}\ldots\mathrm{a}_{i_k} \in W(A)$.

Wir stellen nun einige grundlegende Sachverhalte über Kodierbarkeit von Mengen zusammen:

Da eine kodierbare Menge immer als Bild einer abzählbaren Menge $W(A)$ (vgl. Kapitel 1, Sätze 2 und 3) bei einer eindeutigen Abbildung (nämlich der Kodierung) erscheint, *ist jede kodierbare Menge höchstens abzählbar*. Umgekehrt kann es also für eine überabzählbare Menge (etwa die Menge der reellen Zahlen) keine einheitliche Kodierung im hier definierten Sinne geben.

Jede überhaupt kodierbare Menge ist schon in einem zweielementigen Alphabet kodierbar. Insbesondere ist jede eineindeutig kodierbare Menge schon in einem zweielementigen Alphabet eineindeutig kodierbar. Man hat hierzu nur die nach Voraussetzung existierende Kodierung in der oben beschriebenen Weise unter Benutzung eines zweielementigen Alphabets umzukodieren.

Eine Kodierung einer beliebigen Menge M in einem einelementigen Alphabet bezeichnet man als eine *Numerierung* von M. Dabei werden die als Kodewörter auftretenden Wörter häufig mit den ihnen umkehrbar eindeutig entsprechenden natürlichen Zahlen identifiziert, so daß man eine Numerierung der Menge M auch als eine Abbildung aus der Menge $\mathbf{N}$ der natürlichen Zahlen auf die Menge M definieren kann. Umgekehrt bezeichnet man eine eineindeutige Abbildung g von M in die Menge $\mathbf{N}$ der natürlichen Zahlen als eine *Gödelisierung* von M. Gibt es eine solche Abbildung g, so heißt die Menge M *gödelisierbar*, und für $x \in M$ wird die Zahl $g(x)$ als *Gödelzahl* von x (bezüglich g) bezeichnet. Offenbar ist dann g^{-1} eine eineindeutige Numerierung von M. Demnach ist eine Menge M genau dann gödelisierbar, wenn sie eineindeutig numerierbar ist. (In der Literatur wird der Terminus Gödelisierung, der historisch älter ist als die Begriffe Numerierung und Kodierung, nicht einheitlich, insbesondere häufig im Sinne von Numerierung einer Menge $W(A)$ verwendet.)

Jede Wortmenge $W(A)$ über einem höchstens abzählbaren Alphabet ist gödelisierbar. (Folglich ist jede überhaupt kodierbare Menge numerierbar und jede eineindeutig kodierbare Menge gödelisierbar.)

Zum Beweis geben wir zwei häufig verwendete Gödelisierungen an.

1. Es sei $A = (\mathrm{a}_1, \ldots, \mathrm{a}_n)$ ein endliches geordnetes Alphabet ($n \geqq 2$). Offenbar erhält man eine Gödelisierung der Menge $W(\mathrm{A})$, wenn man sie in der durch die gegebene Ordnung des Alphabets induzierten lexikographischen Ordnung aufzählt (vgl.

Kapitel 1): Die so entstehende Gödelisierung bezeichnen wir im folgenden mit g_{lex}:

$$\begin{array}{llllll} W(A)\colon \Lambda,\ \mathfrak{a}_1, \ldots, \mathfrak{a}_n,\ \mathfrak{a}_1\mathfrak{a}_1, & \ldots, \mathfrak{a}_1\mathfrak{a}_n,\ \mathfrak{a}_2\mathfrak{a}_1, & \ldots, \mathfrak{a}_2\mathfrak{a}_n, \ldots, \mathfrak{a}_n\mathfrak{a}_n, & \mathfrak{a}_1\mathfrak{a}_1\mathfrak{a}_1, & \ldots \\ g_{\mathrm{lex}}\colon\ 0,\ 1, \ldots, n,\ n+1, & \ldots, 2n,\ 2n+1, & \ldots, 3n,\ \ldots, n^2+n, & n^2+n+1, & \ldots \end{array}$$

Prinzipiell kann man zu vorgegebenem $\mathrm{W} \in W(A)$ die Gödelzahl $g_{\mathrm{lex}}(\mathrm{W})$ und umgekehrt zu jeder natürlichen Zahl n das Wort $g_{\mathrm{lex}}^{-1}(n)$ effektiv berechnen, indem man die Wörter aus $W(A)$ in der lexikographischen Anordnung genügend weit aufschreibt und durchzählt. Wir wollen jedoch eine „direkte" Formel für $g_{\mathrm{lex}}(\mathfrak{a}_{i_1} \ldots \mathfrak{a}_{i_m})$ ableiten, indem wir die Anzahl der diesem Wort lexikographisch vorangehenden Wörter abzählen. Das sind:

1 Wort der Länge 0,

n Wörter der Länge 1,

$\ldots,$

n^{m-1} Wörter der Länge $m-1$,

ferner unter den gleichlangen Wörtern der Länge m

$(i_1 - 1)\, n^{m-1}$ Wörter, die mit einem der Buchstaben $\mathfrak{a}_1, \ldots, \mathfrak{a}_{i_1-1}$ beginnen,

$(i_2 - 1)\, n^{m-2}$ Wörter, die mit $\mathfrak{a}_{i_1}\mathfrak{a}_j$ beginnen, wobei $j < i_2$ ist,

$\ldots$

$(i_{m-1} - 1)\, n$ Wörter, die mit $\mathfrak{a}_{i_1}\mathfrak{a}_{i_2} \ldots \mathfrak{a}_{i_{m-2}}\mathfrak{a}_j$ beginnen, wobei $j < i_{m-1}$ ist,

$(i_m - 1)$ Wörter der Form $\mathfrak{a}_{i_1} \ldots \mathfrak{a}_{i_{m-1}}\mathfrak{a}_j$ mit $j < i_m$.

Dem Wort $\mathfrak{a}_{i_1} \ldots \mathfrak{a}_{i_m}$ gehen also insgesamt $\sum_{k=1}^{m} i_k n^{m-k}$ Wörter lexikographisch voran. Daher ist

$$g_{\mathrm{lex}}(\mathfrak{a}_{i_1} \ldots \mathfrak{a}_{i_m}) = \sum_{k=1}^{m} i_k n^{m-k}$$

2. Es sei $A = \{\mathfrak{a}_1, \mathfrak{a}_2, \ldots\}$ ein endliches oder abzählbares Alphabet. Für $n = 1, 2, 3, \ldots$ sei p_n die n-te Primzahl in bezug auf deren natürliche Reihenfolge. Wir definieren

$$g_{\mathrm{prim}}(\Lambda) := 0,$$

$$g_{\mathrm{prim}}(\mathfrak{a}_{i_1} \ldots \mathfrak{a}_{i_m}) := p_1^{i_1} p_2^{i_2} \ldots p_m^{i_m}.$$

Die so definierte Gödelisierung g_{prim} hat gegenüber den Gödelisierungen der Form g_{lex} den Vorteil, daß sie sich unabhängig von der Mächtigkeit des höchstens abzählbaren Alphabets A definieren läßt und insbesondere auch noch im Fall abzählbarer Alphabete verwendbar ist. Sie hat jedoch gegenüber den Gödelisierungen g_{lex} den

Nachteil, daß nur „sehr wenige“ natürliche Zahlen als Gödelzahlen auftreten, nämlich außer 0 genau diejenigen, die mit einem beliebigen Primfaktor alle kleineren Primzahlen als Faktor enthalten (und für die außerdem, falls das Alphabet A endlich ist und aus n Buchstaben besteht, jeder Primfaktor höchstens in der n-ten Potenz vorkommt). Sowohl g_{prim} als auch g_{lex} samt ihren inversen Abbildungen sind in einem noch unpräzisierten Sinne aber unzweifelhaft „effektiv berechenbar“.

Eine Abbildung, deren Definitionsbereich eine Menge von Wörtern oder allgemeiner eine Menge von n-Tupeln von Wörtern und deren Wertebereich ebenfalls eine Menge von Wörtern ist, bezeichnen wir im folgenden als eine *Wortfunktion.* Dabei genügt es grundsätzlich, sich auf die Betrachtung von Abbildungen aus einer Menge $W(A)$ in sich zu beschränken: Sind $A_0, A_1, \ldots, A_n$ beliebige (nicht notwendig verschiedene oder gar disjunkte) Alphabete, so sei $*$ ein beliebiger, in keinem dieser Alphabete vorkommender Buchstabe und

$$A = \bigcup_{\nu=0}^{n} A_\nu \cup \{*\}.$$

Die Betrachtung einer n-stelligen Abbildung f aus $W(A_1) \times \cdots \times W(A_n)$ in $W(A_0)$ kann offenbar durch die Betrachtung der durch

$$f'(W_1 * W_2 * \cdots * W_n) := f(W_1, W_2, \ldots, W_n)$$

definierten Abbildung f' aus $W(A)$ in sich ersetzt werden. Abbildungen aus einer Menge $W(A)$ in sich werden wir kurz als *Wortfunktionen in* $W(A)$ bezeichnen.

Definition 2. Es sei M eine beliebige nichtleere Menge, k eine Kodierung von M in einem Alphabet A und φ eine Abbildung aus M in sich. (Allgemeinere Fälle kann man wieder durch entsprechende Strukturierung der Menge M einbeziehen.) Eine Wortfunktion f in $W(A)$ heißt eine *Darstellung von* φ *bezüglich* k, wenn für jedes Wort $W \in W(A)$ das Ding $\varphi(k(W))$ genau dann existiert, wenn $k(f(W))$ existiert und im Fall der Existenz beide gleich sind.

Die Kenntnis einer Darstellung einer Abbildung bedeutet, daß man das Operieren mit den Objekten der abstrakten Menge in diesem Zusammenhang durch ein konkreteres Operieren mit Kodewörtern ersetzen kann. Ist insbesondere die zugrunde liegende Kodierung k eineindeutig, so ist die Darstellung f einer Abbildung φ durch die Forderung $\varphi \circ k = k \circ f$ eindeutig als $f = k^{-1} \circ \varphi \circ k$ bestimmt, und es ist dann umgekehrt $\varphi = k \circ f \circ k^{-1}$, d. h., um in diesem Fall für ein $x \in M$ das Bildobjekt $\varphi(x)$ zu erhalten, nehme man das Kodewort $k^{-1}(x)$, wende die Wortfunktion f auf dieses Wort an und bestimme die Bedeutung $k(f(k^{-1}(x)))$ des erhaltenen Wortes. Klar ist auch, daß eine Abbildung bezüglich einer nichteineindeutigen Kodierung im allgemeinen viele Darstellungen besitzt.

Beispiel. Es sei $M = \mathbf{N}$ die Menge der natürlichen Zahlen und φ die in 4.5. betrachtete Nachfolgeroperation in $\mathbf{N}$. Bezüglich der eineindeutigen primitiven

Kodierung k_0 von $\mathbf{N}$ im Alphabet $A = \{a\}$ wird φ durch die sehr einfache Wortfunktion $f(W) := Wa$ $(W \in W(\{a\}))$ dargestellt. Die offenbare „effektive Berechenbarkeit" der Wortfunktion f rechtfertigt es, die durch sie dargestellte Funktion φ ebenfalls als „effektiv berechenbar" anzusehen, obwohl das unmittelbare Operieren mit — etwa als Äquivalenzklassen von gleichmächtigen endlichen Mengen aufgefaßten — natürlichen Zahlen sicher nicht „effektiv ausführbar" ist.

Bezüglich der nicht eineindeutigen Dualkodierung k_2' der natürlichen Zahlen im Alphabet $\{0, 1\}$ wird die Nachfolgerfunktion φ unter anderem durch die wie folgt induktiv definierte Wortfunktion f in $W(\{0, 1\})$ dargestellt:

$$f(\Lambda\ ) := 1,$$
$$f(W0) := W1,$$
$$f(W1) := f(W)\,0.$$

Legt man für die Kodierung k_2' etwa die induktive Definition

$$k_2'(\Lambda)\ := null,$$
$$k_2'(W0) := 2 \cdot k_2'(W),$$
$$k_2'(W1) := k_2'(W0) + eins$$

zugrunde, so läßt sich durch Induktion über die Wortlänge leicht zeigen, daß, wie gefordert,

$$k_2'(f(W)) = \varphi(k_2'(W)) \qquad \text{für } W \in W(\{0, 1\})$$

gilt. Für jede natürliche Zahl $n \geqq 1$ sei nun

$$f_n(W) := \underbrace{0 \ldots 0}_{n} f(W).$$

Dann ist für jedes Wort $W \in W(\{0, 1\})$ offenbar $k_2'(f(W)) = k_2'(f_n(W))$ und daher $\varphi(k_2'(W)) = k_2'(f(W)) = k_2'(f_n(W))$, d. h., zugleich mit f sind unter anderem auch alle Wortfunktionen f_n Darstellungen von φ bezüglich der Kodierung k_2'.

7.2. Der anschauliche Algorithmenbegriff, formalisierte Flußdiagramme und deren Interpretation

Das Wort Algorithmus ist durch Verstümmelung (Latinisierung) aus dem Namen des mittelasiatischen Mathematikers MUHAMMAD IBN MUSA „AL-CHORESMI" (d. h. der Choresmier[1])) entstanden, der um 825 in Bagdad wirkte und in seinen Büchern

[1]) In mathematischen und mathematikhistorischen Büchern meist als AL-KHWARIZMI, AL-HUARISMI u. ä. geschrieben. Zur hier benutzten Schreibweise und Deutung des Namens vgl. B. BRENTJES, Die orientalische Welt, DVW, Berlin 1970, S. 356.

unter anderem die Ausführung der Grundrechenarten unter Benutzung des indisch-arabischen Dezimalsystems lehrte, die in dieser Form zuerst durch seine Bücher in Europa bekannt wurde. Demgemäß bedeutete das „Rechnen nach dem Algorithmus" ursprünglich die Anwendung der jedem Schüler geläufigen elementaren dezimalen Rechenverfahren. Im Laufe der Jahrhunderte wurde die Bedeutung des Wortes Algorithmus immer mehr verallgemeinert. Heute versteht man unter einem *Algorithmus* ein schrittweise ablaufendes (Berechnungs-, Entscheidungs-, allgemein Umformungs-) Verfahren für Objekte eines bestimmten Bereichs, das folgenden Bedingungen genügt:

(a) Das Verfahren ist in endlich vielen Regeln unmißverständlich formuliert.

(b) Es gibt eine eindeutig bestimmte Regel, die als erste anzuwenden ist.

(c) Nach Anwendung einer Regel steht fest, ob das Verfahren damit beendet ist bzw. welche Regel als nächste anzuwenden ist.

(d) Jede Regel ist im gegebenen Zusammenhang „effektiv ausführbar" bzw. ohne Anwendung von Intelligenz rein mechanisch realisierbar.

Die genauere Untersuchung vorliegender Algorithmen lehrt, daß man durch geeignete Abänderung der Formulierung jeden Algorithmus unter alleiniger Benutzung zweier spezieller Arten von Regeln formulieren kann, die als *Operations-* bzw. *Prüfregeln* bezeichnet werden. Eine Operationsregel verlangt die Bildung eines neuen Zwischenresultats aus gewissen zu diesem Zeitpunkt vorliegenden Zwischenresultaten bzw. Eingabegrößen. Die als nächste anzuwendende Regel wird, unabhängig vom Zwischenresultat, in der Regel angegeben. Eine Prüfregel verlangt die Entscheidung, welche Regel als nächste anzuwenden ist, in Abhängigkeit davon, ob gewisse zu diesem Zeitpunkt vorliegende Zwischenresultate bzw. Eingabegrößen in gewissen Relationen zueinander stehen bzw. gewisse Eigenschaften haben. Ein neues Zwischenresultat wird bei Anwendung einer Prüfregel nicht gebildet. Die Art und Weise der Aufeinanderfolge der einzelnen Regeln eines Algorithmus läßt sich einfach und zugleich exakt durch einen Graphen darstellen, den man als den Strukturgraphen des betreffenden Algorithmus bezeichnet. Der Definition des Begriffs Strukturgraph stellen wir in aller Kürze einige allgemeine graphentheoretische Definitionen voran.

Definition 1. Ein *endlicher gerichteter Graph G* ist eine endliche nichtleere Menge G von geordneten Paaren (x, y), die als die *Kanten* von G bezeichnet werden. Ein Ding x heißt ein *Knoten* von G, wenn es ein y gibt, so daß $(x, y) \in G$ oder $(y, x) \in G$ ist. Mit $K(G)$ bezeichnen wir die Menge aller Knoten von G. Daher ist stets $G \subseteqq K(G)^2$. Ist $(x, y) \in G$, so heißt *x ein Vorgänger von y* und *y ein Nachfolger von x*. Im folgenden werden nur Graphen betrachtet, in denen jeder Knoten höchstens zwei Nachfolger hat. Knoten mit genau einem Nachfolger heißen dann *Arbeitsknoten*, Knoten mit genau zwei Nachfolgern heißen *Prüfknoten*, und Knoten ohne Nachfolger heißen *Ausgänge*. Sind $x_1, \ldots, x_n$ $(n \geqq 2)$ Knoten von G, so daß $(x_i, x_{i+1}) \in G$ für $i = 1, \ldots,$

$n - 1$ gilt, so heißt die Folge $x_1, \ldots, x_n$ ein *Weg in G von x_1 nach x_n*. Existiert ein solcher Weg, so heißt *x_n von x_1 erreichbar*. Ein Weg heißt *einfach*, wenn seine Knoten paarweise verschieden sind. Ein Weg von x nach x heißt ein *Zyklus*. Ein Graph, in dem es keinen Zyklus gibt, heißt *zyklenfrei*.

Definition 2. Ein Paar (G, e) heißt ein *Strukturgraph*, wenn gilt:

(a) G ist ein Graph und e einer seiner Knoten, der kein Ausgang ist. (e heißt *Eingangsknoten* oder kurz *Eingang* von (G, e).)

(b) Jeder Knoten von G hat höchstens zwei Nachfolger.

(c) Jeder von e verschiedene Knoten ist von e erreichbar. Von jedem Knoten, der kein Ausgang ist, ist wenigstens ein Ausgang erreichbar.

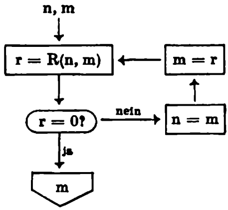
(d) Die beiden Nachfolger eines Prüfknotens sind von diesem und voneinander verschieden.

Ein Strukturgraph wird üblicherweise graphisch dargestellt, indem man jedem Knoten des Graphen ein umrandetes Feld zuordnet, so daß diese Felder sich nicht überlappen. Jede Kante $(x, y) \in G$ wird durch einen Pfeil dargestellt, die von dem x zugeordneten Feld zu dem y zugeordneten Feld führt. Der Eingangsknoten wird durch einen in ihn hineinführenden Pfeil markiert. Es ist üblich (und im Bereich technischer Anwendungen sogar durch TGL geregelt), die verschiedenen Knotenarten zusätzlich durch unterschiedliche Umrahmung zu kennzeichnen, insbesondere Arbeitsknoten rechteckig, Prüfknoten abgerundet. Jedoch ist dies prinzipiell entbehrlich, da die Knotenart bereits aus der Zahl der Nachfolger hervorgeht.

Aus einem Strukturgraphen entsteht ein *Flußdiagramm*, indem man an den Eingangspfeil endlich viele Variablen für Eingabegrößen schreibt, für jeden Prüfknoten die beiden aus ihm herausführenden Pfeile durch „ja" und „nein" oder eine gleichwertige Kennzeichnung unterscheidet und unter Benutzung der Eingabevariablen sowie weiterer Variablen für die Zwischenresultate in jeden Arbeitsknoten die dort auszuführende Operation und in jeden Prüfknoten die dort zu entscheidende Relation hineinschreibt. Je nachdem, ob das Flußdiagramm einen Berechnungs- oder einen Entscheidungsprozeß realisieren soll, können die Ausgangsknoten entweder mit den Variablen für die als Endresultat angesehenen Zwischenresultate oder mit gewissen „Antworten", im einfachsten Fall „ja" und „nein" beschriftet werden. Eine ganz naive Darstellung des inhaltlich als bekannt vorausgesetzten Euklidischen Algorithmus zur Bestimmung des größten gemeinsamen Teilers zweier natürlicher Zahlen könnte demnach wie (1) auf S. 150 aussehen, wobei R die durch

$$\mathrm{R}(\mathrm{n}, \mathrm{m}) := \iota \mathrm{x}\big(0 \leqq \mathrm{x} \wedge \mathrm{x} < \mathrm{m} \wedge \bigvee \mathrm{q}(\mathrm{n} = \mathrm{qm} + \mathrm{x})\big), \textit{ falls } \mathrm{m} \neq 0$$

im Bereich der natürlichen Zahlen definierte partielle Operation bezeichnen soll, die dem Zahlenpaar n, m den Rest bei Division von n durch m zuordnet. Der Leser mache sich klar, daß das Flußdiagramm (1) unverständlich bleibt, wenn man nicht weiß, daß n, m, r Variablen für natürliche Zahlen sind, welche Operation durch R bezeichnet wird und welche spezielle natürliche Zahl das Symbol 0 darstellt. So ergibt sich, daß ein Flußdiagramm in ähnlicher Weise wie ein Ausdruck einer formalisierten Sprache zunächst ein bedeutungsloses syntaktisches Gebilde ist, das im allgemeinen viele verschiedene Interpretationen gestattet. Die folgenden Definitionen stellen den Zusammenhang zwischen Formalismus und möglicher Bedeutung für Flußdiagramme (in Analogie zur Definition der Begriffe Interpretation, Belegung und Wert für Ausdrücke einer formalisierten Sprache) her.



(1)

Definition 3. Es sei S eine formalisierte Sprache, (G, e) ein Strukturgraph, ψ eine Abbildung, die jedem Prüfknoten von G einen seiner beiden Nachfolger zuordnet, φ eine Abbildung, die jedem Prüfknoten von G einen Ausdruck $\mathrm{H} \in S$ und jedem Arbeitsknoten von G eine Termgleichung der Form $\mathbf{x}_i = \mathbf{x}_j$ oder $\mathbf{x}_i = \mathrm{c}$ oder $\mathbf{x}_i = \mathrm{F}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ der Sprache S zuordnet. Dann heißt das System (G, e, ψ, φ) *ein formalisiertes Flußdiagramm über* S.

Beispiel 1. Die Sprache S sei einsortig und enthalte unter anderem die Variablen n, m, r, das zweistellige Operationssymbol R und das Konstantensymbol 0. Der dem bereits diskutierten Flußdiagramm (1) zugrunde liegende Strukturgraph kann in der abstrakten mengentheoretischen Form z. B. als

$$\big(\{(e, b), (b, a), (b, c), (c, d), (d, e)\}, e\big)$$

notiert werden. Hierin ist b der einzige Prüfknoten, und es sei $\psi(b) := a$ (der „ja"-Nachfolger von b). Ferner sei durch φ dem Prüfknoten b der Ausdruck $\mathrm{r} = 0$ und den Arbeitsknoten e, c, d die Termgleichungen $\mathrm{r} = \mathrm{R}(\mathrm{n}, \mathrm{m})$ bzw. $\mathrm{n} = \mathrm{m}$ bzw. $\mathrm{m} = \mathrm{r}$ zugeordnet. Stellt man den angegebenen Strukturgraphen zeichnerisch dar, schreibt die durch φ vorgegebenen Ausdrücke in die entsprechenden Knoten hinein und beschriftet den Pfeil zwischen einem Prüfknoten x und seinem Nachfolger $\psi(x)$ durch „ja", so erhält man ein Flußdiagramm im anschaulichen Sinne, insbesondere in unserem Beispiel das bereits gezeichnete Diagramm (1) bis auf die Beschriftungen des Eingangspfeils und der Ausgänge.

Definition 4. Es sei $(G, e, \psi, \varphi) = \mathfrak{F}$ ein formalisiertes Flußdiagramm über S, ω eine Interpretation von S und f eine *partielle Belegung bezüglich* ω, d. h. eine Abbildung, die einigen Variablen der Sprache S je ein Ding entsprechender Sorte aus der Struktur ω zuordnet. Wir definieren induktiv die (endliche oder abzählbare) Folge $k_1, k_2, k_3, \ldots$ der Knoten von G, die bei Abarbeitung der Anfangsbelegung f in G der Reihe nach durchlaufen werden, und die gleichlange Folge $f_1, f_2, f_3, \ldots$ der Belegungen, die dabei als Zwischenresultate auftreten:

(a) $k_1 := e, f_1 := f$.

Es seien schon der Knoten k_n von G definiert und die Belegung f_n, mit der man „in diesen Knoten hineingeht".

(b) k_n ist ein Arbeitsknoten mit $\varphi(k_n) = \mathbf{x}_i = \mathrm{t}$ (wobei t ein Term der Form $\mathbf{x}_j$ oder c oder $\mathrm{F}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ ist. Falls $Wert(\mathrm{t}, \omega, f_n)$ existiert, d. h., falls f_n für alle in t vorkommenden Variablen definiert und gegebenenfalls die Operation $\omega(\mathrm{F})$ an der betreffenden Stelle definiert ist, sei k_{n+1} der Nachfolger von k_n und

$$f_{n+1}(\mathbf{x}) := \begin{cases} Wert(\mathrm{t}, \omega, f_n) \textit{ für } \mathbf{x} = \mathbf{x}_i, \\ f_n(\mathbf{x}), \textit{falls } f_n(\mathbf{x}) \textit{ definiert und } \mathbf{x} \neq \mathbf{x}_i, \\ \textit{nicht definiert für die übrigen Variablen.} \end{cases}$$

Falls $Wert(\mathrm{t}, \omega, f_n)$ nicht existiert, sind k_{n+1} und f_{n+1} nicht definiert.

(c) k_n ist ein Prüfknoten mit $\varphi(k_n) = \mathrm{H}$. Falls f_n für alle in H frei vorkommenden Variablen definiert ist, sei $f_{n+1} := f_n$ und

$$k_{n+1} := \begin{cases} \psi(k_n), \textit{ falls } Wert(\mathrm{H}, \omega, f_n) = \boldsymbol{W}, \\ \textit{der von } \psi(k_n) \textit{ verschiedene Nachfolger von } k_n, \\ \quad \textit{falls } Wert(\mathrm{H}, \omega, f_n) = \boldsymbol{F}. \end{cases}$$

Falls f_n nicht für alle in H frei vorkommenden Variablen definiert ist, sind k_{n+1} und f_{n+1} nicht definiert.

(d) k_n ist ein Ausgang. Dann sind k_{n+1} und f_{n+1} nicht definiert.

Ist $\mathfrak{F}$ ein formalisiertes Flußdiagramm über der Sprache S und ω eine Interpretation von S, so ist das Paar $(\mathfrak{F}, \omega)$ ein *konkretes Flußdiagramm.* Offenbar ändert sich nichts wesentlich, wenn man annimmt, daß S eine nichtelementare Sprache ist und nur zulässige Interpretationen dieser Sprache betrachtet. In einem konkreten Flußdiagramm spielen die Variablen etwa die Rolle von Adressen (Namen von Speicherplätzen), und die gegebene partielle Anfangsbelegung gibt an, unter welchen Adressen zu Beginn der Abarbeitung welche konkreten Objekte entsprechender Sorte aus der Struktur ω gespeichert sind. Ebenso gibt die partielle Belegung f_n den Belegungszustand des Speichers zu Beginn des n-ten Taktes der Berechnung an. Aus dieser Sicht bedeutet die Ausführung einer in einem Arbeitsknoten stehenden

Anweisung der Form $\mathfrak{x}_i = \mathrm{F}(\mathfrak{x}_1, \ldots, \mathfrak{x}_k)$: Man wende die Operation, die durch F bezeichnet wird, auf diejenigen Dinge $\xi_1, \ldots, \xi_k$ an, die zum betreffenden Zeitpunkt unter den Adressen $\mathfrak{x}_1, \ldots, \mathfrak{x}_k$ gespeichert sind, und speichere das erhaltene Resultat unter der Adresse $\mathfrak{x}_i$, wobei der alte Inhalt der Speicherzelle $\mathfrak{x}_i$, falls sie schon besetzt war, automatisch gelöscht wird. Analog sind die übrigen Knotenbeschriftungen zu verstehen. Stop ohne Resultat, d. h. Abbrechen der Folge der k_n und f_n vor Erreichen eines Ausgangsknotens, tritt ein, wenn eine der aufzurufenden Speicherzellen leer ist oder wenn ein System von aufgerufenen Speicherinhalten nicht im Definitionsbereich der in diesem Takt anzuwendenden Operation liegt. Falls für eine bestimmte Anfangsbelegung f ein Ausgangsknoten erreicht wird, wird man, je nachdem, ob das Flußdiagramm einen Berechnungs- oder einen Entscheidungsprozeß darstellt, einige Komponenten der erhaltenen Endbelegung oder den Ausgang, an dem diese erscheint, als Resultat ansehen. Im ersten Fall definiert das Flußdiagramm eine partielle Operation, die in sehr allgemeiner Weise aus den in den Arbeitsknoten vorkommenden Operationen superponiert ist. (Die gewöhnliche Superposition und die Definition durch Fallunterscheidung (vgl. 4.6.) ergeben sich offenbar als einfache Spezialfälle.) Im zweiten Fall definiert das Flußdiagramm zusammen mit einem bestimmten seiner Ausgänge eine Relation, die für Dinge $\xi_1, \ldots, \xi_n$ der Struktur ω genau dann besteht, wenn die Abarbeitung der durch $f(\mathfrak{x}_i) := \xi_i$ definierten partiellen Belegung zum gewählten Ausgang des Flußdiagramms führt. Zusammenfassend sei gesagt, daß es viele Möglichkeiten gibt, das Endresultat der Anwendung eines konkreten Flußdiagramms $(\mathfrak{F}, \omega)$ auf eine partielle Belegung f bezüglich ω zu definieren, die sich in Abhängigkeit vom jeweiligen Zusammenhang bzw. beabsichtigten Zweck mehr oder weniger eignen. Für uns ist vor allem die Feststellung wichtig, daß

(a) ein durch ein Flußdiagramm definierter Berechnungs- oder Entscheidungsprozeß immer dann effektiv ausführbar ist, wenn die in seinen Arbeitsknoten vorkommenden Operationen effektiv ausführbar und die in seinen Prüfknoten vorkommenden Relationen effektiv entscheidbar sind;

(b) umgekehrt nach unserem heutigen Wissen die Methode, als effektiv vorausgesetzte Berechnungs- und Entscheidungsprozesse durch Flußdiagramme miteinander zu verknüpfen, die allgemeinste Methode zur Erzeugung weiterer Berechnungs- und Entscheidungsprozesse ist, die im gleichen Maße effektiv ausführbar sind wie die gegebenen „Bausteine".

Hat man daher in einem bestimmten Bereich der Mathematik ein System von Operationen und Relationen gewählt, die als effektiv ausführbar bzw. entscheidbar gelten sollen, so ist damit implizit die Gesamtheit aller effektiv ausführbaren Berechnungs- und Entscheidungsalgorithmen fixiert.[1]

[1] Der so präzisierte Algorithmenbegriff ist, insbesondere in Anwendung auf die Geometrie, in [50] behandelt.

7.3. Berechenbarkeit von Wortfunktionen, Entscheidbarkeit und Aufzählbarkeit von Wortmengen, die Hypothese von Church

Während es in beliebigen mathematischen Theorien kaum möglich, ja wahrscheinlich nicht einmal wünschenswert ist, ein für allemal einen absoluten Effektivitätsbegriff für Operationen und Relationen zu definieren, hat es sich als möglich und nützlich erwiesen, einen solchen absoluten Effektivitätsbegriff im Bereich der Wortfunktionen und Wortmengen zu definieren. Da ein Umformungs- oder Entscheidungsprozeß an beliebigen Objekten durch eine Kodierung des betreffenden Bereichs häufig in einen auf Wörter anzuwendenden Prozeß übersetzt werden kann, reichen die Konsequenzen des absoluten Effektivitätsbegriffs für Wörter viel weiter, als es zunächst scheint.

In diesem Abschnitt werden wir aus der Existenz einer mathematisch exakten Definition des Begriffs *berechenbare* Wortfunktion, die alle einfachen, im anschaulichen Sinn gewiß berechenbaren Wortfunktionen umfaßt, alle für die Anwendungen in Kapitel 8 wesentlichen Folgerungen ziehen, ohne zunächst die Definition selbst anzugeben. Eine solche Definition (mittels Turingmaschinen) wird erst in 7.4. gegeben und in 7.5. mit einigen gleichwertigen anderen Definitionen verglichen. Die beiden der mehr technischen Seite der Algorithmentheorie gewidmeten Abschnitte 7.4. und 7.5. sind für das Verständnis von Kapitel 8 nicht unbedingt erforderlich.

Offenbar genügt es grundsätzlich, den Begriff der berechenbaren *arithmetischen* Funktion (d. h. Abbildung aus der Menge $\mathbf{N}$ der natürlichen Zahlen in sich) zu definieren: Ist g eine samt ihrer Inversen effektiv ausführbare Gödelisierung der Menge $W(A)$ (z. B. g_{lex} oder g_{prim}; vgl. 7.1.), so ist offenbar eine Wortfunktion f in $W(A)$ genau dann berechenbar, wenn die arithmetische Funktion

$$\varphi := g \circ f \circ g^{-1} \tag{1}$$

berechenbar ist: Sieht man f als berechenbar an, so ist auch die Hintereinanderausführung der drei berechenbaren Funktionen g^{-1}, f, g effektiv berechenbar. Umgekehrt folgt aus (1)

$$f = g^{-1} \circ \varphi \circ g,$$

so daß mit φ auch f berechenbar ist.

Weiterhin genügt es, den Begriff der entscheidbaren Wortrelation für den Fall einstelliger Relationen zu definieren: Sind $A_1, \ldots, A_n$ beliebige (nicht notwendig verschiedene) Alphabete, so sei $*$ ein in keinem dieser Alphabete vorkommender Trennbuchstabe und $A := \bigcup_{i=1}^{n} A_i \cup \{*\}$. Eine Relation

$$R \subseteq W(A_1) \times W(A_2) \times \cdots \times W(A_n)$$

ist im anschaulichen Sinn offenbar genau dann entscheidbar, wenn die einstellige Relation

$$\{W_1 * W_2 * \cdots * W_n : (W_1, W_2, \ldots, W_n) \in \boldsymbol{R}\} \tag{2}$$

eine entscheidbare Teilmenge von $W(\boldsymbol{A})$ ist. Die Definition der Entscheidbarkeit (der Zugehörigkeit eines Wortes zu) einer Teilmenge $\boldsymbol{M}$ von $W(\boldsymbol{A})$ kann wiederum leicht auf den Begriff der Berechenbarkeit von Wortfunktionen zurückgeführt werden: Ist a ein beliebiger fester Buchstabe des Alphabets $\boldsymbol{A}$, so ist $\boldsymbol{M} \subseteq W(\boldsymbol{A})$ offenbar im anschaulichen Sinne genau dann entscheidbar, wenn die durch

$$\chi_{\boldsymbol{M}}(W) := \begin{cases} \Lambda, \text{ falls } W \in \boldsymbol{M}, \\ a, \text{ falls } W \in W(\boldsymbol{A}) \setminus \boldsymbol{M} \end{cases} \tag{3}$$

definierte *charakteristische Funktion der Menge* $\boldsymbol{M}$ berechenbar ist.

Während vieler Jahrhunderte, in denen die Mathematiker sich darauf beschränkten, für bestimmte (Berechnungs- und Entscheidungs-) Aufgaben Algorithmen zu deren Lösung anzugeben, wurde die Notwendigkeit, den Begriff des Algorithmus bzw. der Entscheidbarkeit und Berechenbarkeit mathematisch exakt zu definieren, nicht empfunden, da an der effektiven Durchführbarkeit der konkreten angegebenen Verfahren nie ein Zweifel bestand. Eine solche Präzisierung wurde aber natürlich nötig, sobald man nachweisen wollte, daß ein bestimmtes Problem grundsätzlich nicht algorithmisch lösbar ist. Tatsächlich standen die ersten Definitionsversuche (Begriff der *primitiv rekursiven arithmetischen Funktion* durch GÖDEL, 1931) in unmittelbarem Zusammenhang mit dem (versuchten) Nachweis, daß ein bestimmtes Problem der mathematischen Logik nicht algorithmisch entscheidbar ist. Wenig später wurde entdeckt, daß die Klasse der primitiv rekursiven Funktionen nicht alle im anschaulichen Sinne berechenbaren arithmetischen Funktionen umfaßt. Daraufhin wurde eine umfassendere Klasse von berechenbaren arithmetischen Funktionen definiert, die als *partiell rekursive Funktionen* bezeichnet wurden, und der amerikanische Mathematiker A. CHURCH stellte die nach ihm benannte Hypothese[1]) auf, daß jede im anschaulichen Sinne berechenbare arithmetische Funktion partiell rekursiv ist, mit anderen Worten, daß der mathematisch exakt definierte Begriff der partiell rekursiven Funktion (und damit der mit ihm durch (1) verknüpfte Begriff der berechenbaren Wortfunktion) sich genau mit dem verschwommenen anschaulichen Begriff der effektiven Berechenbarkeit deckt und daß auch der durch (3) mit dem Begriff der Berechenbarkeit verknüpfte Begriff der rekursiven Entscheidbarkeit sich genau mit dem verschwommenen anschaulichen Begriff der effektiven Entscheidbarkeit deckt.

Sollte die Hypothese von CHURCH falsch sein (woran allerdings gegenwärtig wohl kein Mathematiker der Welt glaubt), so kann sich dies eines Tages durch die Angabe

[1]) A. CHURCH, An unsolvable problem of elementary number theory, Amer. J. Math. 58 (1936), 345–363.

einer Wortfunktion erweisen, die nicht partiell rekursiv, jedoch in irgendeinem vernünftigen anschaulichen Sinne algorithmisch berechenbar ist. Hingegen kann die Richtigkeit der Hypothese von CHURCH grundsätzlich nicht exakt bewiesen werden, wenn wir unter einem exakten Beweis einen Beweis von der Art verstehen, wie er für mathematische Sätze verlangt wird. Jeder solche Beweis müßte sich ja auf eine Definition des Begriffs der anschaulichen Berechenbarkeit bzw. Entscheidbarkeit stützen, und die Annahme, daß diese Definition genau das Richtige trifft, wäre eine neue Hypothese. Es gibt allerdings gegenwärtig mehrere (schon vom begrifflichen Ansatz her) wesentlich verschiedene Wege zur Definition des Begriffs der berechenbaren Wortfunktion bzw. arithmetischen Funktion, die sämtlich als gleichwertig nachweisbar sind. Es ist klar, daß diese Tatsache sehr für die Wahrheit der Hypothese von CHURCH spricht. Wie bereits erwähnt, werden die wichtigsten Methoden zur Definition des Berechenbarkeitsbegriffs in 7.4. und 7.5. kurz vorgestellt und die Beweise der Gleichwertigkeit zumindest skizziert. In den folgenden grundsätzlichen Überlegungen zu den Begriffen Berechenbarkeit, Entscheidbarkeit und Aufzählbarkeit stützen wir uns nicht auf eine solche exakte Definition, deren technische Einzelheiten viele recht klare Sachverhalte eher verwirren. Vielmehr berufen wir uns darauf, daß der präzisierte Begriff der Berechenbarkeit gewiß allen einfachen Forderungen an einen solchen Begriff genügt, da andernfalls die Hypothese von CHURCH längst verworfen bzw. die Definition des Berechenbarkeitsbegriffs entsprechend korrigiert worden wäre. Unser Vorgehen ist natürlich nicht mathematisch exakt, jedoch können alle unbewiesenen Annahmen über den Berechenbarkeitsbegriff mit Hilfe des in 7.4. behandelten Begriffs der Turing-Berechenbarkeit nachträglich gerechtfertigt werden.

Es sei A ein beliebiges endliches Alphabet. Wir setzen vom präzisierten Berechenbarkeitsbegriff voraus, daß jede berechenbare Wortfunktion f in $W(A)$ durch wenigstens einen Algorithmus realisiert wird und daß es ein höchstens abzählbares Alphabet B gibt, so daß jeder Algorithmus $\mathfrak{A}$ vollständig durch ein gewisses Wort $\mathrm{W}_{\mathfrak{A}} \in W(B)$ beschrieben wird. Durch Umkodierung der Menge $W(B)$ im Alphabet A können wir dann gewissen Wörtern $\mathrm{W} \in W(A)$ einen Algorithmus $\mathfrak{A}$ als Bedeutung zuordnen, der eine berechenbare Wortfunktion in $\mathrm{W}(A)$ realisiert. Ist W bezüglich dieser Kodierung ein Kodewort für einen Algorithmus $\mathfrak{A}$, so schreiben wir W als $\langle\mathfrak{A}\rangle$ und nennen das Wort $\langle\mathfrak{A}\rangle$ eine *Niederschrift von* $\mathfrak{A}$. Wir setzen voraus, daß die Menge aller so definierten Niederschriften eine (im präzisierten Sinne) entscheidbare Teilmenge von $W(A)$ ist und daß man bei Kenntnis einer Niederschrift $\langle\mathfrak{A}\rangle$ eines Algorithmus $\mathfrak{A}$ die Folge der Zwischenresultate $\mathfrak{A}^0(\mathrm{W}) = \mathrm{W}, \mathfrak{A}^1(\mathrm{W}), \ldots, \mathfrak{A}^{n+1}(\mathrm{W}) = \mathfrak{A}(\mathfrak{A}^n(\mathrm{W})), \ldots$, die bei der taktweisen Anwendung von $\mathfrak{A}$ auf ein Wort $\mathrm{W} \in W(A)$ entstehen, soweit sie existieren, effektiv berechnen und ferner entscheiden kann, ob ein gewisses Zwischenresultat $\mathfrak{A}^n(\mathrm{W})$ das Endresultat ist, welches dann mit $\mathfrak{A}(\mathrm{W})$ bezeichnet wird.

Am Rande sei bemerkt, daß bei Benutzung üblicher Algorithmenbegriffe und Niederschriftsverfahren die Zuordnung zwischen Algorithmen und ihren Niederschriften sogar eineindeutig ist. Faßt man jedoch die Niederschrift $\langle\mathfrak{A}\rangle$ eines Algorithmus nicht als Kodewort des Algorithmus $\mathfrak{A}$ (d. h. des Berechnungs*verfahrens*), sondern als Kodewort für die durch $\mathfrak{A}$ realisierte Wortfunktion auf, so ist die so definierte Kodierung der Menge der berechenbaren Wortfunktionen in $W(A)$ niemals eineindeutig, selbst wenn die Kodierung der Algorithmen durch ihre Niederschriften eineindeutig ist, da eine überhaupt berechenbare Wortfunktion immer durch unendlich viele verschiedene Algorithmen realisiert wird. Die Menge der berechenbaren Wortfunktionen in $W(A)$ bildet daher ein markantes Beispiel für eine Menge abstrakter Objekte, die sich zwar kodieren lassen, deren ursprüngliche Kodierung aber nicht eineindeutig ist und, wie sich zeigt, auch nicht in konstruktiver Weise zu einer eineindeutigen Kodierung umgestaltet werden kann.

Setzen wir voraus, daß alle konstanten Wortfunktionen berechenbar sind, so gibt es mindestens abzählbar viele berechenbare Wortfunktionen in $W(A)$. Andererseits ist die Menge aller berechenbaren Wortfunktionen in $W(A)$ als kodierbare Menge höchstens abzählbar (vgl. 7.1.), mithin abzählbar. Da die Menge aller Abbildungen aus einer abzählbaren Menge in sich bekanntlich überabzählbar ist, können wir schließen, daß es überabzählbar viele nicht berechenbare Wortfunktionen in $W(A)$ gibt.

Die Menge aller Teilmengen einer abzählbaren Menge ist bekanntlich überabzählbar. Da jede entscheidbare Teilmenge von $W(A)$ eine berechenbare charakteristische Funktion besitzt, kann es höchstens abzählbar viele entscheidbare Teilmengen von $W(A)$ geben. Andererseits sind sicher alle Einermengen von Wörtern aus $W(A)$ entscheidbar. Daher gibt es genau abzählbar viele entscheidbare Teilmengen von $W(A)$ und folglich überabzählbar viele nicht entscheidbare Teilmengen von $W(A)$. Im folgenden geben wir drei spezielle Teilmengen von $W(A)$ an und beweisen, daß diese nicht entscheidbar sind. Jeder dieser nicht entscheidbaren Wortmengen entspricht ein unentscheidbares Problem.

Problem der Selbstanwendbarkeit. Es sei M_1 die Menge aller Niederschriften von Algorithmen $\mathfrak{A}$, für die $\mathfrak{A}(\langle\mathfrak{A}\rangle)$ existiert. Die Annahme, M_1 sei entscheidbar, bedeutet, daß es einen Algorithmus $\mathfrak{B}$ gibt, der auf alle Wörter $\mathrm{W} \in W(A)$ anwendbar ist und für den

$$\mathfrak{B}(\mathrm{W}) = \begin{cases} \Lambda, \text{ falls } \mathrm{W} = \langle\mathfrak{A}\rangle \text{ für einen Algorithmus } \mathfrak{A}, \text{ für den } \mathfrak{A}(\langle\mathfrak{A}\rangle) \text{ existiert,} \\ \mathrm{a} \text{ andernfalls (d. h. falls W nicht Niederschrift eines Algorithmus} \\ \quad \text{oder Niederschrift eines solchen Algorithmus ist, für den } \mathfrak{A}(\langle\mathfrak{A}\rangle) \\ \quad \text{nicht existiert)} \end{cases}$$

gilt. Aus einem solchen Algorithmus $\mathfrak{B}$ kann man sicher einen Algorithmus $\mathfrak{A}_0$ konstruieren, für den gilt

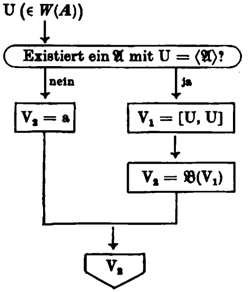
$$\mathfrak{A}_0(\mathrm{W}) = \begin{cases} \mathrm{a}, \text{ falls } \mathfrak{B}(\mathrm{W}) = \mathrm{a}, \\ \text{nicht definiert andernfalls.} \end{cases}$$

Daher existiert $\mathfrak{A}_0(\mathrm{W})$ genau dann, wenn W nicht Niederschrift eines Algorithmus oder Niederschrift eines solchen Algorithmus ist, für den $\mathfrak{A}(\langle\mathfrak{A}\rangle)$ nicht existiert. Insbesondere ergibt sich aus unserer Annahme der Widerspruch, daß $\mathfrak{A}_0(\langle\mathfrak{A}_0\rangle)$ genau dann existiert, wenn $\mathfrak{A}_0(\langle\mathfrak{A}_0\rangle)$ nicht existiert. Demnach ist die Menge $\boldsymbol{M}_1$ und damit das Problem der Selbstanwendbarkeit eines Algorithmus nicht entscheidbar.

Uniformes Halteproblem. Durch geeignete effektive Kodierung sei jedem Paar U, V von Wörtern aus $W(\boldsymbol{A})$ eineindeutig ein Wort $[\mathrm{U}, \mathrm{V}] \in W(\boldsymbol{A})$ zugeordnet. (Falls $\boldsymbol{A}$ wenigstens zwei Buchstaben a, b enthält, kann man z. B. das Wort $\underbrace{\mathrm{aa} \ldots \mathrm{a}}_{g_{\mathrm{lex}}(\mathrm{U})}\mathrm{b}\underbrace{\mathrm{a} \ldots \mathrm{aa}}_{g_{\mathrm{lex}}(\mathrm{V})}$ als [U, V] nehmen. Enthält $\boldsymbol{A}$ nur einen Buchstaben a, so nehme man etwa die Gödelzahl des obigen Wortes.) Es sei

$$\boldsymbol{M}_2 := \{[\langle\mathfrak{A}\rangle, \mathrm{W}]\colon \ \mathfrak{A}(\mathrm{W}) \ \textit{existiert}\}.$$

Aus der Annahme, daß $\boldsymbol{M}_2$ entscheidbar, d. h. die charakteristische Funktion von $\boldsymbol{M}_2$ durch einen Algorithmus $\mathfrak{B}$ berechenbar ist, konstruieren wir einen Algorithmus zur Berechnung der charakteristischen Funktion von $\boldsymbol{M}_1$:



(nach Voraussetzung über die Niederschriften entscheidbar)

(nach Voraussetzung berechenbar)

(nach Voraussetzung ist

$$\mathrm{V}_2 = \begin{cases} \Lambda, & \text{falls } \mathfrak{A}(\langle\mathfrak{A}\rangle) \text{ existiert,} \\ \mathrm{a} & \text{andernfalls.)} \end{cases}$$

Da $\boldsymbol{M}_1$ bereits als nicht entscheidbar erkannt ist, folgt die Unentscheidbarkeit von $\boldsymbol{M}_2$.

Der Beweis der Unentscheidbarkeit von $\boldsymbol{M}_2$ enthüllt beispielhaft einen typischen Zug: Die Tatsache, daß $\boldsymbol{M}_1$ nicht entscheidbar ist, ist „relativ uninteressant", $\boldsymbol{M}_1$ ist offenbar überhaupt nur zu dem Zweck definiert worden, auf möglichst bequeme Weise von einer konkreten Menge die Unentscheidbarkeit nachweisen zu können. Der Beweis der inhaltlich wesentlich interessanteren Unentscheidbarkeit von $\boldsymbol{M}_2$ (sie bedeutet immerhin, daß es kein effektives Verfahren gibt, durch das man im voraus entscheiden kann, ob die Anwendung eines Algorithmus auf ein Wort definiert ist) wird auf die bereits bewiesene Unentscheidbarkeit einer anderen Menge, nämlich

M_1, zurückgeführt, indem man feststellt, daß das unentscheidbare Problem sozusagen ein Spezialfall des neuen Problems ist, also entscheidbar sein müßte, wenn das allgemeinere neue Problem entscheidbar wäre. Nun ist auch das Problem der Entscheidbarkeit von M_2 noch ein „inneres Problem" der Algorithmentheorie, die insgesamt zu dem Zweck entwickelt wurde, die Frage der Entscheidbarkeit gewisser allgemeiner Probleme (z. B. der mathematischen Logik) zu klären. Es wird sich zeigen, daß die Unentscheidbarkeit von logischen und mathematischen Problemen allgemein auf die Unentscheidbarkeit gewisser innerer Probleme der Algorithmentheorie bzw. auf die Existenz solcher Probleme in grundsätzlich ähnlicher Weise zurückgeführt wird wie die Unentscheidbarkeit von M_2 auf die Unentscheidbarkeit von M_1.

Die Unentscheidbarkeit des uniformen Halteproblems schließt nicht aus, daß es für spezielle Algorithmen (möglicherweise sogar für jeden Algorithmus $\mathfrak{A}$?) einen Entscheidungsalgorithmus $\overline{\mathfrak{A}}$ für den Definitionsbereich $D_{\mathfrak{A}}$ von $\mathfrak{A}$ gibt bzw. daß die charakteristische Funktion des Definitionsbereichs für einige berechenbare Funktionen berechenbar ist. Das entsprechende Problem bezeichnet man als *Halteproblem für* $\mathfrak{A}$. Das Halteproblem ist z. B. trivialerweise für jede berechenbare volle Funktion f in $W(A)$ (d. h. $D_f = W(A)$) entscheidbar, da die zugehörige charakteristische Funktion als konstante Funktion sicher berechenbar ist. Wir zeigen, daß es berechenbare Funktionen mit nicht entscheidbarem Definitionsbereich gibt, mit anderen Worten, daß es Algorithmen mit unentscheidbarem Halteproblem gibt:

Aus unseren Voraussetzungen über die taktweise Berechenbarkeit der Zwischenresultate $\mathfrak{A}^*(\mathrm{W})$ bei Vorliegen der Niederschrift $\langle\mathfrak{A}\rangle$ eines Algorithmus und eines Eingabewortes W ergibt sich, daß es einen *universellen Algorithmus* $\mathfrak{A}$ geben muß, der auf beliebige Eingaben der Form $[\langle\mathfrak{A}\rangle, \mathrm{W}]$ genau dann anwendbar ist, wenn $\mathfrak{A}(\mathrm{W})$ existiert, und in diesem Fall als Resultat ebenfalls $\mathfrak{A}(\mathrm{W})$ liefert. Tatsächlich gestattet jede der bisher betrachteten Präzisierungen des Algorithmenbegriffs, einen solchen universellen Algorithmus explizit zu konstruieren, jedoch setzt diese Konstruktion im allgemeinen eine ausgefeilte „Unterprogrammtechnik" in bezug auf den jeweiligen präzisierten Algorithmenbegriff voraus. Das Halteproblem eines universellen Algorithmus kann nicht entscheidbar sein, denn es ist im wesentlichen mit dem uniformen Halteproblem identisch.

Einen gewissen Ersatz für die nicht erfüllte Entscheidbarkeit vieler wichtiger Wortmengen stellt die im folgenden behandelte Aufzählbarkeit dar.

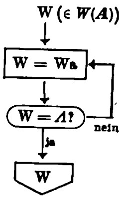
Definition 1. Es sei A ein beliebiges Alphabet, $\mathrm{a} \in A$ ein beliebiger fester Buchstabe. Eine Menge $M \subseteq W(A)$ heißt *aufzählbar*, wenn $M = \emptyset$ oder wenn eine berechenbare Wortfunktion f in $W(A)$ mit dem Definitionsbereich $W(\{\mathrm{a}\})$ und dem Wertebereich M existiert.

Den Zusammenhang mit der anschaulichen Bedeutung des Wortes *aufzählen* sieht man am besten, wenn man für eine durch eine berechenbare Wortfunktion f aufgezählte nichtleere Menge M das Bildwort $f(\underbrace{\mathrm{a} \ldots \mathrm{a}}_{n})$ in der Form W_n schreibt.

Dann ist $M = \{\mathrm{W}_n : n = 0, 1, 2, \ldots\}$, wobei für jede natürliche Zahl n das Wort W_n aus dem Index n effektiv gewonnen werden kann. Man beachte, daß die aufzählende Abbildung nicht eineindeutig sein muß. Insbesondere stellt sich leicht heraus, daß jede endliche Menge aufzählbar ist. Weiterhin ist die Menge $W(A)$ stets (z. B. durch die berechenbare Wortfunktion $f(\underbrace{\mathrm{a} \ldots \mathrm{a}}_{n}) := g_{\mathrm{lex}}^{-1}(n)$) aufzählbar. Da jede aufzählbare Teilmenge von $W(A)$ durch wenigstens eine berechenbare Wortfunktion aufgezählt wird, folgt aus der Abzählbarkeit der Menge der berechenbaren Wortfunktionen in $W(A)$ sofort, daß es höchstens (und mithin genau) abzählbar viele aufzählbare Teilmengen von $W(A)$ gibt.

Satz 1. *Eine Menge $M \subseteq W(A)$ ist aufzählbar genau dann, wenn M Wertebereich einer berechenbaren Wortfunktion in $W(A)$ ist.*

Beweis. In der einen Richtung ist nur noch zu zeigen, daß auch $\emptyset$ als Wertebereich einer berechenbaren Wortfunktion auftritt, mit anderen Worten, daß die nirgends definierte Wortfunktion berechenbar ist. Dies kann man natürlich nur nach Kenntnis einer konkreten Präzisierung des Algorithmenbegriffs nachweisen. Der bei flüchtiger Betrachtung vielleicht wenig anschauliche Sachverhalt wird verständlicher, wenn man berücksichtigt, daß die Menge der berechenbaren Funktionen natürlich bezüglich Verknüpfung mittels Flußdiagrammen abgeschlossen sein soll.[1]) Es ist aber sehr leicht, aus gewiß berechenbaren Funktionen und gewiß entscheidbaren Relationen ein nirgends definiertes Flußdiagramm zu organisieren, z. B.



Es sei nun M der Wertebereich einer berechenbaren Wortfunktion f mit nichtleerem Definitionsbereich. (Falls $D_f = \emptyset$, ist auch $W_f = \emptyset$, folglich nach Definition 1 aufzählbar.) Wir haben zu zeigen, daß M auch als Wertebereich einer auf $W(\{\mathrm{a}\})$ (bzw. zur Vereinfachung der folgenden Formulierungen: auf $\mathbf{N}$) definierten Funktion h

[1]) Das heißt, wir stellen folgende Grundforderung an einen präzisierten Berechenbarkeitsbegriff: Belegt man alle Arbeitsknoten eines beliebigen Strukturgraphen mit im präzisierten Sinn berechenbaren Funktionen und alle seine Prüfknoten mit im präzisierten Sinn entscheidbaren Relationen, so ist die durch das so entstehende Flußdiagramm definierte partielle Funktion im präzisierten Sinn berechenbar (vgl. auch (a) auf S. 152).

erhalten werden kann, die ebenfalls berechenbar ist. Es sei $\mathfrak{A}$ ein Algorithmus, der f berechnet. Ist $\mathfrak{A}^n(\mathrm{W})$ für $\mathrm{W} \in D_f$ und $n \in \mathbf{N}$ das Endresultat $f(\mathrm{W})$, so setzen wir $\mathfrak{A}^m(\mathrm{W}) = \mathfrak{A}^n(\mathrm{W}) \quad$ für $m > n$.

Für $\mathrm{W} \notin D_f$ bricht die Folge der Zwischenresultate $\mathfrak{A}^m(\mathrm{W})$ ohnehin nicht ab. Es sei ferner $\mathrm{W}_0, \mathrm{W}_1, \mathrm{W}_2, \ldots$ eine berechenbare Aufzählung von $W(A)$. Dann ist jedes Element des zweifach unendlichen quadratischen Schemas

$$\begin{array}{cccccc}
\mathfrak{A}^0(\mathrm{W}_0), & \mathfrak{A}^1(\mathrm{W}_0), & \mathfrak{A}^2(\mathrm{W}_0), & \mathfrak{A}^3(\mathrm{W}_0), & \mathfrak{A}^4(\mathrm{W}_0), & \ldots \\
\mathfrak{A}^0(\mathrm{W}_1), & \mathfrak{A}^1(\mathrm{W}_1), & \mathfrak{A}^2(\mathrm{W}_1), & \mathfrak{A}^3(\mathrm{W}_1), & \mathfrak{A}^4(\mathrm{W}_1), & \ldots \\
\mathfrak{A}^0(\mathrm{W}_2), & \mathfrak{A}^1(\mathrm{W}_2), & \mathfrak{A}^2(\mathrm{W}_2), & \mathfrak{A}^3(\mathrm{W}_2), & \mathfrak{A}^4(\mathrm{W}_2), & \ldots \\
\mathfrak{A}^0(\mathrm{W}_3), & \mathfrak{A}^1(\mathrm{W}_3), & \mathfrak{A}^2(\mathrm{W}_3), & \mathfrak{A}^3(\mathrm{W}_3), & \mathfrak{A}^4(\mathrm{W}_3), & \ldots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots
\end{array}$$

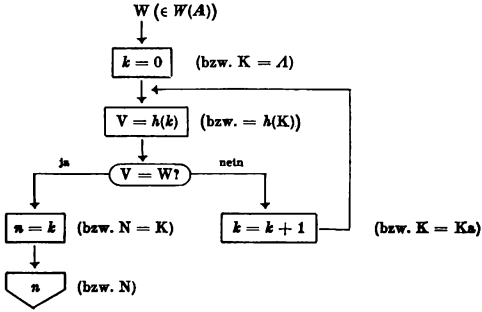
aus vorgegebener Zeilennummer n und Spaltennummer m berechenbar. Berechnet man die Wörter $\mathfrak{A}^m(\mathrm{W}_n)$ in der durch die Pfeile angedeuteten Reihenfolge, so stößt man, da $W_f \neq \emptyset$ ist, irgendwann ein erstes Mal auf ein Endresultat, d. h. ein Element von $W_f = M$, und da alle in einer Zeile auf ein Endresultat folgenden Wörter ebenfalls Endresultate sind, stößt man nach jedem Endresultat (spätestens beim nächsten Durchgang durch die entsprechende Zeile) wieder auf ein Endresultat. Es sei nun $h(0)$ das erste so erhaltene Endresultat und $h(n + 1)$ das erste in der angegebenen Weise nach $h(n)$ erhaltene Endresultat. Es ist klar, daß die so definierte Abbildung h eine berechenbare eindeutige (aber nicht eineindeutige!) Wortfunktion mit dem Definitionsbereich $\mathbf{N}$ (bzw. $W(\{a\})$) und dem Wertebereich M ist.

Satz 2. *Eine Menge $M \subseteq W(A)$ ist aufzählbar genau dann, wenn M Definitionsbereich einer berechenbaren Wortfunktion f in $W(A)$ ist.*

Beweis. Für $M = \emptyset$ ist der Satz auf Grund vorausgegangener Bemerkungen klar. Es sei nun f eine berechenbare Wortfunktion in $W(A)$ mit dem Definitionsbereich M. Dann ist die durch

$$h(\mathrm{W}) := \begin{cases} \mathrm{W}, & \textit{falls } f(\mathrm{W}) \textit{ existiert,} \\ \textit{nicht definiert andernfalls} & \end{cases}$$

definierte Funktion h gewiß auch berechenbar, und es ist $W_h = D_f = M$; daher ist M nach Satz 1 aufzählbar. Umgekehrt sei h eine berechenbare Funktion mit dem Definitionsbereich $\mathbf{N}(W(\{a\}))$ und dem Wertebereich M. Dann ist die Wortfunktion f, die jedem Wort $W \in M$ die kleinste natürliche Zahl n mit $h(n) = W$ zuordnet, gewiß auch berechenbar, wie das folgende Flußdiagramm bestätigt:



Offenbar ist $D_f = M$.

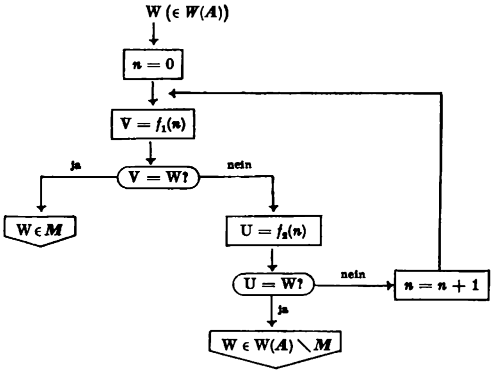
Folgerung. *Es gibt aufzählbare Mengen, die nicht entscheidbar sind.* Der Definitionsbereich einer beliebigen Wortfunktion mit unentscheidbarem Halteproblem (z. B. einer universellen Wortfunktion) hat diese Eigenschaft.

Satz 3. *Eine Menge $M \subseteqq W(A)$ ist entscheidbar genau dann, wenn M und $W(A) \setminus M$ aufzählbar sind. Insbesondere ist jede entscheidbare Menge aufzählbar.*

Wir beweisen zunächst den Zusatz. Ist M entscheidbar, so ist die durch

$$f(W) := \begin{cases} W, & \textit{falls } W \in M, \\ \textit{nicht definiert andernfalls} \end{cases}$$

definierte Wortfunktion f offenbar berechenbar, und es ist $D_f = W_f = M$, folglich ist M aufzählbar. Da mit einer Menge M auch ihre Komplementärmenge $W(A) \setminus M$ entscheidbar ist (durch das gleiche Verfahren, wobei lediglich „ja"- und „nein"-Ausgang zu vertauschen sind), sind folglich M und $W(A) \setminus M$ aufzählbar, falls M entscheidbar ist. Es seien nun M und $W(A) \setminus M$ als aufzählbar vorausgesetzt, wobei wir annehmen können, daß f_1 eine berechenbare Abbildung von $\mathbf{N}$ auf M und f_2 eine berechenbare Abbildung von $\mathbf{N}$ auf $W(A) \setminus M$ ist. Unter dieser Voraussetzung ist ein Entscheidungsverfahren für das Problem $W \in M$? gegeben durch



Folgerung. *Wir können einige der überabzählbar vielen nicht aufzählbaren Teilmengen von $W(A)$ explizit angeben.* Nach Satz 3 ist das Komplement einer aufzählbaren, aber nicht entscheidbaren Menge nicht aufzählbar. Vgl. die vorige Folgerung.

Wie bereits betrachtet, wird die Berechenbarkeit n-stelliger Wortfunktionen auf die Berechenbarkeit spezieller einstelliger Wortfunktionen in einem durch einen Hilfsbuchstaben $*$ vergrößerten Alphabet zurückgeführt. Dem entspricht die Definition eines modifizierten kartesischen Produkts

$$\boldsymbol{M}_1 * \boldsymbol{M}_2 := \{\mathrm{W}_1 * \mathrm{W}_2 \colon \mathrm{W}_1 \in \boldsymbol{M}_1 \wedge \mathrm{W}_2 \in \boldsymbol{M}_2\}, \ \text{\textit{falls}} \ \boldsymbol{M}_1, \boldsymbol{M}_2 \subseteqq \mathrm{W}(\boldsymbol{A}), * \notin \boldsymbol{A}.$$

Das so definierte Produkt von Wortmengen ist, im Gegensatz zum mengentheoretischen kartesischen Produkt, sogar assoziativ.

Satz 4. *Sind $\boldsymbol{M}_1, \boldsymbol{M}_2 \subseteqq W(\boldsymbol{A})$ aufzählbare Mengen, so sind auch $\boldsymbol{M}_1 \cup \boldsymbol{M}_2$, $\boldsymbol{M}_1 \cap \boldsymbol{M}_2$ und $\boldsymbol{M}_1 * \boldsymbol{M}_2$ aufzählbar.*

Beweis. Falls eine der Mengen $\boldsymbol{M}_1, \boldsymbol{M}_2$ (etwa $\boldsymbol{M}_2$) leer ist, so ist $\boldsymbol{M}_1 \cup \boldsymbol{M}_2 = \boldsymbol{M}_1$, $\boldsymbol{M}_1 \cap \boldsymbol{M}_2 = \boldsymbol{M}_1 * \boldsymbol{M}_2 = \emptyset$; folglich sind alle drei Mengen aufzählbar. Es seien nun $\boldsymbol{M}_1$ und $\boldsymbol{M}_2$ nichtleer, f_1 sei eine berechenbare Abbildung von $\mathbf{N}$ auf $\boldsymbol{M}_1$ und f_2 eine berechenbare Abbildung von $\mathbf{N}$ auf $\boldsymbol{M}_2$. Da gewiß entscheidbar ist, ob eine natürliche Zahl n gerade oder ungerade ist, und ebenso gewiß diejenige Zahl k mit $n = 2k$ bzw. $n = 2k + 1$ aus n berechenbar ist, ist die durch

$$f_3(n) := \begin{cases} f_1(k), \ \text{\textit{falls}} \ n = 2k, \\ f_2(k), \ \text{\textit{falls}} \ n = 2k + 1, \end{cases}$$

definierte Funktion f_3 berechenbar und zählt $M_1 \cup M_2$ auf. Wir nehmen weiter an, daß die durch

$$h(2^n(2m+1)) := (n, m)$$

definierte eineindeutige Abbildung von $\mathbf{N} \setminus \{0\}$ auf $\mathbf{N}^2$ berechenbar ist. Dann ist

$$\mathrm{f}_4(2^n(2m+1)) := \mathrm{f}_1(n) * \mathrm{f}_2(m)$$

eine berechenbare Abbildung von $\mathbf{N} \setminus \{0\}$ auf $M_1 * M_2$. Daraus folgt wiederum, daß die durch

$$f_5(n) := \begin{cases} \mathrm{W}, \textit{ falls } \mathrm{f}_4(n) = \mathrm{W} * \mathrm{W}, \\ \textit{nicht definiert andernfalls} \end{cases}$$

definierte Abbildung aus $\mathbf{N}$ auf $M_1 \cap M_2$ berechenbar ist.

Vor der Formulierung des folgenden Satzes erinnern wir daran, daß eine n-stellige Operation f in $W(A)$ als berechenbar gilt, wenn für einen beliebigen Hilfsbuchstaben $* \notin A$ die durch

$$f'(\mathrm{W}_1 * \mathrm{W}_2 * \cdots * \mathrm{W}_n) := f(\mathrm{W}_1, \mathrm{W}_2, \ldots, \mathrm{W}_n)$$

definierte Wortfunktion f' in $W(A \cup \{*\})$ berechenbar ist.

Satz 5. *Ist $M \subseteq W(A)$ aufzählbar und F eine aufzählbare Menge von (eventuell mehrstelligen und zum Teil partiellen) berechenbaren Operationen in $W(A)$ (wobei man sich jedes Element f von F etwa durch die in $W(A)$ formulierte Niederschrift eines Algorithmus zur Berechnung von f gegeben denken kann,) so ist die kleinste bezüglich aller Operationen aus F abgeschlossene Obermenge $\overline{M}$ von M ebenfalls aufzählbar.*

Beweis. Der Satz gilt trivialerweise, falls $M = \emptyset$ oder $F = \emptyset$ ist, da dann $\overline{M} = \emptyset$ bzw. $\overline{M} = M$ ist. Es sei daher $f_0, f_1, f_2, \ldots$ eine berechenbare Aufzählung von F, d. h., man kann bei gegebenem m die Niederschrift eines Algorithmus zur Berechnung von f_m und bei weiter gegebenen Wörtern $\mathrm{W}_1, \ldots, \mathrm{W}_n$ das Wort $f_m(\mathrm{W}_1, \ldots, \mathrm{W}_n)$ berechnen, falls es existiert. Weiterhin sei $\mathrm{V}_0, \mathrm{V}_1, \mathrm{V}_2, \ldots$ eine berechenbare Aufzählung von M. Die Menge $\overline{M}$ ist offenbar gleich der Menge aller Werte, die die Terme einer passend zu wählenden Termsprache annehmen, wenn man ihre Operationssymbole durch die Operationen f_m interpretiert und die Variablen mit Wörtern aus M belegt. Um eine solche Termsprache als Teilmenge eines endlichen Alphabets erhalten zu können, bezeichnen wir die abzählbar vielen Variablen mit x, xo, xoo, xooo usw. Als Operationssymbol für die Funktion f_m wählen wir, falls diese Funktion n-stellig ist, die Zeichenkombination $\underbrace{\mathrm{o} \ldots \mathrm{o}}_{n}\mathrm{F}\underbrace{\mathrm{o} \ldots \mathrm{o}}_{m}$. Damit ist die Menge T der wie üblich hieraus aufgebauten einsortigen Terme eine Teilmenge von $W(\{\mathrm{F}, \mathrm{x}, \mathrm{o}, (,), *\})$ ($*$ dient beim Aufbau der Terme als Komma). Interpretiert man nun, wie beabsichtigt, das Symbol $\underbrace{\mathrm{o}\ldots\mathrm{o}}_{n}\mathrm{F}\underbrace{\mathrm{o}\ldots\mathrm{o}}_{m}$ durch die Operation f_m (diese Interpretation sei mit ω_0 bezeichnet) und belegt man die Variable $\mathrm{x}\underbrace{\mathrm{o}\ldots\mathrm{o}}_{n}$ mit dem Wort V_n (diese Belegung sei mit b bezeichnet), so ist offenbar das Wort $Wert(\mathrm{t}, \omega_0, b)$, falls es in $W(\boldsymbol{A})$ existiert, für jeden Term $\mathrm{t} \in \boldsymbol{T}$ effektiv berechenbar, mit anderen Worten: Es existiert eine berechenbare Abbildung von der Wortmenge $\boldsymbol{T}$ auf die Menge

$$\overline{\boldsymbol{M}} = \{Wert(\mathrm{t}, \omega_0, b) : \mathrm{t} \in \boldsymbol{T}\}.$$

Das bedeutet aber, daß $\overline{\boldsymbol{M}}$ aufzählbar ist.

7.4. Turingmaschinen

Die bis heute wohl verbreitetste und beliebteste Präzisierung des Berechenbarkeitsbegriffs beruht auf dem Begriff der Turingmaschine und wurde 1937 von A. M. TURING[1]) und unabhängig davon in ähnlicher Form bereits 1936 von E. POST[2]) vorgeschlagen. Das Konzept der Turingmaschine erhält man fast zwangsläufig, wenn man einerseits von einer Zerlegung der Wortalgorithmen in kleinstmögliche Einzelschritte und andererseits von der Vorstellung ausgeht, daß ein wirklich streng algorithmischer Wortumformungsprozeß prinzipiell von einer geeignet konstruierten Maschine ausführbar sein muß. Turingmaschinen, die heute in vielen verschiedenen (aber prinzipiell gleichwertigen) Varianten betrachtet werden, existieren nirgendwo in der Welt real. Man kann ihre Bauart und Arbeitsweise mehr oder weniger formal beschreiben und dann ihre Arbeit taktweise in Gedanken verfolgen bzw. mit Papier und Bleistift nachrechnen — in ähnlicher Weise, wie auch in der Physik gewisse Experimente nur in Gedanken durchgeführt werden. Für eine praktische Anwendung in der Rechentechnik sind Turingmaschinen, obwohl ihre technische Realisierung heute leicht möglich wäre, nicht geeignet, da sie jeden noch so einfachen Prozeß in sehr viele Einzelschritte auflösen und daher „sehr langsam" arbeiten. Die Beschäftigung mit Turingmaschinen vermittelt jedoch sowohl grundlegende Erkenntnisse als auch praktische Fertigkeiten, die in der Kybernetik und Rechentechnik sehr nützlich sind. In zunehmendem Maße werden die von der Theorie der Turingmaschinen bereitgestellten Formalismen, Simulationsmethoden u. ä. auch als Zwischenstufe bei der mathematischen Modellierung komplizierter Prozesse in Medizin, Naturwissenschaft, Technik und Gesellschaftswissenschaften genutzt. So kann man feststellen, daß die Turingmaschinen trotz ihrer schattenhaften Existenz im Laufe

[1]) A. M. TURING, On computable numbers, with an application to the Entscheidungsproblem, Proc. London Math. Soc. 42 (1937), 230—265.

[2]) E. POST, Finite combinatory process-formulation I, J. of Symb. Logic 1 (1936), 103—105.

der letzten 30 Jahre einen erheblichen Einfluß auf die Theorie und Praxis des programmgesteuerten Rechnens ausgeübt haben.

Wir geben zunächst eine recht allgemeine Beschreibung der Turingmaschine an, die viele mögliche Varianten umfaßt. Eine Turingmaschine besteht aus einem oder mehreren potentiell unendlichen, in diskrete Zellen eingeteilten Speichern und einem mit diesen Speichern in Wechselwirkung stehenden endlichen Automaten, d. h. einem endlichen diskreten dynamischen System, das durch endlich viele Zustände, endlich viele Ein- und Ausgabesignale und eine endliche Tabelle charakterisiert wird, die die taktweisen Ausgaben und Zustandsänderungen des Systems in Abhängigkeit vom jeweils herrschenden Zustand und den empfangenen Eingabesignalen festlegt. Um die Wechselwirkung eines solchen endlichen Automaten mit dem oder den äußeren Speichern zu ermöglichen, müssen die Zellen der Speicher in der Lage sein, jeweils ein Ausgabesignal bzw. eine (über einen bestimmten Ausgabekanal ausgegebene) Komponente davon zu speichern. Umgekehrt muß der Automat in der Lage sein, den Inhalt gewisser momentan eingestellter Speicherzellen zu lesen (als Eingabesignale zu empfangen) und über gewisse Ausgabekanäle (Transporteinrichtungen) sich relativ zum Speicher zu bewegen. Im folgenden beschreiben wir zunächst anschaulich und dann formalisiert eine möglichst einfache, zur Berechnung von Wortfunktionen geeignete Variante des allgemeinen Konzepts.

Wir nehmen an, daß die Speicherzellen ein links beschränktes, rechts unbeschränktes Band bilden (Abb. 1). Die am weitesten links befindliche Speicherzelle trägt ein Symbol s, das gelesen, aber nicht gelöscht und auch nicht in andere Zellen geschrieben werden kann. In jeder der übrigen Zellen steht zu jedem Zeitpunkt genau ein von s verschiedener Buchstabe des endlichen Bandalphabets $\boldsymbol{B}$, wobei $\boldsymbol{B}$ die Vereinigung des nichtleeren Ein- und Ausgabealphabets $\boldsymbol{A}$ und eines Hilfsalphabets $\boldsymbol{H}$ ist, welches außer s mindestens noch ein sogenanntes Leerzeichen o, außerdem eventuell weitere Buchstaben enthält. Wir nehmen an, daß die eigentliche Maschine eine endliche Menge Q von Zuständen annehmen kann, unter denen sich ein ausgezeichneter Anfangszustand q_1 und eine nichtleere Teilmenge Q_0 von q_1 verschiedener Stopzustände befinden. Die Maschine kann mittels eines sogenannten Lese- und Schreibkopfes (kurz LS-Kopf) jeweils den Inhalt einer Zelle des Bandes lesen. Ihre Arbeitsweise wird durch eine endliche Tabelle beschrieben, in der, in Abhängigkeit vom gerade herrschenden Zustand und vom gerade gelesenen Buchstaben, festgelegt ist: a) der in die gerade gelesene Zelle hineinzuschreibende Buchstabe (wobei der alte Inhalt der Zelle gelöscht wird), b) der im nächsten Takt herrschende Zustand, c) die im nächsten Takt zu bearbeitende Zelle, welche die links benachbarte (Steuerbefehl L), rechts benachbarte (Steuerbefehl R) oder wieder die gleiche Zelle (Steuerbefehl N) sein kann. Wir nehmen schließlich an, daß zu Beginn der Berechnung der LS-Kopf auf die mit s markierte Anfangszelle und der Automat auf den Zustand q_1 eingestellt ist (Abb. 1a), daß sich in endlich vielen, rechts an s anschließenden Zellen ein Wort W $\in W(\boldsymbol{A})$ befindet und daß in den potentiell unendlich vielen weiteren Zellen das Leerzeichen o

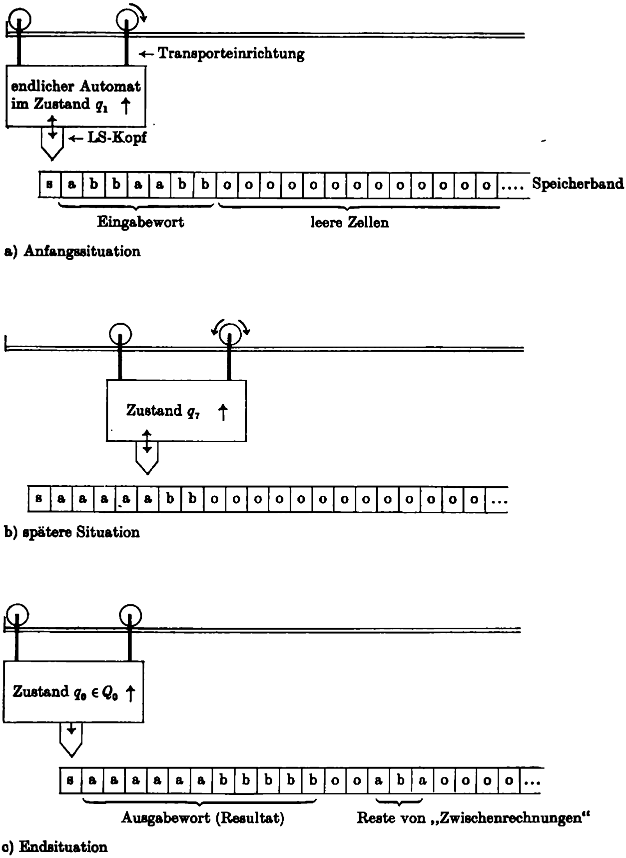


Abb. 1. Situationen einer Turingmaschine

steht. Unter diesen Voraussetzungen sind offenbar die sich Takt für Takt einstellenden Folgesituationen (bestehend aus dem jeweiligen Speicherinhalt, Zustand und Stellung des LS-Kopfes) im anschaulichen Sinne berechenbar. Damit jede so definierte Turingmaschine eine ganz bestimmte (eventuell partielle oder gar nirgends definierte) Wortfunktion realisiert, setzen wir noch fest: Die Programmtafel der Maschine sei so beschaffen, daß ein Zustand aus Q_0 höchstens beim Lesen des Buchstaben s angenommen werden kann und daß in diesem Fall der Steuerbefehl N gegeben wird, so daß der LS-Kopf im Zustand $q_0 \in Q_0$ über der Anfangszelle stehenbleibt. (Es wird nicht verlangt, daß die Maschine bei jeder Rückkehr auf die Anfangszelle stoppt.) Für die Stopzustände $q_0 \in Q_0$ braucht in der Tafel keine Festsetzung von Folgezustand usw. angegeben werden. Erreicht die Maschine, ausgehend von der beschriebenen Anfangssituation, nach endlich vielen Takten eine Situation mit einem Stopzustand (Abb. 1c), so gilt das längste zusammenhängende Wort aus $W(\boldsymbol{A})$, das sich zu diesem Zeitpunkt rechts an die Anfangszelle anschließt, als das dem Eingabewort zugeordnete Resultatwort. Wird eine solche Stopsituation nicht erreicht, so ist die betreffende Wortfunktion für das Eingabewort W nicht definiert.

Wir verwandeln nun die wesentlichen Angaben der anschaulichen Beschreibung unserer Turingmaschinen-Variante in eine abstrakte mathematische Definition.

Definition 1. Eine *Turingmaschine* ist ein System der Form $(\boldsymbol{A}, \boldsymbol{H}, Q, q_1, Q_0, \mathrm{s}, \mathrm{o}, \tau)$, wobei

$\boldsymbol{A}$ eine nichtleere endliche Menge (*Ein- und Ausgabealphabet*) ist,

$\boldsymbol{H}$ eine zu $\boldsymbol{A}$ disjunkte Menge (*Hilfsalphabet*) ist, die mindestens die beiden Elemente s (*Anfangsmarkierung*) und o (*Leerzeichen*) enthält,

Q eine endliche Menge (*Zustandsmenge*) ist, die q_1 (*Anfangszustand*) und die nichtleere Teilmenge Q_0 (der *Stopzustände*) mit $q_1 \notin Q_0$ enthält,

τ eine Abbildung von $(Q \setminus Q_0) \times (\boldsymbol{A} \cup \boldsymbol{H})$ in $Q \times (\boldsymbol{A} \cup \boldsymbol{H}) \times \{\mathrm{L}, \mathrm{N}, \mathrm{R}\}$ (das *Programm*) mit folgenden Eigenschaften ist (für $i = 1, 2, 3$ bezeichne $\tau_i(q, \mathrm{x})$ die i-te Komponente des Tripels $\tau(q, \mathrm{x}) = (q', \mathrm{y}, \mathrm{S})$ $(q, q' \in Q,\ \mathrm{x}, \mathrm{y} \in \boldsymbol{A} \cup \boldsymbol{H} = \boldsymbol{B},\ \mathrm{S} \in \{\mathrm{L}, \mathrm{N}, \mathrm{R}\})$):

a) $\tau_2(q, \mathrm{s}) = \mathrm{s}$ für $q \in Q$ (d. h., s wird nicht gelöscht),

b) $\tau_2(q, \mathrm{x}) \neq \mathrm{s}$ für $\mathrm{x} \neq \mathrm{s}$ (d. h., s wird nicht neu geschrieben),

c) $\tau_3(q, \mathrm{s}) \neq \mathrm{L}$ für $q \in Q$ (d. h., die Maschine darf nicht vom Band herunterlaufen),

d) Ist $\tau_1(q, \mathrm{x}) \in Q_0$, so ist $\mathrm{x} = \mathrm{s}$ und $\tau_3(q, \mathrm{x}) = \mathrm{N}$ (d. h., die Maschine stoppt höchstens auf der Anfangszelle und bleibt dann dort stehen).

Eine konkrete Turingmaschine wird immer durch die Wertetabelle ihres Programms τ gegeben, die wir in rechteckiger Form mit Zeileneingängen für die nicht zu Q_0 gehörigen Zustände und Spalteneingängen für die Buchstaben des Bandalphabets $\boldsymbol{B} = \boldsymbol{A} \cup \boldsymbol{H}$ schreiben. Dabei wird der ersten Zeile immer der Anfangszustand q_1 zugeordnet, und die Stopzustände sind gerade diejenigen Zustände (immer mit q_0 und eventuell oberen Indizes bezeichnet), die zwar im Innern der Tabelle, aber nicht am Zeileneingang auftreten. (Sollte die Maschine noch weitere Stopzustände besitzen, die überhaupt nicht in der Tabelle auftreten, so kann man diese offenbar entfernen, ohne an der Wirkungsweise der Maschine etwas zu ändern.) Somit enthält die Tabelle einer Turingmaschine die Information über alle ihre abstrakten Komponenten, und wir werden daher eine Turingmaschine $(\boldsymbol{A}, \boldsymbol{H}, Q, q_0, Q_0, \mathrm{s}, \mathrm{o}, \tau)$ mit dem Programm bzw. der Tabelle τ im folgenden häufig kurz als Turingmaschine τ bezeichnen.

Beispiel 1. Tabelle einer Turingmaschine.

τ	$\boldsymbol{H}$		$\boldsymbol{A}$	
	s	o	a	b
q_1	q_1 s R	q_2 a L	q_a a R	q_b a R
q_a	—	q_2 a L	q_a a R	q_b a R
q_b	—	q_2 b L	q_a b R	q_b b R
q_2	q_0 s N	—	q_2 a L	q_2 b L

Die durch diese Tabelle gegebene Turingmaschine führt ein beliebiges Eingabewort $\mathrm{W} \in W(\{\mathrm{a}, \mathrm{b}\})$ in folgender Weise in das Wort aW über: Zunächst wird (noch im Zustand q_1) in die rechts neben s befindliche Zelle ein a gedruckt. Dabei „merkt" sich die Maschine den alten Inhalt dieser Zelle (im allgemeinen a oder b), indem sie in den Zustand q_a bzw. q_b übergeht. Die Maschine läuft nun von links nach rechts über das Eingabewort und verschiebt es um eine Zelle nach rechts, indem sie jeweils den Inhalt der i-ten Zelle löscht, sich dann den entsprechenden Zustand merkt und den gelöschten Buchstaben in der $(i+1)$-ten Zelle wieder druckt. In der ersten leeren (d. h. mit o bedruckten) Zelle wird der zuletzt gemerkte Buchstabe abgeladen, worauf die Maschine im Zustand q_2 nach links bis in die Zelle s läuft, wo sie stoppt. Man überprüfe, daß das gewünschte Resultat aW auch im Sonderfall $\mathrm{W} = \Lambda$ erhalten wird, in dem bereits die rechte Nachbarzelle von s leer ist. Die in drei Feldern des Programms stehenden Striche deuten an, daß die entsprechenden Kombinationen von Zustand und Bandsymbol nicht auftreten können. Der konkrete Inhalt dieser Felder hat somit keinen Einfluß auf die Arbeit der Maschine, so daß man sich diese Felder (unter Beachtung der Bedingungen a) bis d) aus Definition 1) beliebig ausgefüllt denken kann. Derartige Striche werden wir auch in Zukunft zur bequemeren Notierung konkreter Programme benutzen. Darüber hinaus vereinbaren wir folgende Abkürzung: Zustände und Bandsymbole, die gegenüber dem vorigen Takt konstant bleiben, werden nicht aufgeschrieben, d. h.

$$\tau(q, \mathrm{x}) = (q', \mathrm{R}) \text{ bedeutet } \tau(q, \mathrm{x}) = (q', \mathrm{x}, \mathrm{R}),$$

$$\tau(q, \mathrm{x}) = (\mathrm{y}, \mathrm{L}) \text{ bedeutet } \tau(q, \mathrm{x}) = (q, \mathrm{y}, \mathrm{L}),$$

$$\tau(q, \mathrm{x}) = \mathrm{R} \text{ bedeutet } \tau(q, \mathrm{x}) = (q, \mathrm{x}, \mathrm{R}) \text{ usw.}$$

Unter Benutzung dieser Abkürzungen ist das obige Programm wie folgt zu schreiben:

τ	H		A	
	s	o	a	b
q_1	R	q_2 a L	q_a R	q_b a R
q_a	—	q_2 a L	R	q_b a R
q_b	—	q_2 b L	q_a b R	R
q_2	q_0 N	—	L	L

Definition 2. Es sei $(\boldsymbol{A}, \boldsymbol{H}, Q, q_1, Q_0, \mathrm{s}, \mathrm{o}, \tau)$ eine Turingmaschine. Eine *Konstellation* bezüglich dieser Turingmaschine ist ein Wort im Alphabet $\boldsymbol{B} \cup (\boldsymbol{B} \times Q)$ $(\boldsymbol{B} = \boldsymbol{A} \cup \boldsymbol{H})$, das mit s oder einem Buchstaben (s, q) beginnt und genau einen Buchstaben aus $\boldsymbol{B} \times Q$ enthält.

Der Begriff der Konstellation (in der angloamerikanischen Literatur mit *instantaneous description*, kurz *ID*, bezeichnet) gibt offenbar genau das Wesentliche einer momentanen Stellung der Turingmaschine wieder, die vorher als Situation bezeichnet wurde, wenn man annimmt, daß die unendlich vielen Zellinhalte, die auf den letzten angegebenen Buchstaben folgen, o sind. Um den Zusammenhang mit dem anschaulichen Situationsbegriff noch deutlicher zu machen, werden wir den Buchstaben der Form (x, q) einer Konstellation, der Zustand und Stellung des LS-Kopfes angibt, als $\overset{q}{\mathrm{x}}$ schreiben. Die Anwendung der in Beispiel 1 angegebenen Maschine auf das Eingabewort baabb kann dann durch folgende endliche Folge von Konstellationen exakt beschrieben werden:

$$\overset{q_1}{\mathrm{s}}\ \mathrm{b\ a\ a\ b\ b}$$

$$\mathrm{s}\ \overset{q_1}{\mathrm{b}}\ \mathrm{a\ a\ b\ b}$$

$$\mathrm{s\ a}\ \overset{q_b}{\mathrm{a}}\ \mathrm{a\ b\ b}$$

$$\mathrm{s\ a\ b}\ \overset{q_a}{\mathrm{a}}\ \mathrm{b\ b}$$

$$\mathrm{s\ a\ b\ a}\ \overset{q_a}{\mathrm{b}}\ \mathrm{b}$$

$$\mathrm{s\ a\ b\ a\ a}\ \overset{q_b}{\mathrm{b}}$$

$$\mathrm{s\ a\ b\ a\ a\ b}\ \overset{q_b}{\mathrm{o}}$$

$$\mathrm{s\ a\ b\ a\ a}\ \overset{q_2}{\mathrm{b}}\ \mathrm{b}$$

$$\mathrm{s\ a\ b\ a}\ \overset{q_2}{\mathrm{a}}\ \mathrm{b\ b}\ \cdot$$

$$\mathrm{s\ a\ b}\ \overset{q_2}{\mathrm{a}}\ \mathrm{a\ b\ b}$$

$$\mathrm{s\ a}\ \overset{q_2}{\mathrm{b}}\ \mathrm{a\ a\ b\ b}$$

$$\mathrm{s}\ \overset{q_2}{\mathrm{a}}\ \mathrm{b\ a\ a\ b\ b}$$

$$\overset{q_2}{\mathrm{s}}\ \mathrm{a\ b\ a\ a\ b\ b}$$

$$\overset{q_0}{\mathrm{s}}\ \mathrm{a\ b\ a\ a\ b\ b}$$

Definition 3. Bezüglich einer Turingmaschine τ sei K eine Konstellation, die den Buchstaben $\overset{q}{x}$ enthält. Falls $q \notin Q_0$ ist, sei die *Folgekonstellation* $\tau(K)$ durch folgende Tabelle definiert:

K	$\tau(K)$	
$\mathrm{U}\overset{q}{\mathrm{x}}\mathrm{V}$	$\mathrm{U}\overset{q'}{\mathrm{y}}\mathrm{V}$	falls $\tau(q, \mathrm{x}) = (q', \mathrm{y}, \mathrm{N})$
$\mathrm{U}\overset{q}{\mathrm{z}}\mathrm{x}\mathrm{V}$	$\mathrm{U}\overset{q'}{\mathrm{z}}\mathrm{y}\mathrm{V}$	falls $\tau(q, \mathrm{x}) = (q', \mathrm{y}, \mathrm{L})$[1]
$\mathrm{U}\overset{q}{\mathrm{x}}\mathrm{z}\mathrm{V}$	$\mathrm{U}\mathrm{y}\overset{q'}{\mathrm{z}}\mathrm{V}$	falls $\tau(q, \mathrm{x}) = (q', \mathrm{y}, \mathrm{R})$
$\mathrm{U}\overset{q}{\mathrm{x}}$	$\mathrm{U}\mathrm{y}\overset{q'}{\mathrm{o}}$	

[1]) Dann ist $\mathrm{x} \neq \mathrm{s}$ nach Definition 1, Bedingung c).

Falls $q \in Q_0$ ist, heiße K *Endkonstellation*, und $\tau(K)$ sei in diesem Fall nicht definiert.

Es sei $\tau^1(K) = \tau(K)$; ist die Konstellation $\tau^n(K)$ schon definiert und keine Endkonstellation von τ, so sei $\tau^{n+1}(K) = \tau(\tau^n(K))$. Jede Turingmaschine τ mit dem Ein- und Ausgabealphabet A realisiert (berechnet) eine wie folgt definierte Wortfunktion f_τ in $W(A)$:

$f_\tau(\mathrm{W}) := \mathrm{V}$ genau dann, wenn eine natürliche Zahl n, ein Hilfsbuchstabe $\mathrm{x} \notin A$ (eventuell $\mathrm{x} = \mathrm{o}$) und ein (eventuell leeres) Wort $\mathrm{R} \in W(A \cup H)$ existieren, so daß

$$\tau^n(\overset{q_1}{\mathrm{s}}\mathrm{W}) = \overset{q_0}{\mathrm{s}}\mathrm{V} \quad \text{oder} \quad \tau^n(\overset{q_1}{\mathrm{s}}\mathrm{W}) = \overset{q_0}{\mathrm{s}}\mathrm{VxR},$$

wobei q_0 ein Stopzustand von τ ist (d. h., $\tau^n(\overset{q_1}{\mathrm{s}}\mathrm{W})$ ist eine Endkonstellation von τ). Eine Wortfunktion f in $W(A)$ heißt *Turing-berechenbar*, wenn eine Turingmaschine τ mit dem Ein- und Ausgabealphabet A existiert, so daß $f = f_\tau$ ist.

Die Bestimmung, daß in der bei Abarbeitung der Konstellation $\overset{q_1}{\mathrm{s}}\mathrm{W}$ erhaltenen

Endkonstellation auf das Ausgabewort $V \in W(A)$ ein Rest folgen darf, der dann aber mit einem Buchstaben $x \notin A$ beginnen muß, präzisiert die schon in der anschaulichen Beschreibung der Turingmaschinen getroffene Festsetzung, daß derjenige Teil des letzten Bandinhalts als Resultat gedeutet wird, der in der rechten Nachbarzelle von s beginnt und in der linken Nachbarzelle des ersten Buchstaben $x \notin A$ endet. Insbesondere ist also $f_\tau(W) = \Lambda$, falls für ein n

$$\tau^n(\overset{q_1}{s}\,W) = \overset{q_0}{s}\,x\,R$$

gilt, wobei $x \notin A$ ist. Unter Bezug auf Definition 3 können wir der in 7.3. vorläufig formulierten Hypothese von Church folgende präzisierte Fassung geben:

Jede im anschaulichen Sinn effektiv berechenbare Wortfunktion ist Turing-berechenbar.

Definition 4. Konstellationen K_1, K_2 einer Turingmaschine heißen *äquivalent*, wenn eine natürliche Zahl $m \geqq 0$ existiert, so daß $K_1 = K_2\underbrace{o \dots o}_{m}$ oder $K_2 = K_1\underbrace{o \dots o}_{m}$ ist.

Demnach sind äquivalente Konstellationen nichts anderes als verschiedene endliche Beschreibungen des gleichen (unendlichen) Bandinhalts nebst Zustand und Stellung des LS-Kopfes. Es ist leicht zu sehen, daß die Eigenschaften Reflexivität, Symmetrie und Transitivität einer Äquivalenzrelation hier tatsächlich erfüllt sind. Ferner ergibt sich aus Definition 3, daß für äquivalente Konstellationen K_1, K_2 und beliebige natürliche Zahlen n genau dann $\tau^n(K_1)$ existiert, wenn $\tau^n(K_2)$ existiert, und daß im Fall der Existenz beide Konstellationen ebenfalls äquivalent sind. Insbesondere ist daher $f_\tau(W) = V$ genau dann, wenn eine natürliche Zahl n, ein Hilfsbuchstabe $x \notin A$ und ein Wort $R \in W(A \cup H)$ existieren, so daß für jede Konstellation K der Form $\overset{q_1}{s}\,W o \dots o$ die Endkonstellation $\tau^n(K)$ zur Konstellation $\overset{q_0}{s}\,VxR$ äquivalent ist (mithin genau dann zu $\overset{q_0}{s}\,V$ äquivalent ist, falls hierbei $xR \in W(\{o\})$ gilt). Unter Bezug auf diese Tatsachen werden wir im folgenden häufig äquivalente Konstellationen identifizieren.

Wir weisen nun die Turing-Berechenbarkeit einiger einfacher Wortfunktionen nach, indem wir jeweils ein die betreffende Funktion realisierendes Turingprogramm angeben. Als Ein- und Ausgabealphabet benutzen wir dabei der Einfachheit halber in allen Fällen $A = \{a, b\}$. Die Beispiele sind jedoch so beschaffen, daß man ohne weiteres sieht, wie die Turing-Berechenbarkeit der für andere Alphabete analog definierten Wortfunktionen zu zeigen ist. Dem Leser sei dringend empfohlen, für einige weitere Wortfunktionen selbständig Turingprogramme zu konstruieren.

Beispiel 2. Die nirgends definierte Wortfunktion in $W(A)$ wird durch die Maschine

	s	o	a	b
q_1	R	R	R	R

realisiert, der außerdem ein in der Tabelle selbst nicht vorkommender Stopzustand q_J zuzuordnen ist.

Beispiel 3. Die Funktion $f(\mathrm{W}) := \Lambda$ wird durch die Maschine

	s	o	a	b
q_1	R	q_2 L	o R	o R
q_2	q_0 N	L	—	—

realisiert.

Beispiel 4. Die Funktion $f(\mathrm{W}) := \mathrm{Wa}$ wird durch die Maschine

	s	o	a	b
q_1	R	q_2 a L	R	R
q_2	q_0 N	—	L	L

realisiert.

Beispiel 5. Es sei $f(\Lambda) = \Lambda$ und, falls $f(\mathrm{W})$ schon definiert ist, $f(\mathrm{Wx}) = \mathrm{x}f(\mathrm{W})$ für $\mathrm{x} \in \boldsymbol{A}$, d. h., $f(\mathrm{W})$ ist das „umgedrehte“ Wort W, im folgenden mit $\overleftarrow{\mathrm{W}}$ bezeichnet. Diese Wortfunktion f wird durch die folgende Maschine realisiert:

	s	o	a	b
q_1	R	q_6 L	q_a o R	q_b o R
q_a	—	q_2 a L	R	R
q_b	—	q_2 b L	R	R
q_2	—	q_4 R	q_a' o L	q_b' o L
q_a'	—	q_3 a R	L	L
q_b'	—	q_3 b R	L	L
q_3	—	q_4 R	q_a o R	q_b o R
q_4	—	q_5 L	R	R
q_5	—	—	q_a'' o L	q_b'' o L
q_a''	—	q_6 a L	L	q_b'' a L
q_b''	—	q_6 b L	q_a'' b L	L
q_6	q_0 N	—	L	L

Definition 5. Eine Turingmaschine bzw. ihr Programm τ heißen *sauber*, wenn für alle $W \in W(A)$ (A Ein- und Ausgabealphabet) gilt: Falls $V = f_\tau(W)$ existiert, hat die Endkonstellation die Form $\overset{q_0}{s}\,Vo\ldots$, d. h., der in Definition 3 zugelassene nichtleere Rest R des Bandinhalts enthält keine vom Leerzeichen o verschiedenen Buchstaben.

Eine saubere Turingmaschine hinterläßt demnach keine Spuren von Zwischenrechnungen auf dem Band. Alle bisher betrachteten Beispielprogramme sind sauber.

Definition 6. Turingprogramme τ_1, τ_2 mit gleichem Ein- und Ausgabealphabet A (aber eventuell verschiedenen Hilfsalphabeten) heißen *äquivalent* ($\tau_1 \cong \tau_2$), wenn $f_{\tau_1} = f_{\tau_2}$ ist.

Satz 1. *Zu jedem Turingprogramm τ kann man ein äquivalentes sauberes Turingprogramm τ konstruieren.*

Beweis. Die Idee besteht darin, die eventuell zunächst nicht saubere Turingmaschine nach Abschluß der eigentlichen Berechnung ein letztes Mal nach rechts laufen zu lassen, wobei alle hinter dem letzten Buchstaben des Resultats gefundenen, von o verschiedenen Buchstaben gelöscht werden. Die Maschine kann jedoch zu diesem Zeitpunkt im allgemeinen nicht mehr „wissen", wie weit diese Suche fortzusetzen ist, bevor sie endgültig nach s zurückkehren kann, da die eigentlichen Stücke des zu beseitigenden „Abfalls" im allgemeinen durch beliebig lange Lücken aus Leerzeichen getrennt sein können. Um diese Schwierigkeit zu umgehen, ersetze man im ursprünglichen Programm τ jede Anweisung, o zu drucken, durch die Anweisung, ein gewisses neues Hilfszeichen $*$ zu drucken, welches beim Lesen in allen Fällen wie o behandelt wird. Dadurch bilden die im Laufe der Arbeit der Maschine jemals berührten Speicherzellen ein zusammenhängendes (d. h. nicht durch Leerzeichen getrenntes) Wort. Für jeden Stopzustand q_0 des ursprünglichen Programms hänge man nun folgende Zeilen an das Programm an:

	s	o	$*$	$x \in H$ $x \neq s, o, *$	$x \in A$
q_0	R	q_0^2 L	q_0^1 o R	q_0^1 o R	R
q_0^1	—	q_0^2 L	o R	o R	o R
q_0^2	q_0' N	L	—	—	L

und nehme $\{q_0' : q_0 \in Q_0\}$ als neue Stopzustände.

Satz 2. *Sind f_1 und f_2 Turing-berechenbare Wortfunktionen in $W(A)$, so ist auch $f_2 \circ f_1$ Turing-berechenbar. Ein Turingprogramm zur Berechnung von $f_2 \circ f_1$ läßt sich nach einem einheitlichen Verfahren aus gegebenen Programmen für f_1 und f_2 aufstellen.*

Beweis. Wir können annehmen, daß das gegebene Programm τ_1 zur Berechnung von f_1 sauber ist und nur einen Stopzustand besitzt. (Andernfalls ersetze man alle Stopzustände von τ_1 durch einen einzigen, wodurch man offenbar ein äquivalentes

Programm erhält.) Ist $f_2 = f_{\tau_2}$, so kann man durch entsprechende Umbenennung der Zustände (eigentlich Übergang zu einem isomorphen und daher erst recht äquivalenten Programm) erreichen, daß der Anfangszustand von τ_2 mit dem Stopzustand q_0 von τ_1 übereinstimmt und die übrigen Zustände von τ_2 keine Zustände von τ_1 sind. Schließlich können wir erreichen, daß beide Programme das gleiche Hilfsalphabet $\boldsymbol{H}$ benutzen, indem wir $\boldsymbol{H} = \boldsymbol{H}_1 \cup \boldsymbol{H}_2$ setzen, wobei $\boldsymbol{H}_i$ das Hilfsalphabet von τ_i ist, und $\tau_1(q, \mathrm{x})$ beliebig für $\mathrm{x} \in \boldsymbol{H}_2 \setminus \boldsymbol{H}_1$ bzw. $\tau_2(q, \mathrm{y})$ beliebig für $\mathrm{y} \in \boldsymbol{H}_1 \setminus \boldsymbol{H}_2$ definieren (d. h. an den entsprechenden Stellen der Tabellen Striche setzen). Schreibt man unter all diesen Voraussetzungen die Zeilen des Programms τ_2 einfach unter die Programmzeilen von τ_1, nimmt den Anfangszustand von τ_1 als neuen Anfangszustand und die Stopzustände von τ_2 als neue Stopzustände, so erhält man schon ein Programm τ zur Berechnung der Funktion $f_2 \circ f_1$. Für $\mathrm{W} \in W(A)$ stimmen die Konstellationen $\tau^{\pi}\bigl(\overset{q_1}{\mathrm{s}}\,\mathrm{W}\bigr)$ so lange mit $\tau_1^{\pi}\bigl(\overset{q_1}{\mathrm{s}}\,\mathrm{W}\bigr)$ überein, bis eventuell der Zustand q_0 erreicht wird. Da τ_1 sauber ist, ist die entsprechende Konstellation äquivalent zu $\overset{q_0}{\mathrm{s}} f_1(\mathrm{W})$. Auf eine solche Konstellation wirkt nun τ genau wie τ_2, so daß genau dann nach endlich vielen Takten ein Stopzustand von τ_2 (und damit von τ) erreicht wird, wenn $f_2\bigl(f_1(\mathrm{W})\bigr)$ existiert. In diesem Fall hat die Endkonstellation bei Anwendung von τ auf W demnach die Gestalt $\overset{q}{\mathrm{s}} f_2\bigl(f_1(\mathrm{W})\bigr)\,\mathrm{R}$, wobei q ein Stopzustand von τ_2 ist und R leer ist oder mit einem Hilfsbuchstaben beginnt. (τ_2 ist nicht als sauber vorausgesetzt.)

Beispiel 6. Aus den in den Beispielen 3 und 4 angegebenen Programmen zur Berechnung von $f_1(\mathrm{W}) = \Lambda$ und $f_2(\mathrm{W}) = \mathrm{Wa}$ soll ein Programm zur Berechnung von $f_3(\mathrm{W}) = f_2(f_1(\mathrm{W})) = \mathrm{a}$ zusammengesetzt werden. Da beide Programme die gleichen Hilfsbuchstaben benutzen und das angegebene Programm für f_1 bereits sauber ist, brauchen wir die Programme nur untereinanderzuschreiben, wobei die Bezeichnungen der Zustände des zweiten Programms entsprechend abzuändern sind:

	s	o	a	b	
q_1	R	q_2 L	o R	o R	Programm für f_1
q_2	q_0 N	L	—	—	
q_0 (statt q_1)	R	q_3 a L	R	R	
q_3 (statt q_2)	q_0' N	—	L	L	Programm für f_2

(q_0' statt q_0 neuer Stopzustand).

Aus dem allgemeinen Zusammenhang zwischen Berechenbarkeit und Entscheidbarkeit (vgl. 7.3.) ergibt sich

Definition 7. Eine Menge $\boldsymbol{M} \subseteq W(\boldsymbol{A})$ heißt *Turing-entscheidbar*, wenn ihre charakteristische Funktion Turing-berechenbar ist, d. h. wenn es eine Turingmaschine τ mit dem Ein- und Ausgabealphabet $\boldsymbol{A}$ gibt, so daß

$$f_\tau(\mathrm{W}) = \begin{cases} \Lambda, & \textit{falls} \quad \mathrm{W} \in \boldsymbol{M}, \\ \mathrm{a}, & \textit{falls} \quad \mathrm{W} \in W(\boldsymbol{A}) \setminus \boldsymbol{M} \end{cases} \qquad (\mathrm{a} \in \boldsymbol{A} \text{ beliebig})$$

gilt.

Für den Einbau von Entscheidungsverfahren als Unterprogramme in kompliziertere Turingprogramme erweist sich jedoch eine Variante dieses Begriffs als günstiger, bei der überdies die Willkür in der Wahl der beiden Werte der charakteristischen Funktion wegfällt.

Satz 3. *Eine Menge* $\boldsymbol{M} \subseteq W(\boldsymbol{A})$ *ist Turing-entscheidbar genau dann, wenn es ein Turingprogramm* τ *mit dem Ein- und Ausgabealphabet* A *und genau zwei Stopzuständen* q_0^+ *und* q_0^- *gibt, so daß für* $\mathrm{W} \in W(\boldsymbol{A})$

$$\tau^n\begin{pmatrix} q_1 \\ \mathrm{s\,W} \end{pmatrix} = \overset{q_0^+}{\mathrm{s}}\,\mathrm{W} \text{ für eine natürliche Zahl } n, \textit{ falls } \mathrm{W} \in \boldsymbol{M},$$

$$\tau^n\begin{pmatrix} q_1 \\ \mathrm{s\,W} \end{pmatrix} = \overset{q_0^-}{\mathrm{s}}\,\mathrm{W} \text{ für eine natürliche Zahl } n, \textit{ falls } \mathrm{W} \in W(\boldsymbol{A}) \setminus \boldsymbol{M}.$$

Ein Programm τ der in Satz 3 beschriebenen Art heißt ein *Entscheidungsprogramm* (*für* $\boldsymbol{M}$).

Wir beweisen Satz 3. Zunächst sei τ ein Entscheidungsprogramm für $\boldsymbol{M}$. Wir konstruieren aus τ ein Programm zur Berechnung der charakteristischen Funktion von $\boldsymbol{M}$, indem wir an τ folgende Zeilen anfügen:

	s	o	$\mathrm{x} \in \boldsymbol{H}$, $\mathrm{x} \neq \mathrm{s}, \mathrm{o}$	$\mathrm{x} \in \boldsymbol{A}$
q_0^+	R	q^1 L	—	o R
q^1	q_0 N	L	—	—
q_0^-	R	q^2 L	—	o R
q^2	q^3 R	L	—	—
q^3	—	q^4 a L	—	—
q^4	q_0 N	—	—	—

Dabei sollen q^1, q^2, q^3, q^4, q_0 neue (d. h. in τ nicht vorkommende) Zustände sein; q_0 sei der Stopzustand des neuen Programms. Die Wirkungsweise ist klar. Ebenso klar ist, daß man, ausgehend von einem Entscheidungsprogramm für $\boldsymbol{M}$, nach Belieben Unterprogramme an q_0^+ bzw. q_0^- anhängen kann, die das Resultat der Entscheidung auf andere Weise als durch Λ bzw. a auf dem Band sichtbar werden lassen. Zum

Beispiel kann man die Wörter ja bzw. nein ausdrucken lassen, falls die dazu benötigten Buchstaben zu A gehören. Die so berechnete Wortfunktion ist in gewissem Sinne auch eine „charakteristische Funktion“ von M. Allgemein kann jede auf $W(A)$ voll definierte Wortfunktion, die genau zwei Werte annimmt, als charakteristische Funktion gedeutet werden.

Es sei nun τ ein Programm zur Berechnung der charakteristischen Funktion (im ursprünglichen Sinne) einer Menge M. Wir können annehmen, daß τ sauber ist, den Anfangszustand q_1, genau einen Stopzustand q_0 und das Hilfsalphabet H hat. Das Programm τ' geht aus τ hervor, indem man überall s durch einen neuen Hilfsbuchstaben $*$ ersetzt. Aus τ' konstruieren wir nun ein Entscheidungsprogramm für M, wobei wir hier den Fall $A = \{a, b\}$ explizit aufschreiben. Es ist dann leicht zu sehen, wie man im Fall eines anderen Ein- und Ausgabealphabets zu verfahren hat (vgl. die folgende Tabelle). In kurzen Worten läßt sich die Wirkungsweise des so ergänzten Programms τ' wie folgt beschreiben: Zunächst wird hinter das Eingabewort

	s	o	$*$	$x \in H$ $x \neq$ s, o, $*$	a	b	
q^1	R	$q_1 * N$	—	—	q_a o R	q_b o R	führt $\overset{q^1}{\mathrm{s}}$ o in s $\overset{q_1}{*}$
q_a	—	q^2 a R	—	—	R	q_b a R	
q_b	—	q^2 b R	—	—	q_a b R	R	und $\overset{q^1}{\mathrm{s}}$ W x in
q^2	—	$q^4 *$ L	—	—	—	—	
q^3	—	—	q^4 L	—	L	L	s W x o $\overset{q_1}{*}$ W x für
q^4	—	q_1 R	—	—	q^5 L	q^5 L	$x \in \{a, b\}$, $W \in W(\{a, b\})$
q^5	—	q^6 R	—	—	L	L	über
q^6	—	—	—	—	q_a^1 o L	q_b^1 o L	
q_a^1	—	q_a^2 a R	—	—	—	—	τ' führt
q_b^1	—	q_b^2 b R	—	—	—	—	
q_a^2	—	R	q_a^3 R	—	R	R	s … $\overset{q_1}{*}$ W in
q_b^2	—	R	q_b^3 R	—	R	R	s … $\overset{q_0}{*}$ o für $W \in M$,
q_a^3	—	q^3 a L	—	—	R	R	
q_b^3	—	q^3 b L	—	—	R	R	s … $\overset{q_0}{*}$ a für $W \notin M$
							über
			τ'				
q_0	—	q^7 L	R	—	q^9 o L	—	führt s W o $\overset{q_0}{*}$ o
q^7	—	—	q^8 o L	—	—	—	
q^8	q_0^+ N	L	—	—	L	L	in $\overset{q_0^+}{\mathrm{s}}$ W und
q^9	—	—	q^{10} o L	—	—	—	
q^{10}	q_0^- N	L	—	—	L	L	s W o $\overset{q_0}{*}$ a in $\overset{q_0^-}{\mathrm{s}}$ W über

W ein Trennzeichen $*$ geschrieben. Ist $W \neq \Lambda$, so wird dabei gleichzeitig W um eine Zelle nach rechts verschoben, so daß zwischen s und dem ersten Buchstaben von W ein Leerzeichen o steht. Hinter $*$ wird nun W buchstabenweise noch einmal abgeschrieben, wobei nach jedem Abschreiben eines Buchstaben das anfangs vorn stehende Leerzeichen um eine Stelle nach rechts verschoben wird, um so beim nächsten Rücklauf die Stellung des zu übertragenden Buchstaben zu markieren. Dieser Prozeß wird schließlich dadurch beendet, daß einmal beim Rücklauf das Leerzeichen unmittelbar links von * gefunden wird. Auf das hinter * stehende zweite Exemplar von W wirkt nun τ' genauso wie τ, wobei lediglich * statt s als „Pseudoanfangszelle“ benutzt wird, so daß das links von * stehende erste Exemplar von W dabei nicht berührt wird. Die Berechnung der charakteristischen Funktion endet damit, daß der LS-Kopf im Zustand q_0 (dem Stopzustand von τ bzw. τ') auf * stehenbleibt. Ausgehend von dieser Situation wird durch die letzten Programmzeilen getestet, ob rechts neben * das Zeichen o oder a steht. Je nachdem, welches Zeichen dort steht, kehrt die Maschine in zwei parallel laufenden Unterprogrammen zur Zelle s zurück, wobei im ersten Fall *, im zweiten Fall *a gelöscht wird und im ersten Fall der Stopzustand q_0^+, im zweiten Fall der Stopzustand q_0^- erreicht wird.

Beispiel 7. Wir wollen die Turing-Entscheidbarkeit der Menge aller Wörter $W \in W(\{a, b\})$ nachweisen, die die Buchstaben a und b in gleicher Anzahl enthalten. Dazu geben wir zunächst ein Programm zur Berechnung der charakteristischen Funktion dieser Menge an. Dem hierin steckenden eigentlichen Prüfprozeß liegt die Idee zugrunde, daß die Maschine durch Umordnen der Buchstaben eines beliebigen Eingabewortes W versucht, dieses auf die Form abab ... ab zu bringen. So weit, wie das Eingabewort bereits diese Bauart hat, läuft die Maschine von links nach rechts darüber hinweg, wobei sie abwechselnd die Zustände q_1 und q_3 annimmt. Im Zustand q_1 „erwartet“ sie ein a und im Zustand q_3 ein b. Falls einmal nicht der gewünschte sondern der andere Buchstabe gefunden wird, löscht die Maschine diesen und sucht rechts davon den ersten zum Austausch brauchbaren Buchstaben. Der Prozeß bricht genau dann mit positivem Resultat ab, wenn im Zustand q_1 statt a das Leerzeichen o gefunden wird. Das Programm hat demnach folgende Gestalt:

	s	o	a	b	
q_1	R	q_2 L	q_3 R	q_4 o R	a erwarten
q_2	q_0 N	L	o L	o L	positiver Ausgang, Eingabe löschen
q_3	—	q_8 L	q_6 o R	q_1 R	b erwarten
q_4	—	q_8 L	q_5 b L	R	a zum Austausch suchen
q_5	—	q_3 a R	—	L	Rücklauf dazu
q_6	—	q_8 L	R	q_7 a L	b zum Austausch suchen
q_7	—	q_1 b R	L	—	Rücklauf dazu
q_8	q_9 R	L	o L	o L	negativer Ausgang, Eingabe löschen
q_9	—	q_2 a L	—	—	a drucken

Der Leser führe dieses Programm nach dem im Beweis von Satz 3 angegebenen Verfahren in ein Entscheidungsprogramm über und prüfe dieses an den Eingabewörtern Λ, ab, aba und abbab.

Satz 4. *Ist $M \subseteqq W(A)$ Turing-entscheidbar und sind f_1 und f_2 Turing-berechenbare Wortfunktionen in $W(A)$, so ist die durch*

$$f_3(\mathrm{W}) := \begin{cases} f_1(\mathrm{W}), & \textit{falls } \mathrm{W} \in M \textit{ und } f_1(\mathrm{W}) \textit{ existiert,} \\ f_2(\mathrm{W}), & \textit{falls } \mathrm{W} \in W(A) \setminus M \textit{ und } f_2(\mathrm{W}) \textit{ existiert,} \\ \textit{nicht definiert andernfalls} \end{cases}$$

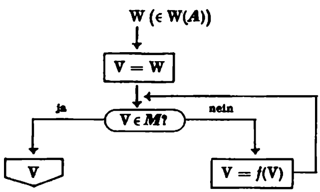
definierte Wortfunktion f_3 Turing-berechenbar. Aus Turingprogrammen zur Entscheidung von M (bzw. zur Berechnung der charakteristischen Funktion von M) und zur Berechnung von f_1 und f_2 kann nach einem einheitlichen Verfahren ein Turingprogramm zur Berechnung von f_3 konstruiert werden.

Beweis. Es sei τ_0 ein Entscheidungsprogramm für M mit dem Anfangszustand q_1 und den Stopzuständen q_0^+ und q_0^-. Ferner seien τ_1 bzw. τ_2 Turingprogramme zur Berechnung von f_1 bzw. f_2, wobei τ_1 den Anfangszustand q_0^+ und den Stopzustand q_0 und τ_2 den Anfangszustand q_0^- und den Stopzustand q_0 hat und die drei Programme außer den genannten Fällen keine Zustände gemeinsam haben. Wir können weiter voraussetzen, daß alle drei Programme das gleiche Hilfsalphabet benutzen, indem wir sie andernfalls für die ursprünglich nicht benutzten Buchstaben des kleinsten gemeinsamen Erweiterungsalphabets beliebig (d. h. durch Striche in der Tafel) fortsetzen. Unter diesen Voraussetzungen ist offenbar schon $\tau_0 \cup \tau_1 \cup \tau_2$ ein Programm zur Berechnung von f_3, das man also im konkreten Fall einfach durch Aneinanderreihen der Programmzeilen der drei Programme erhält. Als Anfangszustand für das Gesamtprogramm dient dabei q_1, als Stopzustand q_0.

Satz 5. *Ist $M \subseteqq W(A)$ und f eine Wortfunktion in $W(A)$, so sei f^M die durch*

$$f^M(\mathrm{W}) := \begin{cases} f^n(\mathrm{W}), \textit{ falls } f^0(\mathrm{W}) = \mathrm{W},\ f^1(\mathrm{W}) = f(\mathrm{W}), \ldots, f^n(\mathrm{W}) = f(f^{n-1}(\mathrm{W})) \\ \qquad \textit{existieren, } f^i(\mathrm{W}) \notin M \textit{ für } i < n,\ f^n(\mathrm{W}) \in M, \\ \textit{nicht definiert andernfalls} \end{cases}$$

definierte bzw. durch das Flußdiagramm



dargestellte Wortfunktion in $W(\boldsymbol{A})$. Ist $\boldsymbol{M}$ Turing-entscheidbar und f Turing-berechenbar, so ist $f^{\boldsymbol{M}}$ Turing-berechenbar, und ein Programm zur Berechnung von $f^{\boldsymbol{M}}$ kann aus gegebenen Programmen zur Entscheidung von $\boldsymbol{M}$ und Berechnung von f nach einem einheitlichen Verfahren konstruiert werden.

Beweis. Es sei τ_0 ein Entscheidungsprogramm für $\boldsymbol{M}$ mit dem Anfangszustand q_1 und den Stopzuständen $q_0{}^+$ und $q_0{}^-$, τ_1 ein sauberes Programm zur Berechnung von f mit dem gleichen Hilfsalphabet, dem Anfangszustand $q_0{}^-$ und dem Stopzustand q_1, so daß beide Programme keine weiteren Zustände gemeinsam haben. Dann ist $\tau_0 \cup \tau_1$ mit dem Anfangszustand q_1 und dem Stopzustand $q_0{}^+$ ein Programm zur Berechnung von $f^{\boldsymbol{M}}$.

Die in den Beweisen der Sätze 2 bis 5 zusammengefaßten Methoden zur Verdopplung von Eingabewörtern und Verknüpfung von Programmen bezeichnet man als *Unterprogrammtechnik* für Turingmaschinen. Das Wesentliche dieser Unterprogrammtechnik besteht in der geeigneten Identifizierung von Anfangs- und Stopzuständen verschiedener Programme mit im übrigen disjunkten Zustandsmengen. Die Einfachheit dieser Technik (insbesondere im Vergleich zur analogen Programmverknüpfung für andere Präzisierungen des Algorithmenbegriffs) ist einer der Gründe für die Bevorzugung der Turingmaschinen.

Ist $\boldsymbol{A}$ ein beliebiges Alphabet und sind $|$, $*$ zwei nicht in $\boldsymbol{A}$ enthaltene Zeichen, so sei q_i durch $\underbrace{|\ldots|}_{i+1}$ verschlüsselt. Die in einem beliebigen Turingprogramm mit dem Ein- und Ausgabealphabet $\boldsymbol{A}$ benutzten Hilfsbuchstaben kann man in der Form h_1 $(= \mathrm{s})$, h_2 $(= \mathrm{o})$, h_3, h_4, ... annehmen und daher h_i durch $\underbrace{*\ldots*}_{i}$ verschlüsseln. Dann entspricht jedem möglichen Quintupel eines Turingprogramms τ mit dem Ein- und Ausgabealphabet $\boldsymbol{A}$, d. h. jeder Vorschrift der Form

$$\tau(q_i, \mathrm{x}) = (q_j, \mathrm{y}, \mathrm{S})$$

ein Kodewort $\underbrace{|\ldots|}_{i+1}\mathrm{X}\underbrace{|\ldots|}_{j+1}\mathrm{YS}$ im festen Alphabet $\boldsymbol{B} = \boldsymbol{A} \cup \{|, *, \mathrm{L}, \mathrm{N}, \mathrm{R}\}$, wobei $\mathrm{X}, \mathrm{Y} \in \boldsymbol{A} \cup W(\{*\})$ ist. Die Gesamtheit aller dieser Kodewörter bildet ein System von Pseudoatomen (vgl. 7.1.), d. h., man kann jedes als Verkettung solcher Wörter geschriebene Wort eindeutig wieder in seine Bestandteile zerlegen. Da die Reihenfolge bzw. Anordnung der einzelnen Befehle eines Turingprogramms überhaupt keine Rolle spielt (formal mengentheoretisch ist ja τ als Abbildung nichts anderes als eine ungeordnete Menge von Quintupeln), läßt sich jedes Programm bezüglich des festen Alphabets $\boldsymbol{A}$ durch ein Kodewort im Alphabet $\boldsymbol{B}$ darstellen, wobei jedoch diese Darstellung nicht eindeutig bestimmt, sondern die Reihenfolge der darin vorkommenden Pseudoatome beliebig ist. Zum Beispiel ist

$$||*||*\mathrm{R}||**|||**\mathrm{L}||\mathrm{a}||**\mathrm{R}||\mathrm{b}||**\mathrm{R}|||*|*\mathrm{N}|||**|||**\mathrm{L}$$

ein Kodewort für das in Beispiel 3 angegebene Turingprogramm. Wir haben es hier mit dem allgemeinsten Typ einer Kodierung (vgl. 7.1.) zu tun, d. h., nicht alle (sogar nur sehr wenige) Wörter aus $W(\boldsymbol{B})$ sind Kodewörter eines Turingprogramms, und viele verschiedene Kodewörter haben jeweils die gleiche Bedeutung. Das letzte könnte man zunächst vermeiden, indem man irgendeine kanonische Reihenfolge der einzelnen Quintupel eines Programms vorschreibt. Faßt man jedoch die Kodewörter als Kodewörter der durch das jeweilige Programm berechneten Wortfunktion auf, so läßt sich die Mehrdeutigkeit nicht mehr vermeiden, da ja eine überhaupt Turing-berechenbare Wortfunktion durch viele verschiedene Programme realisierbar ist, deren Äquivalenz nicht einmal entscheidbar ist. Es ist jedoch (zunächst im anschaulichen Sinne) entscheidbar, ob ein gegebenes Wort aus $W(\boldsymbol{B})$ Kodewort einer Turing-berechenbaren Wortfunktion in $W(\boldsymbol{A})$ ist. Hierfür ist wichtig, daß man die in Definition 1 formulierten Bedingungen a) bis d) am Kodewort eines Programms effektiv überprüfen kann und daß nach unseren Festsetzungen über die Deutung des Resultats bei einem eventuell nicht sauberen Programm jedes Turingprogramm eine Wortfunktion definiert. Es ist nun nicht schwer, durch Umkodierung eine Niederschrift $\langle\tau\rangle$ der über dem Ein- und Ausgabealphabet $\boldsymbol{A}$ operierenden Turingprogramme τ so zu definieren, daß $\langle\tau\rangle \in W(\boldsymbol{A})$ ist und die oben diskutierte Entscheidbarkeit erhalten bleibt. Damit ist eine Grundforderung, die in 7.3. an den präzisierten Berechenbarkeitsbegriff gestellt wurde, erfüllt. Bezüglich eines solchen Niederschriftsbegriffs für Turingprogramme (für den es natürlich viele konkrete Möglichkeiten gibt) kann man nun unter Benutzung der Unterprogrammtechnik auch eine universelle Turingmaschine explizit angeben. Die Konstruktion einer solchen Maschine, deren Angabe den Rahmen dieses Buches überschreiten würde, erfordert weit mehr Schreib- als Kopfarbeit. Der interessierte Leser sei auf [19] verwiesen.

7.5. Andere Präzisierungen des Berechenbarkeitsbegriffes und deren Gleichwertigkeit

1. Varianten des Begriffs der Turingmaschine. In der Theorie der Turingmaschinen wurden und werden zahlreiche Varianten des in 7.4. benutzten Formalismus der Turingmaschinen untersucht. Dabei stellte sich bisher in allen Fällen die prinzipielle Gleichwertigkeit hinsichtlich der Berechenbarkeit von Wortfunktionen heraus. Das „Umsteigen" von einem Programm bezüglich einer gewissen Variante zu einem gleichwertigen Programm bezüglich einer anderen Variante erfordert im allgemeinen lediglich eine größere Zahl von Zuständen oder Takten bzw. gestattet die Verkleinerung der Zustandszahl oder Rechenzeit. Daher werden derartige Untersuchungen schon seit einiger Zeit nicht mehr unter dem Aspekt durchgeführt, die Hypothese von Church zu bestätigen oder gar zu widerlegen, sondern im Zusammen-

hang mit dem sehr komplizierten Begriff der Kompliziertheit eines Algorithmus oder einer Wortfunktion. Ohne daß hier auf Einzelheiten solcher Untersuchungen eingegangen werden kann, sei erwähnt, daß man zwischen der *Berechnungskompliziertheit* eines Algorithmus, d. h. der zur Erreichung des Resultats (in Abhängigkeit vom Eingabewort) benötigten Schrittzahl, und der (etwa durch die Anzahl der Regeln bzw. der Zustände eines Turingprogramms gegebenen) *Beschreibungskompliziertheit* eines Algorithmus zu unterscheiden hat. Aus der Fülle der noch nicht oder nicht restlos geklärten Probleme in diesem Bereich sei die Frage angeführt, ob bzw. wie man sinnvoll die Kompliziertheit einer Wortfunktion unabhängig vom Berechnungsformalismus definieren kann. Wir skizzieren nun einige wesentliche Varianten des Begriffs der Turingmaschine, wobei wir den in 7.4. definierten Maschinentyp als *Standard-Turingmaschine* bezeichnen werden.

1.1. *Turingmaschinen mit beidseitig unbegrenztem Speicherband.* Es ist klar, daß man jede Standard-Turingmaschine auf einer solchen Maschine simulieren kann, indem man eine beliebige Zelle des beidseitig unbegrenzten Speichers mit s bedruckt und dann das gleiche Programm verwendet, wobei der links von s befindliche Teil des Speichers natürlich nicht benutzt wird. Umgekehrt kann man eine Maschine τ mit

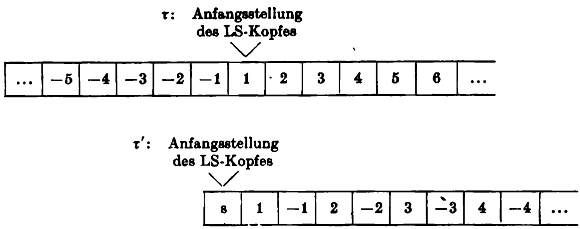


Abb. 2

beidseitig unbegrenztem Speicher auf einer Standard-Turingmaschine in folgender Weise simulieren: Die Zellen des unbeschränkten Speichers werden den Zellen des links beschränkten Speichers, wie in Abb. 2 dargestellt, zugeordnet. Jedem Zustand q des Programms τ entsprechen im allgemeinen sechs Zustände $q, q_L, q_R, \bar{q}, \bar{q}_L, \bar{q}_R$ des neuen Programms τ' für die Standard-Turingmaschine. Ist q_1 der Anfangszustand von τ, so sei q_1^* der Anfangszustand von τ', und es sei $\tau'(q_1^*, \mathrm{s}) := (q_1, \mathrm{s}, \mathrm{R})$, $\tau'(q_1^*, \mathrm{x})$ beliebig für $\mathrm{x} \neq \mathrm{s}$, d. h., im ersten Takt wird die simulierende Maschine τ' auf die richtige Anfangszelle (vgl. Abb. 2) eingestellt. Für jeden Befehl der Form $\tau(q, \mathrm{x}) = (q', \mathrm{y}, \mathrm{R})$ wird

$$\left.\begin{array}{ll} \tau'(q, \mathrm{x}) = (q_R', \mathrm{y}, \mathrm{R}), & \tau'(q_R', \mathrm{z}) = (q', \mathrm{z}, \mathrm{R}) \\ \tau'(\bar{q}, \mathrm{x}) = (\bar{q}_L', \mathrm{y}, \mathrm{L}), & \tau'(\bar{q}_L', \mathrm{z}) = (\bar{q}', \mathrm{z}, \mathrm{L}) \end{array}\right\} \text{für alle z des Bandalphabets,}$$

in τ' aufgenommen, d. h., wenn die Maschine sich im Zustand q auf einer Zelle mit positiver Nummer befindet und x liest, so druckt sie y, geht nach rechts und in einen Wartezustand q_R', in dem einerseits der im übernächsten Takt einzunehmende Zustand q', andererseits die Bewegungsrichtung R und schließlich das Arbeiten auf der rechten Hälfte des simulierten Bandes, d. h. auf den Zellen mit positiver Nummer, verschlüsselt sind. In diesen Zuständen der Form q_R werden demnach gerade die Zellen mit negativer Nummer passiert, ohne deren Inhalt zur Kenntnis zu nehmen oder zu ändern. Sinngemäß ist ein Rechtslauf der Maschine τ auf den Zellen mit negativen Nummern im allgemeinen durch einen Linkslauf der simulierenden Maschine τ' zu ersetzen, wobei in den Zwischenzuständen der Form $\bar{q}_L$ die Zellen mit positiver Nummer passiert werden, ohne deren Inhalt zur Kenntnis zu nehmen oder zu ändern. Es ist nun schon fast klar, wie ein Befehl der Form $\tau(q, \mathrm{x}) = (q', \mathrm{y}, \mathrm{L})$ in das Programm τ' zu übernehmen ist, nämlich durch

$$\left.\begin{array}{ll} \tau'(q, \mathrm{x}) = (q_L', \mathrm{y}, \mathrm{L}), & \tau'(q_L', \mathrm{z}) = (q', \mathrm{z}, \mathrm{L}) \\ \tau'(\bar{q}, \mathrm{x}) = (\bar{q}_R', \mathrm{y}, \mathrm{R}), & \tau'(\bar{q}_R', \mathrm{z}) = (\bar{q}', \mathrm{z}, \mathrm{R}) \end{array}\right\} \text{ für alle z des Bandalphabets.}$$

Einem Befehl der Form $\tau(q, \mathrm{x}) = (q', \mathrm{y}, \mathrm{N})$ entsprechen die Befehle

$$\tau'(q, \mathrm{x}) = (q', \mathrm{y}, \mathrm{N}), \quad \tau'(\bar{q}, \mathrm{x}) = (\bar{q}', \mathrm{y}, \mathrm{N}).$$

Es ist noch der Übergang von der positiv numerierten Bandseite auf die negativ numerierte und umgekehrt zu realisieren. Dies leisten offenbar die Befehle

$$\tau'(q_L, \mathrm{s}) = (\bar{q}_R, \mathrm{s}, \mathrm{R}), \quad \tau'(\bar{q}, \mathrm{s}) = (q, \mathrm{s}, \mathrm{R}).$$

Sind Programm und Anfangsinhalt des Speichers für eine Maschine mit unbegrenztem Band so beschaffen, daß der nichtleere Teil des Speichers in jedem Takt ein zusammenhängendes Wort bildet (dies kann man z. B. durch Benutzung eines Pseudo-Leerzeichens erreichen; vgl. den Beweis von Satz 1 in 7.4.), so kann man diese Maschine einfacher als durch den oben angegebenen Prozeß dadurch auf einer Standard-Turingmaschine simulieren, daß man jedesmal, wenn ein s gelesen wird, in ein Unterprogramm springt, durch welches der gesamte nichtleere Speicherinhalt um eine Zelle nach rechts verschoben wird, wonach der LS-Kopf auf die nun leere rechte Nachbarzelle von s (bzw. linke Nachbarzelle des nichtleeren Speicherinhalts) zurückkehrt. Mit anderen Worten: Wenn die Maschine am linken Ende von ihrem jeweiligen Zwischenresultat herunterrutscht bzw. dort Platz braucht, so schiebt sie das gesamte Zwischenresultat um jeweils eine Zelle nach rechts.

1.2. *Turingmaschinen im Quadrupel- und Tripel-Formalismus.* Hier wird die Arbeit einer Turingmaschine (unabhängig von der Speichergestalt) in noch kleinere Einzelschritte zerlegt. Im ursprünglichen Quadrupel-Formalismus haben die Befehle die Form $\tau(q, \mathrm{x}) = (q', \mathrm{Y})$, wobei Y ein Bandsymbol oder ein Steuerbefehl ist, d. h., in jedem Takt kann, in Abhängigkeit vom Zustand q und vom gelesenen Symbol x, nur entweder der Zelleninhalt geändert oder eine Bewegung des

LS-Kopfes ausgeführt werden. Zunächst ist klar, daß man jedes im Quadrupel-Formalismus geschriebene Programm in ein Takt für Takt gleichwertiges Programm im Quintupel-Formalismus umschreiben kann. Umgekehrt kann man zu jedem Programm im Quintupel-Formalismus ein äquivalentes Programm im Quadrupel-Formalismus angeben, das jeden Takt der Quintupel-Maschine in zwei Takten realisiert. Man ersetze z. B. den Befehl $\tau(q, \mathrm{x}) = (q', \mathrm{y}, \mathrm{R})$ durch

$$\tau'(q, \mathrm{x}) = (q_{\mathrm{R}}', \mathrm{y}), \quad \tau'(q_{\mathrm{R}}', \mathrm{y}) = (q', \mathrm{R}).$$

In analoger Weise sind auch die beiden anderen kombinatorisch möglichen Quadrupel-Formalismen (im ersten Fall wird in jedem Takt der Zellinhalt geändert, aber als zweites wahlweise ein neuer Zustand oder ein Steuerbefehl vorgeschrieben; im zweiten Fall wird in jedem Takt ein Steuerbefehl ausgeführt, aber als zweites nur entweder eine Zustandsänderung oder eine Änderung des Zellinhalts) zum klassischen Quintupel-Formalismus äquivalent. Es ist etwas überraschend, daß man diesen klassischen Formalismus sogar noch in einem Tripel-Formalismus simulieren kann, bei dem in jedem Takt nur eine der drei Komponenten Zustand, Zellinhalt, gelesene Zelle geändert werden darf.[1])

1.3. *Turingmaschinen mit mehreren Bändern oder mehreren LS-Köpfen auf einem Band.* Es sei $A_\varkappa$ das Alphabet der auf dem $\varkappa$-ten Band ($1 \leqq \varkappa \leqq k$) lesbaren und schreibbaren Zeichen. Eine *k-Band-Maschine* mit der Zustandsmenge Q und der Menge $Q_0 \subset Q$ von Stopzuständen wird dann durch eine Abbildung τ von $(Q \setminus Q_0) \times A_1 \times \cdots \times A_k$ in $Q \times A_1 \times \cdots \times A_k \times S^k$ beschrieben, wobei $S = \{\mathrm{L}, \mathrm{N}, \mathrm{R}\}$ ist. Das bedeutet, daß in Abhängigkeit vom jeweiligen Zustand q und den auf den k Bändern gelesenen Symbolen $\mathrm{x}_1, \ldots, \mathrm{x}_k$

a) der neue Zustand q',

b) für jedes der k Bänder der in die gelesene Zelle zu schreibende Buchstabe $\mathrm{x}_1', \ldots, \mathrm{x}_k'$,

c) für jedes der k Bänder die Bewegung des LS-Kopfes

festzulegen ist. Die Simulation einer solchen k-Band-Maschine auf einer 1-Band-Maschine kann nach ähnlichen Prinzipien erfolgen, wie sie in 1.1. benutzt wurden. Die Stellung der k LS-Köpfe in einem bestimmten Takt ist dabei durch k Hilfsbuchstaben zu markieren. Der eine LS-Kopf der simulierenden Maschine verändert nun nach den Anweisungen des simulierten Programms die Stellungen dieser k Pseudo-Köpfe und die Inhalte der entsprechenden Zellen auf den k Pseudo-Bändern, deren jedes durch jede k-te Zelle des einen Bandes dargestellt wird. Es ist klar, daß hier die Simulation eines einzigen Taktes der k-Band-Maschine ein umfangreiches Unterprogramm und sehr viele Schritte der simulierenden Maschine erfordert. Umgekehrt sind natürlich viele prinzipiell algorithmisch lösbare Aufgaben auf einer

[1]) Siehe hierzu P. C. Fisher, On Formalism for Turing Machines, J. Assoc. Comput. Machin. 12 (1965), 570–580.

Mehrband-Maschine viel leichter und schneller zu bewältigen als auf einer Einband-Maschine. Zum Beispiel erfordert die Addition zweier Dualzahlen auf zwei Bändern, wobei das Resultat auf dem zweiten Band erscheinen soll, nur zwei Zustände und so viele Takte, wie der Stellenzahl des längeren Summanden entsprechen:

1. Band 2. Band	0 0	0 1	1 0	1 1	o 0	o 1	0 o	1 o	o o
q_1	oL q_1 0L	oL q_1 1L	oL q_1 1L	oL q_2 0L	oN q_1 0L	oN q_1 1L	oL q_1 0L	oL q_1 1L	oN q_0 oN
q_2	oL q_1 1L	oL q_2 0L	oL q_2 0L	oL q_2 1L	oN q_1 1L	oN q_2 0L	oL q_1 1L	oL q_2 0L	oN q_0 1N

Dabei bedeutet o das Leerzeichen, 0 bzw. 1 die Dualziffern, q_1 Addition ohne Übertrag von der vorhergehenden Stelle und q_2 Addition mit Übertrag, und wir gehen der Einfachheit halber von der in Abb. 3 dargestellten Anfangskonstellation aus.

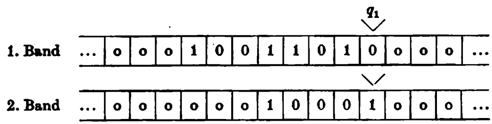


Abb. 3

Wie man sieht, ist die Tabelle des Turing-Programms in diesem Fall kaum etwas anderes als die vollständige Niederschrift des beim handschriftlichen Addieren befolgten Algorithmus. Es sei dem Leser empfohlen, die Addition zweier durch ein Zeichen * getrennter Dualzahlen auf einer Standard-Turingmaschine zu programmieren und Überlegungen zur hierbei benötigten Zustandszahl und Rechenzeit (Taktzahl) anzustellen. Fazit: Turingmaschinen mit zwei Bändern sind ein für die Addition von Dualzahlen „problemorientierter" Formalismus. Standard-Turingmaschinen lösen diese Aufgabe prinzipiell auch, sind aber im Vergleich extrem ungeeignet dafür.

Unter Benutzung zweier Bänder wird es insbesondere relativ leicht, eine universelle Turingmaschine für alle Standard-Turingmaschinen zu konstruieren. Dazu schreibe man das Programm der zu simulierenden Maschine auf das eine und das Eingabewort auf das andere Band. Unter Berufung auf die prinzipielle Simulierbarkeit jeder 2-Band-Maschine durch eine 1-Band-Maschine erhält man so die Existenz einer universellen Turingmaschine im ursprünglich geforderten Sinn.

1.4. *Turingmaschinen mit ebenem Speicher.* Hier wird der Rahmen der Wortalgorithmen in gewisser Weise überschritten, da als Ein- und Ausgabegrößen wie auch als Zwischenresultate kreuzworträtselartige Gebilde (Abb. 4) auftreten. An die Stelle der Steuerbefehle L, N, R treten Befehle L (linke Nachbarzelle), R (rechte Nachbarzelle), O (obere Nachbarzelle), U (untere Nachbarzelle), N (gleiche Zelle wie vorher). Im übrigen entspricht der Formalismus dem der klassischen Turingmaschinen mit Bandspeicher. Schon bei einem nur aus Leerzeichen und einem davon verschiedenen Symbol bestehenden Alphabet ergeben sich nun natürlich viele interes-

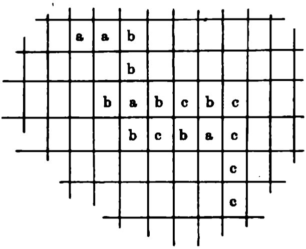


Abb. 4

sante Fragen hinsichtlich der Entscheidbarkeit gewisser „geometrischer“ Eigenschaften der Eingabegebilde und der „architektonischen“ Umgestaltung. Allerdings sind die „kreuzworträtselartigen Gebilde“ in naheliegender Weise eineindeutig durch Wörter kodierbar, und die Zuordnung zwischen Kodewörtern der Ein- und Ausgabegrößen, die durch eine solche ebene Turingmaschine ineinander übergeführt werden, erweist sich als Turing-berechenbar im Standardsinne. Demnach wird die Hypothese von Church durch die Existenz derartiger Maschinen nicht unmittelbar angetastet. Ebene Turingmaschinen sind jedoch in noch extremerer Weise als Mehrbandmaschinen für gewisse Aufgaben „problemorientiert“, deren Überführung in eine klassische Berechnungs- bzw. Entscheidungsaufgabe durch entsprechende Kodierung das ursprüngliche Problem unvorstellbar komplizieren würde. Zum Beispiel sind ebene Turingmaschinen problemorientiert für die Frage nach der Orientierungsfähigkeit von Automaten in Labyrinthen, da man solche Labyrinthe in einem ebenen Speicherfeld unmittelbar vorgeben kann.[1])

2. **Rekursive Funktionen.** Die Klasse der primitiv rekursiven Funktionen und danach die umfassendere Klasse der partiell rekursiven Funktionen werden induktiv als kleinste Klasse von arithmetischen Funktionen definiert, die gewisse einfache (im anschaulichen Sinne berechenbare) Funktionen enthält und bezüglich gewisser

[1]) Vgl. hierzu etwa K. Döpp, Automaten in Labyrinthen, EIK 7 (1971), 79—94, 167—190, sowie P. Schreiber, Theseus im Labyrinth als Turingmaschine, Zeitschr. f. Math. Logik u. Grdlg. d. Math. 17 (1971), 57—60.

Erzeugungsprinzipien (die von im anschaulichen Sinne berechenbaren Funktionen zu im gleichen Maße berechenbaren Funktionen führen) abgeschlossen ist:

Definition 1.

a) Die *Nullfunktion* f^0, definiert durch $f^0(n) := 0$, ist primitiv rekursiv.

b) Die *Nachfolgerfunktion* f^1, definiert durch $f^1(n) := n + 1$, ist primitiv rekursiv.

c) Die *Projektionsfunktionen* $f_m{}^i$ $(m \geqq 1, 1 \leqq i \leqq m)$, definiert durch die Beziehung $f_m{}^i(n_1, \ldots, n_m) := n_i$, sind primitiv rekursiv.

d) Sind die m-stellige Funktion g und die m k-stelligen Funktionen f_i $(1 \leqq i \leqq m)$ primitiv rekursiv, so ist auch die aus ihnen durch *Superposition* gewonnene Funktion h, definiert durch $h(n_1, \ldots, n_k) := g(f_1(n_1, \ldots, n_k), \ldots, f_m(n_1, \ldots, n_k))$, primitiv rekursiv.

e) Sind die m-stellige Funktion g_0 und die $(m + 2)$-stellige Funktion g_1 primitiv rekursiv, so ist auch die aus ihnen durch *primitive Rekursion* gewonnene, durch

$$\begin{aligned} h(n_1, \ldots, n_m, 0) &:= g_0(n_1, \ldots, n_m), \\ h(n_1, \ldots, n_m, n + 1) &:= g_1(n_1, \ldots, n_m, n, h(n_1, \ldots, n_m, n)) \end{aligned} \tag{1}$$

definierte $(m + 1)$-stellige Funktion h primitiv rekursiv.

Die Klasse der *primitiv rekursiven Funktionen* ist demnach die kleinste Funktionenmenge, die die Nullfunktion, die Nachfolgerfunktion und alle Projektionsfunktionen enthält und bezüglich Superposition und primitiver Rekursion abgeschlossen ist.

Beispiel. Wir wollen zeigen, daß die durch $f_+(m, n) := m + n$ definierte zweistellige Funktion f_+ primitiv rekursiv ist. Dazu setzen wir im Schema (1) der primitiven Rekursion $m = 1$, wählen für g_0 die primitiv rekursive Funktion $f_1{}^1$ und für g_1 die durch Superposition aus den primitiv rekursiven Funktionen f^1 und $f_3{}^3$ entstehende und also auch primitiv rekursive Funktion

$$g_1(m, n, k) := f^1(f_3{}^3(m, n, k)) = k + 1.$$

Die bekannte rekursive Definition der Addition

$$m + 0 := m, \quad m + (n + 1) := (m + n) + 1$$

läßt sich nun offenbar in der Form (1) als

$$f_+(m, 0) := g_0(m), \quad f_+(m, n + 1) := g_1(m, n, f_+(m, n))$$

schreiben. Folglich ist f_+ primitiv rekursiv.

Definition 2.

a) Die Nullfunktion, die Nachfolgerfunktion und alle Projektionsfunktionen sind partiell rekursiv.

b) Die Klasse der partiell rekursiven Funktionen sei bezüglich Superposition und primitiver Rekursion abgeschlossen.

c) Ist die $(m + 1)$-stellige Funktion g $(m \geqq 1)$ partiell rekursiv, so ist die m-stellige, durch

$$h(n_1, \ldots, n_m) := \begin{cases} \textit{kleinste Zahl } n \textit{ mit } g(n_1, \ldots, n_m, n) = 0, \textit{ falls eine solche} \\ \textit{existiert und } g(n_1, \ldots, n_m, k) \textit{ für } k < n \textit{ stets definiert ist,} \\ \textit{nicht definiert andernfalls} \end{cases}$$

definierte Funktion h partiell rekursiv.

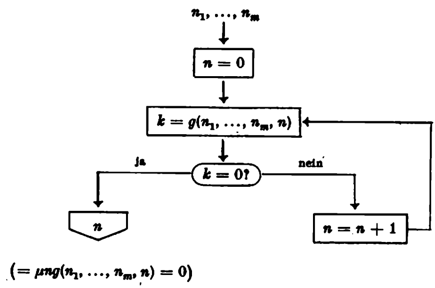
Für das in c) benutzte Definitionsprinzip schreibt man kurz

$$h(n_1, \ldots, n_m) := \mu n[g(n_1, \ldots, n_m, n) = 0]$$

(gelesen: das kleinste n, für das $g(n_1, \ldots, n_m, n) = 0$ ist) und bezeichnet die Operation, die auf diese Weise jeder $(m + 1)$-stelligen arithmetischen Funktion eine m-stellige arithmetische Funktion zuordnet, als *μ-Operator*. Mit dieser Bezeichnung kann man Definition 2 auch so aussprechen: Die Klasse der *partiell rekursiven Funktionen* ist die kleinste Funktionenmenge, die die Nullfunktion, die Nachfolgerfunktion und alle Projektionsfunktionen enthält und bezüglich Superposition, primitiver Rekursion und μ-Operator abgeschlossen ist. Aus den Definitionen folgt sofort, daß jede primitiv rekursive Funktion erst recht partiell rekursiv ist. Es ist leicht zu sehen, daß die Klasse der partiell rekursiven Funktionen die der primitiv rekursiven Funktionen echt umfaßt: Alle primitiv rekursiven Funktionen sind volle, d. h. überall definierte Funktionen. Jedoch führt der μ-Operator aus der Klasse der vollen Funktionen heraus. So ergibt seine Anwendung auf die primitiv rekursive Funktion f_+ die Funktion

$$h(m) := \mu n\,[m + n = 0] = \begin{cases} 0, \textit{ falls } m = 0, \\ \textit{nicht definiert für } m > 0, \end{cases}$$

die folglich partiell rekursiv, aber nicht primitiv rekursiv ist. Die Berechtigung, den μ-Operator als Erzeugungsprinzip für berechenbare Funktionen zu benutzen, ergibt sich aus dem folgenden Flußdiagramm, das sicher einen im anschaulichen Sinne effektiven Berechnungsprozeß darstellt, falls die Funktion g berechenbar ist:



Definition 3. Eine partiell rekursive Funktion heißt *allgemein rekursiv*, wenn sie überall definiert ist.

Aus dieser Definition folgt sofort, daß jede primitiv rekursive Funktion allgemein rekursiv ist, und daß die allgemein rekursiven Funktionen eine echte Teilmenge der Klasse der partiell rekursiven Funktionen bilden. Man kann zeigen, daß es allgemein rekursive Funktionen gibt, die nicht primitiv rekursiv sind. Dies ist jedoch natürlich nicht ganz so leicht wie der Nachweis, daß es partiell rekursive Funktionen gibt, die nicht allgemein rekursiv sind. Merkwürdigerweise wurde die Hypothese von CHURCH, die sich, wie bereits in 7.3. bemerkt, ursprünglich auf die Klasse der primitiv rekursiven Funktionen bezog, zunächst auf Grund der Entdeckung von allgemein, jedoch nicht primitiv rekursiven Funktionen berichtigt. Die uns heute viel näher liegende Überlegung, daß die Klasse aller im anschaulichen Sinne berechenbaren Funktionen selbstverständlich auch viele partielle Funktionen enthalten muß und daß die Klasse der primitiv rekursiven Funktionen schon deshalb nicht alle im anschaulichen Sinne berechenbaren Funktionen umfassen kann, drang erst relativ spät ins Bewußtsein der Mathematiker.

Die (auf arithmetische Funktionen bezogene) heutige Fassung der Hypothese von CHURCH, daß jede im anschaulichen Sinne berechenbare Funktion partiell rekursiv ist, wirft sofort für eine Unzahl von Erzeugungsprinzipien für im anschaulichen Sinne berechenbare Funktionen (z. B. viele weitere Formen der Substitution und Rekursion) die Frage auf, ob sie denn nicht aus der Klasse der partiell rekursiven Funktionen herausführen. Tatsächlich erscheinen ja die zur Definition dieser Klasse benutzten Grundfunktionen und Erzeugungsprinzipien auf den ersten Blick sehr „schwach". Hierzu ist zu sagen, daß die heutige „elegante" Definition das Resultat jahrzehntelanger mühsamer Untersuchungen ist, in deren Verlauf die ursprünglich zur Definition benutzten Prinzipien reduziert und die Entbehrlichkeit vieler weiterer Erzeugungsprinzipien nachgewiesen wurden (vgl. hierzu insbesondere [31]). Die folgenden Sätze erschließen den tatsächlichen Umfang bzw. Reichtum der Klasse der partiell rekursiven Funktionen auf einem anderen Weg.

Satz 1. *Eine m-stellige* ($m \geqq 1$) *arithmetische Funktion* f *ist genau dann partiell rekursiv, wenn die durch*

$$\bar{f}(\underbrace{a \dots a}_{n_1}b\underbrace{a \dots a}_{n_2}b \dots b\underbrace{a \dots a}_{n_m}) := \underbrace{a \dots a}_{f(n_1, \dots, n_m)}$$

definierte partielle Wortfunktion $\bar{f}$ *in* $W(\{a, b\})$ *Turing-berechenbar ist.*

Wir skizzieren den Beweis dieses Satzes: a) Der Beweis, daß für alle partiell rekursiven Funktionen f die zugehörige Wortfunktion $\bar{f}$ Turing-berechenbar ist, wird durch Induktion bezüglich Definition 2 geführt. Es ist also zu zeigen: $\overline{f^0}, \overline{f^1}, \overline{f_m{}^i}$ sind Turing-berechenbar. (Entsprechende Turingprogramme kann der Leser selbst mühelos konstruieren.) Mit Hilfe der Unterprogrammtechnik für Turingmaschinen ist ferner zu zeigen: Entsteht h durch Superposition aus g und $f_1, \ldots, f_m$ und sind $\bar{g}, \bar{f}_1, \ldots, \bar{f}_m$ Turing-berechenbar, so ist auch $\bar{h}$ Turing-berechnebar. Entsteht h durch primitive Rekursion aus g_0 und g_1 (bzw. durch Anwendung des μ-Operators auf g) und sind $\bar{g}_0, \bar{g}_1$ (bzw. $\bar{g}$) Turing-berechnebar, so ist auch $\bar{h}$ Turing-berechenbar.

b) Der Beweis, daß f partiell rekursiv ist, wenn $\bar{f}$ Turing-berechenbar ist, erfordert den Nachweis, daß gewisse spezielle Funktionen partiell rekursiv sind. Setzt man dies als bekannt voraus, so läßt sich der Beweisweg leicht skizzieren. Die Wortfunktion $\bar{f}$ werde durch das Turingprogramm τ berechnet. Man wählt eine eineindeutige Gödelisierung g von der Menge aller Konstellationen K bezüglich τ auf die Menge der natürlichen Zahlen und zeigt, daß folgende Funktionen partiell (und wie sich zeigt, sogar primitiv) rekursiv sind:

$$h_1(n_1, \ldots, n_m) := g\Bigl(\mathrm{s}\overset{q_1}{\,}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{n_1}\mathrm{b}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{n_2}\mathrm{b} \ldots \mathrm{b}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{n_m}\Bigr),$$

$$h_2(m) := \begin{cases} n, \textit{ falls } m \textit{ Gödelzahl einer Endkonstellation der Form} \\ \quad \mathrm{s}\overset{q_0}{\,}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{n} \textit{ bzw. } \mathrm{s}\overset{q_0}{\,}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{n}\mathrm{x} \ldots (\mathrm{x} \neq \mathrm{a}), \\ 0 \textit{ sonst}, \end{cases}$$

$$h_3(m) := \begin{cases} 0, \textit{ falls } m \textit{ Gödelzahl einer Endkonstellation}, \\ 1 \textit{ sonst}, \end{cases}$$

$$h_4\bigl(g(K)\bigr) := g\bigl(\tau(K)\bigr) \quad (K \textit{ beliebige Konstellation}).$$

(Damit h_4 überall definiert ist, muß $\tau(K)$ für jede Konstellation K definiert sein. Man setzt zweckmäßigerweise $\tau(K) = K$, falls K Endkonstellation ist.) Mit h_4 ist auch

$$h_5\bigl(g(K), n\bigr) := g\bigl(\tau^n(K)\bigr) \qquad \bigl(\tau^0(K) := K\bigr)$$

primitiv rekursiv, da h_5 den Rekursionsgleichungen

$$h_5(m, 0) = m, \qquad h_5(m, n+1) = h_4\bigl(h_5(m, n)\bigr)$$

genügt. Mit h_3 und h_5 ist als Superposition auch

$$h_6(m, n) := h_3\bigl(h_5(m, n)\bigr)$$

primitiv rekursiv, und es ist offenbar

$$h_6\bigl(g(K), n\bigr) = \begin{cases} 0, \textit{ falls } \tau^n(K) \textit{ Endkonstellation}, \\ 1 \textit{ sonst}. \end{cases}$$

Daher ist

$$h_7(m) := \mu n h_6(m, n) = 0$$

partiell rekursiv. Die Funktion h_7 ordnet $g(K)$ die Nummer desjenigen Taktes zu, in dem die Maschine τ bei Anwendung auf die Konstellation K stoppt, falls ein solcher Takt existiert. Die Funktion f läßt sich nun offenbar durch

$$f(n_1, \ldots, n_m) = h_2(h_5(h_1(n_1, \ldots, n_m), h_7(h_1(n_1, \ldots, n_m))))$$

darstellen, d. h. als Superposition partiell rekursiver Funktionen. Folglich ist f selbst partiell rekursiv. Zugleich haben wir eine wichtige Teilaussage des Darstellungssatzes von KLEENE bewiesen, nämlich:

Satz 2. *Jede partiell rekursive Funktion läßt sich aus primitiv rekursiven Funktionen durch Superposition und höchstens einmalige Anwendung des μ-Operators erzeugen.*

Den Zusammenhang zwischen der Turing-Berechenbarkeit von Wortfunktionen und der partiellen Rekursivität von arithmetischen Funktionen kann man auch noch auf eine andere Weise formulieren:

Satz 3. *Es sei g eine Gödelisierung der Wortmenge $W(A)$, so daß g und g^{-1} Turing-berechenbar sind (d. h., für beliebiges $\mathrm{a} \in A$ sind die Wortfunktionen $f_1(\mathrm{W}) := \underbrace{\mathrm{a} \ldots \mathrm{a}}_{g(\mathrm{W})}$ und $f_2 = f_1^{-1}$ Turing-berechenbar). Dann ist eine beliebige Wortfunktion f in $W(A)$ genau dann Turing-berechenbar, wenn die durch*

$$\varphi_f(g(\mathrm{W})) := g(f(\mathrm{W}))$$

definierte arithmetische Funktion φ_f partiell rekursiv ist.

Beweis. Ist f Turing-berechenbar, so ist zum Beweis, daß φ_f partiell rekursiv ist, nach Satz 1 nur zu zeigen: Die durch

$$f_3(\underbrace{\mathrm{a} \ldots \mathrm{a}}_{g(\mathrm{W})}) := \underbrace{\mathrm{a} \ldots \mathrm{a}}_{g(f(\mathrm{W}))}$$

definierte Wortfunktion f_3 ist Turing-berechenbar. Da $f_3 = f_1 \circ f \circ f_2$ ist und f, f_1, f_2 Turing-berechenbar sind, folgt dies aus 7.4., Satz 2. Ist umgekehrt φ_f partiell rekursiv, so ist nach Satz 1 die Wortfunktion f_3 Turing-berechenbar. Damit ist aber auch $f = f_2 \circ f_3 \circ f_1$ Turing-berechenbar.

3. Normale Algorithmen. Die normalen Algorithmen wurden 1951 von A. A. MARKOV als Präzisierung des Algorithmenbegriffs vorgeschlagen. Sie stellten den ersten unmittelbar auf Wortfunktionen zugeschnittenen Algorithmenbegriff dar. (Turingmaschinen wurden ursprünglich nur zur Berechnung arithmetischer Funktionen genutzt.) Analog zu den Turingmaschinen existieren heute zum Begriff der normalen Algorithmen mehrere äquivalente Varianten.

Als *Markov-Substitution* (*Mark*) bezeichnen wir die wie folgt definierte dreistellige volle und im anschaulichen Sinne gewiß effektiv ausführbare Wortoperation:

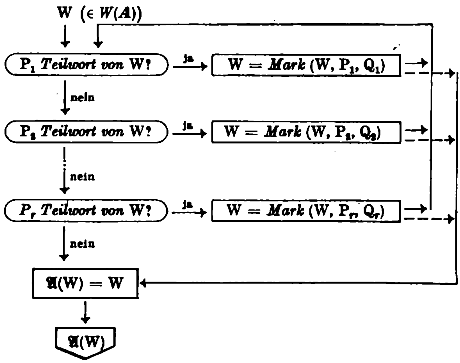
$$Mark(\mathrm{W}, \mathrm{P}, \mathrm{Q}) := \begin{cases} \mathrm{W}, & \textit{falls } \mathrm{P} \textit{ nicht Teilwort von } \mathrm{W} \textit{ ist,} \\ \mathrm{UQR}, & \textit{falls } \mathrm{W} = \mathrm{UPR} \textit{ und } \mathrm{P} \textit{ in keinem} \\ & \textit{echten Anfangsstück von } \mathrm{UP} \textit{ vorkommt.} \end{cases}$$

Demnach wird, falls P überhaupt in W vorkommt, das am weitesten links gelegene Vorkommen von P in W durch Q ersetzt. Insbesondere ist $Mark(\mathrm{W}, \Lambda, \mathrm{Q}) = \mathrm{QW}$, da das leere Wort stets in W am Anfang als Teilwort vorkommt, und $Mark(\mathrm{W}, \mathrm{P}, \Lambda)$ das Resultat des Streichens des ersten Vorkommens von P in W.

Es sei $\boldsymbol{A}$ ein endliches Alphabet, das die Zeichen $\rightarrow$ und $\bullet$ nicht enthält. Eine *nichtabbrechende* (Markovsche) *Regel* ist ein Wort der Form $\mathrm{P} \rightarrow \mathrm{Q}$ mit $\mathrm{P}, \mathrm{Q} \in W(\boldsymbol{A})$. Eine *abbrechende Regel* ist ein Wort der Form $\mathrm{P} \rightarrow \bullet\, \mathrm{Q}$ mit $\mathrm{P}, \mathrm{Q} \in W(\boldsymbol{A})$. In beiden Fällen wird P als *linke Seite* und Q als *rechte Seite* der betreffenden Regel bezeichnet. Ein *normaler Algorithmus im Alphabet* $\boldsymbol{A}$ ist eine endliche geordnete Liste von (abbrechenden oder nichtabbrechenden) Regeln, allgemein in der Form

$$\mathfrak{A} = \begin{cases} \mathrm{P}_1 \rightarrow (\bullet)\, \mathrm{Q}_1 \\ \mathrm{P}_2 \rightarrow (\bullet)\, \mathrm{Q}_2 \\ \vdots \\ \mathrm{P}_r \rightarrow (\bullet)\, \mathrm{Q}_r \end{cases}$$

geschrieben. Dabei deutet die Schreibweise $(\bullet)$ an, daß im konkreten Algorithmus jede Regel unabhängig von den anderen abbrechend oder nichtabbrechend sein kann. Die von MARKOV vorgeschriebene Art der Anwendung eines solchen Algorithmus läßt sich exakt durch folgendes Flußdiagramm beschreiben:



Dabei gilt im konkreten Fall jeweils der durchgezogene Pfeil, falls die entsprechende Regel nichtabbrechend ist, und der gestrichelte Pfeil, falls sie abbrechend ist. In Worten kann man demnach die Anwendung eines normalen Algorithmus auf ein Wort etwa wie folgt beschreiben:

Eine Regel ist anwendbar auf ein Wort W, wenn ihre linke Seite in W vorkommt. Die Regeln werden in der gegebenen Reihenfolge auf ihre Anwendbarkeit geprüft. Falls keine Regel anwendbar ist, ist $\mathfrak{A}(\mathrm{W}) = \mathrm{W}$. Andernfalls wird die erste anwendbare Regel angewendet, indem das erste Vorkommen ihrer linken Seite in W durch die entsprechende rechte Seite ersetzt wird. War die angewendete Regel abbrechend, so ist das erhaltene Wort $\mathfrak{A}(\mathrm{W})$, andernfalls beginnt man von vorn mit der Prüfung der Anwendbarkeit der ersten Regel.

Beispiel. Es sei $\boldsymbol{A} = \{\mathrm{a}, \mathrm{b}\}$ und $\mathfrak{A} = \begin{cases} \mathrm{bab} \to \mathrm{ba}, \\ \mathrm{ba} \to \bullet\, \mathrm{a}, \\ \Lambda \to \mathrm{ba}. \end{cases}$

Anwendung auf W = b liefert der Reihe nach folgende Zwischenresultate

b (1. und 2. Regel nicht anwendbar, 3. Regel anwendbar),
bab (1. Regel anwendbar),
ba (1. Regel nicht anwendbar, 2. Regel anwendbar),
a ($= \mathfrak{A}(\mathrm{b})$, da die zuletzt angewendete Regel abbrechend ist).

Anwendung auf a ergibt

a (1. und 2. Regel nicht anwendbar, 3. Regel anwendbar),
baa (1. Regel nicht anwendbar, 2. Regel anwendbar),
aa ($= \mathfrak{A}(\mathrm{a})$, da die letzte Regel abbrechend ist).

Es zeigt sich, daß man nicht jede Turing-berechenbare Wortfunktion in $W(\boldsymbol{A})$ durch einen normalen Algorithmus *in* $W(\boldsymbol{A})$ berechnen kann. Dazu werden im allgemeinen sogenannte *Hilfsbuchstaben* benötigt, die in den Regeln und daher auch in den Zwischenresultaten, jedoch nicht im Endresultat vorkommen dürfen. Ein normaler Algorithmus im Alphabet $\boldsymbol{A} \cup \boldsymbol{H}$ ($\boldsymbol{H}$ Alphabet der Hilfsbuchstaben) heißt ein *Algorithmus über* $\boldsymbol{A}$, wenn für $\mathrm{W} \in W(\boldsymbol{A})$ stets $\mathfrak{A}(\mathrm{W}) \in W(\boldsymbol{A})$ ist (falls $\mathfrak{A}(\mathrm{W})$ existiert). Als Beispiel hierzu betrachten wir die Berechnung der Wortfunktion

$$f(\mathrm{W}) := \mathrm{Wa} \qquad \big(\mathrm{W} \in W(\{\mathrm{a}, \mathrm{b}\})\big)$$

durch einen normalen Algorithmus $\mathfrak{B}$ über $W(\{\mathrm{a}, \mathrm{b}\})$ (in $W(\{\mathrm{a}, \mathrm{b}, *\})$. Die bei der Lösung dieser Aufgabe zu überwindende Schwierigkeit besteht darin, daß nicht das erste, sondern das letzte Vorkommen von Λ in W durch a ersetzt werden soll. ($f'(\mathrm{W}) = \mathrm{aW}$ wäre sehr leicht durch den Algorithmus $\mathfrak{B}' = \{\Lambda \to \bullet\, \mathrm{a}\}$ zu realisieren.) Der für die normalen Algorithmen fast typische Kunstgriff besteht in der Einführung eines Hilfsbuchstaben $*$ vorn, der dann das Wort durchläuft und an dessen Ende durch a ersetzt wird:

$$\mathfrak{B} = \begin{cases} * \mathrm{a} \to \mathrm{a} * \\ * \mathrm{b} \to \mathrm{b} * \\ * \to \bullet \mathrm{a} \\ \Lambda \to * \end{cases}$$

Bei Anwendung von $\mathfrak{B}$ auf baab erhält man der Reihe nach die Zwischenresultate

$$\begin{array}{l} \mathrm{b\ a\ a\ b} \\ * \mathrm{\ b\ a\ a\ b} \\ \mathrm{b} * \mathrm{a\ a\ b} \\ \mathrm{b\ a} * \mathrm{a\ b} \\ \mathrm{b\ a\ a} * \mathrm{b} \\ \mathrm{b\ a\ a\ b} * \\ \mathrm{b\ a\ a\ b\ a} \end{array}$$

Man beachte, daß die dritte Regel abbrechend sein muß, da andernfalls immer wieder die letzte Regel angewendet werden könnte!

Markov arbeitete die Theorie der normalen Algorithmen mit vielen Sätzen und mit Methoden zur Unterprogrammtechnik bis zur expliziten Konstruktion eines *universellen* normalen Algorithmus $\mathfrak{U}$ aus, der bei Anwendung auf die geeignet kodierte Niederschrift eines Paares $(\mathfrak{A}, \mathrm{W})$ das Resultat $\mathfrak{A}(\mathrm{W})$ liefert, falls dieses existiert. Wir zeigen nun

Satz 4. *Eine Wortfunktion in $W(A)$ ist genau dann Turing-berechenbar, wenn sie sich durch einen normalen Algorithmus über $W(A)$ realisieren läßt.*[1])

Beweis. a) Zum Beweis, daß jede durch einen normalen Algorithmus realisierte Wortfunktion Turing-berechenbar ist, genügt es auf Grund der Unterprogrammtechnik für Turingmaschinen, zu jedem Paar P, Q von Wörtern ein Programm τ mit zwei Stopzuständen q_0^+ und q_0^- anzugeben, so daß $f_\tau(\mathrm{W}) = Mark\ (\mathrm{W}, \mathrm{P}, \mathrm{Q})$ ist und dabei τ in q_0^+ stoppt, falls P in W vorkommt, bzw. in q_0^-, falls P nicht in W vorkommt. Wir geben ein solches Programm für den Fall an, daß das Alphabet des Algorithmus gleich $\{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$ und $\mathrm{P} = \mathrm{ba}$, $\mathrm{Q} = \mathrm{ccb}$ ist. Man erkennt daraus, wie im allgemeinen vorzugehen ist.

---

[1]) Dabei kann man immer mit einem einzigen Hilfsbuchstaben auskommen. Vgl. H. M. Нагорный, О минимальном алфавите алгорифмов над данным алфавитом, Тр. матем. ин.-та им. Стеклова АН СССР 52 (1958), 66–74, und P. Schreiber, Über die Entbehrlichkeit von Hilfsbuchstaben bei der Berechnung mehrstelliger Wortfunktionen durch Markowsche Algorithmen, Zeitschr. f. Math. Logik u. Grdlg. d. Math. 12 (1966), 241–242.

	s	o	a	b	c	ba → ccb
q_1	R	q_5 L	R	q_2 R	R	b wird gesucht
q_2	—	q_5 L	q_3 c L	q_2 R	q_1 R	a (nach b) wird gesucht
q_3	—	—	—	c R	q_b R	ba ist gefunden und wird durch cc ersetzt
q_a	—	q_4 a L	R	q_b a R	q_c a R	b wird nach cc eingefügt und der Rest des Wortes um eine Zelle nach rechts verschoben
q_b	—	q_4 b L	q_a b R	R	q_c b R	
q_c	—	q_4 c L	q_a c R	q_b c R	R	
q_4	q_0^+ N	—	L	L	L	Substitution beendet, LS-Kopf kehrt nach s zurück
q_5	q_0^- N	—	L	L	L	ba kommt nicht im Eingabewort vor, LS-Kopf kehrt zu s zurück

b) Jede Turing-berechenbare Wortfunktion ist durch einen normalen Algorithmus realisierbar. Dazu simulieren wir die Arbeit einer Turingmaschine τ, indem wir die Zustände von τ als Hilfsbuchstaben benutzen, die jeweils links neben dem gerade gelesenen Buchstaben eingefügt werden. Die Zwischenresultate eines so konstruierten normalen Algorithmus entsprechen im wesentlichen den Konstellationen derjenigen Turingmaschine, von der wir voraussetzen, daß $\boldsymbol{A}$ ihr Ein- und Ausgabealphabet, $\boldsymbol{H} = \{\mathrm{s}, \mathrm{o}, \ldots\}$ das Alphabet ihrer Hilfsbuchstaben, q_1 ihr Anfangszustand und Q_0 die Menge ihrer Stopzustände ist. Der simulierende normale Algorithmus hat dann folgende Gestalt:

$$\left\{\begin{array}{llll}
(1) & \mathrm{zqx} \to \mathrm{q'zy} & (\mathrm{z} \in \boldsymbol{A} \cup \boldsymbol{H}) & \text{für } \tau(\mathrm{q}, \mathrm{x}) = (\mathrm{q'}, \mathrm{y}, \mathrm{L}) \\
(2) & \mathrm{qx} \to \mathrm{q'y} & & \text{für } \tau(\mathrm{q}, \mathrm{x}) = (\mathrm{q'}, \mathrm{y}, \mathrm{N}) \\
(3) & \mathrm{qx} \to \mathrm{yq'} & & \text{für } \tau(\mathrm{q}, \mathrm{x}) = (\mathrm{q'}, \mathrm{y}, \mathrm{R}) \\
(4) & \mathrm{q} \to \mathrm{yq'} & & \text{für } \tau(\mathrm{q}, \mathrm{o}) = (\mathrm{q'}, \mathrm{y}, \mathrm{R}) \\
(5) & \mathrm{q} \to \mathrm{q'y} & & \text{für } \tau(\mathrm{q}, \mathrm{o}) = (\mathrm{q'}, \mathrm{y}, \mathrm{N}) \\
(6) & \mathrm{zq} \to \mathrm{q'zy} & & \text{für } \tau(\mathrm{q}, \mathrm{o}) = (\mathrm{q'}, \mathrm{y}, \mathrm{L}) \\
(7) & \mathrm{q_0s} \to * & & \text{für } \mathrm{q_0} \in Q_0 \\
(8) & * \mathrm{x} \to \mathrm{x} * & (\mathrm{x} \in \boldsymbol{A}) & \\
(9) & * \to + & & \\
(10) & + \mathrm{x} \to + & (\mathrm{x} \in \boldsymbol{A} \cup \boldsymbol{H}, \mathrm{x} \neq \mathrm{s}) & \\
(11) & + \to \bullet \Lambda & & \\
(12) & \Lambda \to \mathrm{q_1s} & &
\end{array}\right.$$

Die angegebenen Regeln sind bis auf (7), (9), (11) und (12) endliche Schemata von Regeln, die im konkreten Fall jeweils für alle dahinter angegebenen Fälle aufzunehmen sind. Die Reihenfolge der Regeln, die im allgemeinen in normalen Algorithmen wesentlich ist, ist hier, von den folgenden Ausnahmen abgesehen, beliebig, da im allgemeinen auf die Zwischenresultate genau eine der Regeln (außer (12)) anwendbar ist. Die Ausnahmen sind folgende: (12) muß letzte Regel sein, da sie immer anwendbar ist und daher alle unter ihr stehenden Regeln blockieren würde. Die Regeln der Form (4), (5), (6) müssen unterhalb der jeweiligen Regeln der Form (1), (2), (3) stehen, also z. B. $q \to q'y$ nach $qo \to q'y$ und $zq \to q'zy$ nach $zqo \to q'zy$ usw. Die Regeln der Form (4), (5), (6) sollen ja nur dann angewendet werden, wenn der LS-Kopf der simulierten Turingmaschine über der ersten leeren Zelle hinter der momentanen Konstellation steht. Die linken Seiten der zugehörigen Markovschen Regeln kommen jedoch als Teilwort auch in denjenigen Zwischenresultaten vor, auf die die entsprechende Regel der Form (1), (2) oder (3) angewendet werden sollte. Schließlich muß (9) hinter (8) und (11) hinter (10) stehen. Die Regeln (7) bis (11) dienen, wie man leicht einsieht, dem Zweck, die Endkonstellation in das ihr entsprechende Resultatwort zu verwandeln, d. h. den Anfang $q_0 s$ und alle eventuell auf das dahinter stehende Resultatwort folgenden Leerzeichen und sonstigen Buchstaben zu löschen. Falls die simulierte Turingmaschine sauber ist, kann man (9), (10), (11) durch

$$* o \to *$$

$$* \to \bullet \Lambda$$

ersetzen. Die außer den Zuständen $q \in Q$ und Buchstaben $h \in \boldsymbol{H}$ vom normalen Algorithmus benutzten Hilfsbuchstaben $*$ und $+$ dürfen natürlich in $\boldsymbol{A} \cup \boldsymbol{H} \cup Q$ nicht vorkommen.

8. Entscheidbarkeit und Aufzählbarkeit von Sprachen und Theorien

8.1. Entscheidbarkeit formalisierter Sprachen

Damit man die Frage nach der Entscheidbarkeit einer formalisierten Sprache S exakt (etwa im Sinne der Turing-Entscheidbarkeit) stellen kann, muß S Teilmenge einer Menge $W(A)$ mit endlichem Alphabet A sein. Dies ist bei der ursprünglichen Auffassung weder für kanonische noch für modifizierte Sprachen der Fall, da schon die Menge aller Variablen einer Sprache stets abzählbar ist. Man kann diese Forderung jedoch durch naheliegende und geringfügige Modifizierung (Umkodierung) der Sprachgrundbestandteile erfüllen. Im Fall n-sortiger Sprachen wähle man etwa für die Variablensorten n paarweise verschiedene Grundbuchstaben $\mathrm{x}, \mathrm{y}, \ldots, \mathrm{z}$ und einen von diesen verschiedenen Buchstaben a zur Indexkodierung und schreibe die Variablen $\mathrm{x}_i, \mathrm{y}_j, \ldots$ in der Form $\mathrm{x}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{i}$, $\mathrm{y}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{j}$ $(i, j \geqq 0)$. Im Fall abzählbar vieler Variablensorten ist in analoger Weise auch die Sortennummer zu kodieren, d. h., man schreibt etwa $\mathrm{x}\underbrace{\mathrm{b} \ldots \mathrm{b}}_{j}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{i}$ statt $\mathrm{x}_i{}^j$. Werden nur endlich viele verschiedene Relationssymbole R_i benötigt, so kann man diese für den jetzigen Zweck als paarweise verschiedene Grundbuchstaben ansehen. Andernfalls sind die Nummer des Relationssymbols und eventuell auch dessen Typ in analoger Weise wie die Indizes der Variablen zu kodieren. Gleiches gilt für Operations- und Konstantensymbole.

Wir behaupten nun, daß jede kanonische und jede „vernünftig“ definierte modifizierte formalisierte Sprache $S \subseteqq W(A)$, wobei A ein endliches Alphabet ist, eine entscheidbare Teilmenge von $\mathrm{W}(A)$ ist, d. h., daß es ein auf alle Wörter $\mathrm{W} \in W(A)$ anwendbares algorithmisches Verfahren gibt, das für $\mathrm{W} \in S$ eine positive und für $\mathrm{W} \in W(A) \setminus S$ eine negative Antwort gibt. Bei der Angabe eines solchen Verfahrens werden wir uns, wie in solchen Fällen üblich, auf die inhaltliche Beschreibung beschränken und es den mißtrauischen Lesern überlassen, diese Beschreibung in ein adäquates Turingprogramm oder einen Markovschen Algorithmus überzuführen. Es sei jedoch auch bemerkt, daß das Problem der Entscheidung, ob eine beliebige Zeichenreihe $W \in W(A)$ ein Ausdruck einer Sprache $S \subseteq W(A)$ ist, nicht ganz so trivial ist, wie es manchem Leser auf den ersten Blick erscheinen mag, da er gewohnt ist, an relativ „überschaubare" Zeichenreihen zu denken. Ein allgemeines Verfahren muß aber prinzipiell auf beliebig lange Wörter anwendbar sein.

Von den inhaltlich verschiedenen Möglichkeiten bzw. Wegen, die Entscheidbarkeit der formalisierten Sprachen einzusehen, behandeln wir eine Methode, die man als „Einschmelzung" bezeichnen könnte, da sie im wesentlichen darin besteht, zunächst von kürzestmöglichen und dann von immer längeren Teilwörtern des gegebenen Wortes W auf Grund der induktiven Definition von S festzustellen, daß sie Terme bzw. Ausdrücke von S sind, wobei die bereits erkannten Teilstücke durch solche nicht selbst zu S gehörige Wörter möglichst einfacher Bauart ersetzt werden, die lediglich die für das weitere Schließen notwendigen Informationen enthalten, d. h. ob es sich um einen Term (dann welcher Sorte) oder einen Ausdruck handelt und welche Variablen darin vollfrei bzw. gebunden vorkommen.

Wir demonstrieren dieses Verfahren zunächst am Beispiel der Sprache S_{eukl} der ebenen euklidischen Geometrie (vgl. 3.3., Beispiel 2), die wir für den jetzigen Zweck noch wie folgt abändern: Als Punktvariablen dienen $\mathrm{P}\underbrace{\mathrm{a}\ldots\mathrm{a}}_{i}$ $(i \geqq 0)$, als Geradenvariablen $\mathrm{g}\underbrace{\mathrm{a}\ldots\mathrm{a}}_{i}$ $(i \geqq 0)$, wobei wir diese Zeichenreihen jedoch im folgenden abkürzend wieder mit P_i bzw. g_i bezeichnen. Statt P_i auf g_j schreiben wir $\mathrm{P}_i \in \mathrm{g}_j$, statt $[\mathrm{P}_i, \mathrm{P}_j, \mathrm{P}_k]$ schreiben wir $\mathrm{ZP}_i\mathrm{cP}_j\mathrm{cP}_k\mathrm{c}$. Damit ist S_{eukl} eine Menge von Wörtern im endlichen Alphabet

$$A = \{\mathrm{P}, \mathrm{g}, \mathrm{a}, \mathrm{c}, \in, \mathrm{Z}, \cong, =, \neg, \wedge, \vee, \rightarrow, \leftrightarrow, \bigwedge, \bigvee, (,)\}.$$

$\mathrm{H}(\mathrm{P}_{i_1}\mathrm{c}\ldots\mathrm{cP}_{i_n}\mathrm{cg}_{j_1}\mathrm{c}\ldots\mathrm{cg}_{j_m}\mathrm{cdP}_{k_1}\mathrm{c}\ldots\mathrm{cP}_{k_r}\mathrm{cg}_{l_1}\mathrm{c}\ldots\mathrm{cg}_{l_s}\mathrm{c})$ bedeute einen Ausdruck, der die Punkte- bzw. Geradenvariablen, die vor dem Trennzeichen d stehen, vollfrei und die hinter d stehenden Variablen gebunden enthält. Dabei brauchen die aufgeführten Variablen nicht unbedingt paarweise verschieden zu sein. Den induktiven Definitionsregeln für die Sprache S_{eukl} entsprechen nun folgende Einschmelzungsregeln der Form $\mathrm{U} \Rightarrow \mathrm{W}$, wobei eine einzelne Anwendung einer solchen Regel darin besteht, ein beliebiges Teilwort der Form U durch das entsprechende Teilwort W zu ersetzen:

$$\mathrm{P}_i = \mathrm{P}_j \Rightarrow \mathrm{H}(\mathrm{P}_i\mathrm{cP}_j\mathrm{cd})$$

$$\mathrm{g}_i = \mathrm{g}_j \Rightarrow \mathrm{H}(\mathrm{g}_i\mathrm{cg}_j\mathrm{cd})$$

$$\mathrm{P}_i \in \mathrm{g}_j \Rightarrow \mathrm{H}(\mathrm{P}_i\mathrm{cg}_j\mathrm{cd})$$

$$\mathrm{ZP}_i\mathrm{cP}_j\mathrm{cP}_k\mathrm{c} \Rightarrow \mathrm{H}(\mathrm{P}_i\mathrm{cP}_j\mathrm{cP}_k\mathrm{cd})$$

$$P_iP_j \cong P_kP_l \Rightarrow H(P_icP_jcP_kcP_lcd)$$

$$\neg H \Rightarrow H$$

$$\big(H(\ldots) \circ H(\ldots)\big) \Rightarrow H(\ldots), \qquad (\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}),$$

wobei in der neuen Zeichenreihe alle diejenigen Variablen hinter dem Trennzeichen d, d. h. als gebunden aufgezählt werden, die in wenigstens einer der beiden alten Zeichenreihen als gebunden aufgezählt waren, während diejenigen Variablen vor dem Trennzeichen d als vollfrei aufgezählt werden, die in einer der beiden alten Zeichenreihen als vollfrei und in der anderen nicht als gebunden aufgezählt waren.

$$\bigwedge P_iH(UP_icVdW) \Rightarrow H(RdWP_ic), \qquad \text{wobei } R = \mathit{Sub}(UP_icV, P_ic, \Lambda), \qquad (1)$$

d. h., falls eine Teilzeichenreihe die Form $\bigwedge P_iH(\ldots)$ hat und unter den in $H(\ldots)$ als vollfrei aufgezählten Variablen P_i wenigstens einmal vorkommt, wird der Anfang $\bigwedge P_i$ gestrichen, desgleichen alle Vorkommen von P_ic vor dem Trennzeichen d. Statt dessen wird P_ic am Ende der Liste der gebundenen Variablen neu hinzugefügt. Man beachte die Rolle des hier mit c bezeichneten „Kommas"! Würde man die Variablen einfach als Pa ... aPa ... aPa ... a usw. aneinanderreihen, so würden beim Streichen aller Vorkommen von P_i auch die entsprechenden Anfänge aller P_j mit $j > i$ gestrichen.

Zu (1) analoge Regeln sind nun noch für $\bigwedge g_iH(\ldots)$, $\bigvee P_iH(\ldots)$, $\bigvee g_iH(\ldots)$ zu formulieren.

Zwei Beispiele mögen die Wirkung des Systems der Einschmelzungsregeln verdeutlichen. Anwendung auf die Zeichenreihe

$$\bigvee P_1\big((P_1P_2 \cong P_3P_5 \wedge ZP_1cP_3cP_4c) \rightarrow \bigvee g_2(P_1 \in g_2 \wedge \bigwedge P_2ZP_2cP_1cP_3c)\big)$$

liefert nach und nach folgende Zwischenresultate:

$$\bigvee P_1\Big(\big(H(P_1cP_2cP_3cP_5cd) \wedge H(P_1cP_3cP_4cd)\big) \rightarrow \bigvee g_2\big(H(P_1cg_2cd) \wedge \bigwedge P_2\; H(P_2cP_1cP_3cd)\big)\Big),$$

$$\bigvee P_1\Big(H(P_1cP_2cP_3cP_4cP_5cd) \rightarrow \bigvee g_2\big(H(P_1cg_2cd) \wedge H(P_1cP_3cdP_2c)\big)\Big),$$

$$\bigvee P_1\big(H(P_1cP_2cP_3cP_4cP_5cd) \rightarrow \bigvee g_2H(P_1cP_3cg_2cdP_2c)\big),$$

$$\bigvee P_1\big(H(P_1cP_2cP_3cP_4cP_5cd) \rightarrow H(P_1cP_3cdP_2cg_2c)\big),$$

$$\bigvee P_1H(P_1cP_3cP_4cP_5cdP_2cg_2c),$$

$$H(P_3cP_4cP_5cdP_2cg_2cP_1c),$$

d. h., die untersuchte Zeichenreihe ist ein Ausdruck, in dem die Variablen P_3, P_4, P_5 vollfrei und die Variablen P_1, P_2, g_2 gebunden vorkommen.

Wendet man das gleiche Verfahren auf die Zeichenreihe

$$\bigwedge g_3 \bigvee P_1(P_1P_2 \cong P_1P_3 \wedge P_2 \in g_3) \rightarrow P_2 \in g_3)$$

an, so erhält man nach einigen Zwischenschritten

$$H(P_2cP_3cdP_1cg_3c) \rightarrow H(P_2cg_3cd)),$$

worauf wegen der vorn fehlenden Klammer keine Regel mehr anwendbar ist, d. h., diese Zeichenreihe ist kein Ausdruck.

Prinzipiell kommt es offenbar nicht auf die Reihenfolge der Anwendung von Einschmelzungsregeln an. Um jedoch einen Algorithmus zu erhalten, kann man etwa vereinbaren, die Liste der Regeln als Programm im Markovschen Sinne anzusehen, d. h. die Regeln in der Reihenfolge der Liste auf ihre Anwendbarkeit zu prüfen, jeweils die am weitesten links im bearbeiteten Wort ersetzbaren Teilwörter umzuformen und nach jeder Anwendung einer Regel zur Prüfung der Anwendbarkeit der ersten Regel zurückzukehren. Das Verfahren ist dann beendet, wenn keine Regel mehr anwendbar ist, so daß das Resultat $\mathfrak{A}(W)$ für beliebige $W \in W(A)$ eindeutig bestimmt ist. Es ist nun zu zeigen:

Für $W \in W(A)$ *existiert* $\mathfrak{A}(W)$, *und* $\mathfrak{A}(W)$ *hat genau dann die Form* H(U), *wobei in* U *das Zeichen*) *nicht vorkommt, wenn* $W \in S_{eukl}$ *ist.*

Der erste Teil der Behauptung folgt sofort daraus, daß die Gesamtzahl der Buchstaben $=, \in, Z, \cong, \neg, \wedge, \vee, \rightarrow, \leftrightarrow, \bigwedge, \bigvee$ bei jeder Anwendung einer Regel um eins abnimmt, so daß das Resultat spätestens dann erreicht wird, wenn diese Gesamtzahl null ist.

Zum Beweis der zweiten Behauptung zeigt man

a) durch Induktion über die Kompliziertheit der Ausdrücke H: *Für* $H \in S_{eukl}$ *hat* $\mathfrak{A}(H)$ *die Form* H(U), *wobei*) *nicht in* U *vorkommt*;

b) durch Induktion über die Schrittzahl: *Hat* $\mathfrak{A}(W)$ *die Form* H(... d ...), *und ist dieses Resultat in n Schritten erreicht worden, so ist* W *ein Ausdruck von* S_{eukl}, *in dem genau die vor bzw. hinter dem Trennzeichen* d *aufgezählten Variablen vollfrei bzw. gebunden vorkommen.*

Für b) skizzieren wir den Beweis. Wird das Resultat in einem Schritt erreicht, so muß W ein prädikativer Ausdruck von S_{eukl} sein, da dann nur eine der fünf ersten Regeln in Frage kommt. Alle übrigen Regeln enthalten nämlich in ihren linken Seiten den Buchstaben H, der im Eingabewort W nicht vorkommen kann. Die Behauptung sei nun schon für alle $W \in W(A)$ bewiesen, für die das Resultat in höchstens n Schritten erreicht wird, und W sei jetzt ein solches Wort, für das das Resultat $\mathfrak{A}(W)$ in $n + 1$ Schritten erhalten wird. Man hat nun eine umfangreiche Fallunterscheidung durchzuführen, je nachdem, welche Regel als letzte angewendet wurde. War dies etwa die Regel (1), so muß das letzte Zwischenresultat von der Form $\wedge P_iH(\ldots P_i \ldots d \ldots)$ gewesen und in n Schritten erreicht worden sein. Da keine der Regeln ein Schreiben von $\bigwedge P_i$ vor ein Wort vorsieht, muß schon das Eingabewort die Form $\bigwedge P_iR$ gehabt haben, so daß in den gleichen n Schritten R in $H(\ldots P_i \ldots d \ldots)$ übergeführt wurde. Nach Induktionsannahme ist daher R ein Ausdruck aus S_{eukl}, in dem genau die vor bzw. hinter d aufgezählten Variablen vollfrei bzw. gebunden vorkommen. Daher ist $W = \wedge P_i R$ ein Ausdruck, der die gleichen Variablen vollfrei bzw. gebunden enthält, abgesehen davon, daß P_i jetzt gebunden statt vollfrei vorkommt. Laut Regel (1) wird aber P_i beim $(n+1)$-ten Schritt vor d gelöscht und hinter d neu geschrieben.

Um das Vorgehen im Fall von Sprachen mit „echten" Termen zu erläutern, betrachten wir noch die Sprache der Vektorrechnung, in der Termgleichungen die einzigen prädikativen Ausdrücke sind. Wir nehmen jetzt an, daß alle Skalarvariablen die Form s_i und alle Vektorvariablen die Form v_j haben (ausführlich $s\underbrace{a \ldots a}_{i}$ bzw. $v\underbrace{a \ldots a}_{j}$). Ferner erinnern wir an die Bedeutung folgender Symbole (vgl. 3.3., Beispiel 3):

o: Nullvektor,

0 bzw. 1: Skalarnull bzw. -eins,

$\oplus$: Vektoraddition,

$+$ bzw. $\cdot$: Addition bzw. Multiplikation der Skalare,

$\circ$: Skalarmultiplikation.

$V(s_{i_1}c \ldots s_{i_n}c\, v_{j_1}c \ldots v_{j_m}c)$ bezeichne nun einen Term der Sorte Vektor, der genau die angegebenen Variablen enthält, analog $S(s_{i_1}c \ldots s_{i_n}cv_{j_1}c \ldots v_{j_m}c)$ einen Term der Sorte Skalar. Insbesondere bezeichnen S bzw. V allein variablenfreie Terme. An die Stelle der für S_{sukl} angegebenen Einschmelzungsregeln für prädikative Ausdrücke treten nun folgende Regeln (die übrigen können sinngemäß übernommen werden):

$$o \Rightarrow V \qquad s_i \Rightarrow S(s_i c)$$

$$0 \Rightarrow S \qquad v_j \Rightarrow V(v_j c)$$

$$1 \Rightarrow S$$

$$\big(V(s_{i_1}c \ldots s_{i_n}cv_{j_1}c \ldots v_{j_m}c) \oplus V(s_{k_1}c \ldots s_{k_r}sv_{l_1}c \ldots v_{l_s}c)\big)$$

$$\Rightarrow V(s_{i_1}c \ldots s_{i_n}cs_{k_1}c \ldots s_{k_r}cv_{j_1}c \ldots v_{j_m}c\, v_{l_1}c \ldots v_{l_s}c)$$

analog

$$\big(S(\ldots) + S(\ldots)\big) \Rightarrow S(\ldots)$$

$$\big(S(\ldots) \cdot S(\ldots)\big) \Rightarrow S(\ldots)$$

$$\big(S(\ldots) \circ V(\ldots)\big) \Rightarrow V(\ldots)$$

$$S(\ldots) = S(\ldots) \Rightarrow H(\ldots d)$$

$$V(\ldots) = V(\ldots) \Rightarrow H(\ldots d).$$

Neben und vor der Entscheidbarkeit der Sprache $\boldsymbol{S}$ ist hier natürlich die Entscheidbarkeit der Mengen $\boldsymbol{T}^V$ bzw. $\boldsymbol{T}^S$ aller Terme der Sorte Vektor bzw. Skalar von Interesse. Es ist plausibel (und analog zum Satz über die Entscheidbarkeit von $\boldsymbol{S}_{eukl}$ zu beweisen), daß genau die Zeichenreihen aus $\boldsymbol{T}^V$ sich durch Anwendung von Einschmelzungsregeln auf die Form V(...) bringen lassen und daß Analoges für $\boldsymbol{T}^S$ gilt.

Als Folgerung aus der Entscheidbarkeit der formalisierten Sprachen entnehmen wir für die beiden folgenden Abschnitte:

Ist $\boldsymbol{A}$ ein endliches Alphabet, $\boldsymbol{S} \subseteq W(\boldsymbol{A})$ eine formalisierte Sprache und existiert ein auf alle Wörter $\mathrm{W} \in \boldsymbol{S}$ anwendbares Entscheidungsverfahren für eine Teilmenge $\boldsymbol{M}$ von $\boldsymbol{S}$, so ist $\boldsymbol{M}$ auch eine entscheidbare Teilmenge von $W(\boldsymbol{A})$. Um $\mathrm{W} \in \boldsymbol{M}$? für $\mathrm{W} \in W(\boldsymbol{A})$ zu entscheiden, hat man zunächst $\mathrm{W} \in \boldsymbol{S}$ zu testen und im Fall positiven Ausgangs anschließend das für Elemente von $\boldsymbol{S}$ verwendbare Entscheidungsverfahren für $\boldsymbol{M}$ anzuwenden.

8.2. Entscheidbarkeit formalisierter Theorien, Satz von Church

Es sei $(\boldsymbol{S}, \sigma)$ eine beliebige (insbesondere eventuell elementare) Sprache, wobei wir annehmen, daß ein endliches Alphabet $\boldsymbol{A}$ mit $\boldsymbol{S} \subset W(\boldsymbol{A})$ existiert (vgl. 8.1.). Eine (durch ein Axiomensystem $\boldsymbol{X}$ gegebene) in $(\boldsymbol{S}, \sigma)$ formulierte Theorie $\boldsymbol{T} = \boldsymbol{Fl}_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$ heißt *entscheidbar*, wenn $\boldsymbol{T}$ eine (etwa im Sinne der Turing-Entscheidbarkeit) entscheidbare Teilmenge von $W(\boldsymbol{A})$ ist. Als Spezialfall ergibt sich für $\boldsymbol{X} = \emptyset$ die Frage nach der Entscheidbarkeit der Mengen $\boldsymbol{Fl}_{\boldsymbol{S}}{}^{\sigma}(\emptyset)$, d. h. nach der Entscheidbarkeit der logischen Allgemeingültigkeit in der Sprache $(\boldsymbol{S}, \sigma)$. Die folgenden Überlegungen werden zeigen, daß die meisten „wichtigen" bzw. „interessanten" mathematischen Theorien nicht entscheidbar sind und daß insbesondere die logische Allgemeingültigkeit nur für sehr spezielle Sprachen entscheidbar ist.

Definition 1. Es sei $\boldsymbol{B}$ ein beliebiges endliches Alphabet. Eine Theorie $\boldsymbol{Fl}_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X}) = \boldsymbol{T} \subset W(\boldsymbol{A})$ heißt ***B-ausdrucksfähig***, wenn es eine berechenbare eineindeutige Abbildung (Wortfunktion) h von $W(\boldsymbol{B})$ in die Menge der variablenfreien Terme von $\boldsymbol{S}$ und zu jeder berechenbaren Wortfunktion f in $W(\boldsymbol{B})$ einen Ausdruck $H_f(\mathrm{x}, \mathrm{y})$ mit genau zwei vollfreien und keinen weiteren freien Variablen gibt, so daß für alle Wörter $\mathrm{U}, \mathrm{V} \in W(\boldsymbol{B})$ gilt:

$$f(\mathrm{U}) = \mathrm{V} \textit{ genau dann, wenn } \mathrm{H}_f(h(\mathrm{U}), h(\mathrm{V})) \in \boldsymbol{T}.$$

Insbesondere heißt eine Sprache $(\boldsymbol{S}, \sigma)$ $\boldsymbol{B}$-ausdrucksfähig, wenn die Theorie $\boldsymbol{Fl}_{\boldsymbol{S}}{}^{\sigma}(\emptyset)$ $\boldsymbol{B}$-ausdrucksfähig ist.

Definition 1 bedeutet im wesentlichen, daß eine Theorie bzw. Sprache $\boldsymbol{B}$-ausdrucksfähig ist, wenn man in ihr über die Wörter des Alphabets $\boldsymbol{B}$ sprechen und einen präzisierten Berechenbarkeitsbegriff für die Wortfunktionen in $W(\boldsymbol{B})$ definieren kann.

Satz 1. *Zu jedem endlichen Alphabet* $\boldsymbol{B}$ *gibt es eine* $\boldsymbol{B}$*-ausdrucksfähige elementare Theorie* $\boldsymbol{T} = Fl_{\boldsymbol{S}}(\boldsymbol{X})$ *mit endlichem Axiomensystem* $\boldsymbol{X}$.

Zum Beweis von Satz 1 wählen wir eine einsortige Sprache $\boldsymbol{S}$ zur Formalisierung des elementaren Teils der Theorie der freien Halbgruppen (vgl. Kapitel 1). Ist $\boldsymbol{B} = \{b_1, \dots, b_m\}$ das gegebene Alphabet, so enthalte $\boldsymbol{S}$ Variablen x, x_i für Wörter, ein Konstantensymbol 0 zur Bezeichnung des leeren Wortes, Konstantensymbole $c_1, \dots, c_m$ zur Bezeichnung der Buchstaben $b_1, \dots, b_m$ und für spätere Zwecke noch weitere Konstantensymbole a, b, c, d, ferner ein zweistelliges Operationssymbol $+$ zur Bezeichnung der Verkettung mit der Vorschrift, Terme nach dem Schema $(t_1 + t_2)$ zu bilden. In der so beschriebenen Sprache $\boldsymbol{S}$ können wir ein endliches Axiomensystem $\boldsymbol{X}$ formulieren, das alle in Kapitel 1 formulierten Halbgruppenaxiome bis auf das nichtelementare Axiom A3 umfaßt:

$$\big((x_1 + x_2) + x_3\big) = \big(x_1 + (x_2 + x_3)\big) \qquad \text{(entspricht A1)},$$

$$\left.\begin{array}{l} (x + 0) = x \\ (0 + x) = x \end{array}\right\} \text{(entspricht A2)},$$

$$(x_1 + x_2) = 0 \to x_1 = 0 \wedge x_2 = 0 \qquad \text{(entspricht A4a)},$$

$$(x_1 + y) = (x_2 + y) \to x_1 = x_2 \qquad (y = c_1, \dots, c_m, a, b, c, d),$$

$$\neg\, (x_1 + y) = (x_2 + z) \qquad (y, z = c_1, \dots, c_m, a, b, c, d, y \neq z)$$

(Die endlich vielen Axiome, die durch alle genannten Einsetzungen für y, z aus diesen beiden abkürzenden Schemata entstehen, entsprechen zusammen A4b.) Hieraus folgt

$$\neg\, y = z \qquad (y, z = c_1, \dots, c_m, a, b, c, d, y \neq z)$$

(Diese Sätze bedeuten, daß die von 0 verschiedenen Konstanten paarweise verschiedene „Buchstaben" bezeichnen. Hieraus folgt insbesondere A5 (vgl. Kapitel 1)).

Die völlige Unterdrückung von A3 hat zur Folge, daß es in einem beliebigen Modell ω von $\boldsymbol{X}$ außer den Wörtern im Alphabet

$$\{\omega(c_1), \dots, \omega(c_m), \omega(a), \omega(b), \omega(c), \omega(d)\}$$

eventuell noch weitere Objekte gibt, z. B. Folgen unendlicher Länge oder Wörter, die unter Verwendung weiterer „Buchstaben" aufgebaut sind. Jedoch sichern die formulierten Axiome bereits, daß es in jedem Modell von $\boldsymbol{X}$ einen zur Wortmenge $W(\{c_1, \dots, c_m, a, b, c, d\})$ isomorphen Teilbereich gibt. Bei weiteren Formulierungen in der Sprache S werden wir zur Abkürzung alle Termklammern und das Verkettungssymbol $+$ wieder weglassen und ferner die bereits in Kapitel 1 eingeführten abkürzenden Ausdrücke „$\mathbf{x}_i$ beginnt mit $\mathbf{x}_j$" (d. h. $\bigvee \mathbf{x}\mathbf{x}_i = \mathbf{x}_j\mathbf{x}$), „$\mathbf{x}_i$ endet mit $\mathbf{x}_j$" und „$\mathbf{x}_i$ kommt in $\mathbf{x}_j$ vor" benutzen. (Natürlich hätten wir unsere Sprache S gleich der beim Behandeln von Wörtern üblicherweise verwendeten Sprache angleichen können. Dann würde jedoch vielleicht für manchen Leser der Unterschied zwischen den formalisierten Ausdrücken der Sprache S und den daneben „metatheoretisch" benutzten Aussagen über Wörter, z. B. über die Zeichenreihen der Sprache S, zu undeutlich.) Die durch

$$h(\Lambda) := 0, \quad h(\mathrm{Wb}_i) := \big(h(\mathrm{W}) + \mathrm{c}_i\big) \quad (i = 1, \ldots, m)$$

induktiv über die Wortkompliziertheit in $W(\boldsymbol{B})$ definierte Abbildung h ist offenbar eine eineindeutige Abbildung von $W(\boldsymbol{B})$ in (nicht auf!) die Menge der variablenfreien Terme von S, womit die erste an die $\boldsymbol{B}$-Ausdrucksfähigkeit von S gestellte Forderung erfüllt ist. Unter Benutzung weiterer Konstanten a, b, c, d formalisieren wir nun in S die Turing-Berechenbarkeit der Wortfunktionen in $W(\boldsymbol{B})$. Dazu kodieren wir durch $\mathrm{b}\underbrace{\mathrm{a} \ldots \mathrm{a}}_{n}\mathrm{b}$ $(n \geqq 1)$ die in beliebigen Turingprogrammen benutzten Hilfsbuchstaben, insbesondere bedeuten bab bzw. baab das Startzeichen s und das Leerzeichen o und werden im folgenden auch so abgekürzt. Ganz ausführlich gesagt bedeutet also der Buchstabe s in Ausdrücken der Sprache S eine Abkürzung für den variablenfreien Term $\big((\mathrm{b} + \mathrm{a}) + \mathrm{b}\big)$. Analog kodieren wir durch bcb, bccb, bcccb die Steuerbefehle L, N, R und kürzen diese Terme von S im folgenden auch so ab. Durch $\mathrm{b}\underbrace{\mathrm{c} \ldots \mathrm{c}}_{n}\mathrm{b}$ $(n \geqq 4)$ kodieren wir schließlich die in einem beliebigen Turingprogramm benötigten Zustände q_i und kürzen diese Terme im folgenden auch durch q_i ab. Nach diesen Vorbereitungen wollen wir zu gegebener Turing-berechenbarer Wortfunktion f in $W(\boldsymbol{B})$ einen Ausdruck $\mathrm{H}_f(\mathbf{x}_1, \mathbf{x}_2) \in S$ mit den in der Definition 1 verlangten Eigenschaften konstruieren. Es sei τ ein sauberes Turingprogramm zur Berechnung von f. Den in 7.4. definierten Konstellationsbegriff modifizieren wir hier so, daß die Zustandsbezeichnung jeweils links neben der gelesenen Zelle steht. Wir geben nun zunächst einen Ausdruck $\mathrm{H}_f'(\mathbf{x}_3, \mathbf{x}_4)$ mit folgender Bedeutung an: $\mathbf{x}_3$ und $\mathbf{x}_4$ sind Konstellationen bezüglich τ, so daß $\mathbf{x}_3$ in einem Takt durch τ in $\mathbf{x}_4$ übergeführt wird. Im konkreten Fall ergibt sich $\mathrm{H}_f'(\mathbf{x}_3, \mathbf{x}_4)$ aus der Tabelle von τ im wesentlichen durch alternative Verknüpfung aller endlich vielen möglichen Fälle von Zustand und gelesenem Buchstaben. Dies sei am einfachen Beispiel des Programms

τ	s	o	b_1
q_1	R	q_2 L	o R
q_2	q_0 N	L	—

demonstriert. Der Ausdruck $H_f'(x_3, x_4)$ lautet in diesem Fall wie folgt:

$$\bigvee x_5(x_3 = q_1 s x_5 \wedge x_4 = s q_1 x_5)$$

$$\vee \bigvee x_5(x_3 = x_5 o q_1 \wedge x_4 = x_5 q_2) \vee (x_3 = s q_1 \wedge x_4 = q_2 s)$$

$$\vee \bigvee x_5 x_6(x_3 = x_5 q_1 c_1 x_6 \wedge x_4 = x_5 o q_1 x_6) \vee (x_3 = q_2 s \wedge x_4 = q_0 s)$$

$$\vee \bigvee x_5(x_3 = x_5 o q_2 \wedge x_4 = x_5 q_2) \vee (x_3 = s q_2 \wedge x_4 = q_2 s).$$

Dabei gehört das erste Alternativglied zum Fall (q_1, s), die nächsten beiden Alternativglieder gehören zum Fall (q_1, o), und zwar das erste zu dem Unterfall, daß in der linken Nachbarzelle das Leerzeichen steht, das letzte zu dem Unterfall, daß in der linken Nachbarzelle das Zeichen s steht. Das vierte Alternativglied gehört zum Fall (q_1, b_1), wobei b_1 in der Sprache S die Konstante c_1 entspricht. Das fünfte Alternativglied gehört zum Fall (q_2, s) und die letzten beiden zum Fall (q_2, o).

Es läge nun nahe, den Ausdruck $H_f(x_1, x_2)$ wie folgt zu bilden:

$$\bigvee n x_3 x_4 \ldots x_n\big(H_f'(q_1 s x_1, x_3) \wedge H_f'(x_3, x_4) \wedge H_f'(x_4, x_5) \wedge \ldots \wedge H_f'(x_n, q_0 s x_2)\big)$$

Etwas Derartiges ist jedoch in einer elementaren Sprache nicht formulierbar. Wir können uns hier dadurch helfen, daß wir die endliche Folge der Konstellationen von τ, die schließlich von der Anfangskonstellation $q_1 s x_1$ zur Endkonstellation $q_0 s x_2$ führt, mit Hilfe des bisher noch nicht verwendeten „Buchstaben" d zu einem langen Wort $q_1 s x_1 d x_3 d x_4 d \ldots d x_n d q_0 s x_2$ zusammenkleben. Die Bauart eines solchen Wortes läßt sich in der Sprache S elementar beschreiben, und der gesuchte Ausdruck $H_f(x_1, x_2)$ lautet

$$\bigvee x\big(x \text{ beginnt mit } q_1 s x_1 d \wedge x \text{ endet mit } d q_0 s x_2$$

$$\wedge \bigwedge x_3 x_4\big((x \text{ beginnt mit } x_3 d x_4 d \vee d x_3 d x_4 d \text{ kommt in } x \text{ vor} \vee x \text{ endet mit }$$

$$d x_3 d x_4)$$

$$\wedge \neg\, d \text{ kommt in } x_3 x_4 \text{ vor} \rightarrow H_f'(x_3, x_4)\big)\big).$$

Damit beschließen wir den Beweis von Satz 1.

Satz 2. *Ist eine Theorie* ***B****-ausdrucksfähig für ein beliebiges endliches Alphabet* ***B****, so ist sie auch* ***C****-ausdrucksfähig für jedes andere endliche Alphabet* ***C****.*

Beweis. Es sei $Fl_S^\sigma(X)$ eine B-ausdrucksfähige Theorie, C ein beliebiges endliches Alphabet. Dann gibt es eine eineindeutige Abbildung k von $W(C)$ in $W(B)$, so daß k und k^{-1} berechenbare Wortfunktionen sind (vgl. 7.1.). Ist h eine eineindeutige berechenbare Wortfunktion von $W(B)$ in die Menge der variablenfreien Terme der Sprache S, so ist $h \cdot k$ eine eineindeutige berechenbare Abbildung von $W(C)$ in die Menge der variablenfreien Terme von S. Ist f eine berechenbare Wortfunktion in $W(C)$, so ist

$$\bar{f} := k \cdot f \cdot k^{-1}$$

eine berechenbare Wortfunktion in $W(\boldsymbol{B})$. Daher existiert nach Voraussetzung ein Ausdruck $H_{\bar{f}}(x, y) \in \boldsymbol{S}$, so daß für $U, V \in W(\boldsymbol{B})$ gilt: $\bar{f}(U) = V$ genau dann, wenn $H_{\bar{f}}(h(U), h(V)) \in Fl_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$. Daher gilt insbesondere für Wörter $R, S \in W(\boldsymbol{C})$: $f(R) = S$ genau dann, wenn $\bar{f}(k(R)) = k(S)$, und das letzte ist genau dann der Fall, wenn $H_{\bar{f}}(h(k(R)), h(k(S))) \in Fl_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$ gilt. Das heißt aber, $Fl_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$ ist $\boldsymbol{C}$-ausdrucksfähig.

Satz 2 rechtfertigt die

Definition 3. Eine Theorie $Fl_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$ heißt *ausdrucksfähig*, wenn sie ausdrucksfähig für wenigstens ein endliches Alphabet (und damit für jedes endliche Alphabet) $\boldsymbol{B}$ ist.

Satz 3. *Jede ausdrucksfähige Theorie ist unentscheidbar.*

Beweis. Es sei $\boldsymbol{T} = Fl_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$ ausdrucksfähig und dabei $\boldsymbol{A}$ ein endliches Alphabet mit $\boldsymbol{S} \subset W(\boldsymbol{A})$. Dann ist $\boldsymbol{T}$ insbesondere $\boldsymbol{A}$-ausdrucksfähig, d. h., es gibt eine eineindeutige berechenbare Wortfunktion h von $W(\boldsymbol{A})$ in die Menge der variablenfreien Terme von $\boldsymbol{S}$ und zu jeder berechenbaren Wortfunktion f in $W(\boldsymbol{A})$ einen Ausdruck $H_f(x, y) \in \boldsymbol{S}$, so daß für $U, V \in W(\boldsymbol{A})$

$$f(U) = V \text{ genau dann, wenn } H_f(h(U), h(V)) \in \boldsymbol{T}, \tag{1}$$

gilt. Dabei kann man offenbar bei beliebiger Funktion f immer die gleichen Variablen x, y in den Ausdrücken $H_f(x, y)$ benutzen. Da die durch $g(U) := Sub(U, x, h(U))$ definierte Wortfunktion in $W(\boldsymbol{A})$ überall definiert und offenbar berechenbar ist, folgt aus der Annahme der Entscheidbarkeit von $\boldsymbol{T}$, daß auch die durch

$$f(U) := \begin{cases} \Lambda, & \textit{falls } g(U) \in W(\boldsymbol{A}) \setminus \boldsymbol{T}, \\ \textit{nicht definiert, falls } g(U) \in \boldsymbol{T} \end{cases} \tag{2}$$

definierte Wortfunktion f berechenbar ist. Dies führen wir zum Widerspruch, indem wir das Wort $U = H_f(x, h(\Lambda))$ betrachten. Nach (1) ist $f(U) = \Lambda$ genau dann, wenn $H_f(h(U), h(\Lambda)) \in \boldsymbol{T}$. Nach (2) ist $f(U) = \Lambda$ genau dann, wenn $g(U) \notin \boldsymbol{T}$, aber offenbar ist $g(U) = H_f(h(U), h(\Lambda))$.

Satz 4. *Existiert in einer Sprache* $(\boldsymbol{S}, \sigma)$ *ein endliches Axiomensystem* $\boldsymbol{X}$, *so daß die Theorie* $Fl_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$ *unentscheidbar ist, so ist auch* $Fl_{\boldsymbol{S}}{}^{\sigma}(\emptyset)$ *unentscheidbar.*

Beweis. Ist $\boldsymbol{X}$ endlich, so können wir $Fl_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$ auch als $Fl_{\boldsymbol{S}}{}^{\sigma}(\{H_0\})$ mit $H_0 \in \bar{\boldsymbol{S}}$ erhalten, indem wir von den ursprünglichen Axiomen zunächst zu gleichbedeutenden abgeschlossenen Ausdrücken übergehen und diese dann konjunktiv miteinander zu einem einzigen Axiom $H_0 \in \bar{\boldsymbol{S}}$ verknüpfen. Dann gilt für $H \in \boldsymbol{S}$: $H \in Fl_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$ genau dann, wenn $H_0 \to H \in Fl_{\boldsymbol{S}}{}^{\sigma}(\emptyset)$. Gäbe es daher ein Entscheidungsverfahren für das letzte Problem, so hätte man es nur auf den aus H effektiv erhältlichen Ausdruck $H_0 \to H$ anzuwenden.

Satz 5. *Ist* (S', σ') *eine definitorische Erweiterung von* (S, σ) *mit berechenbarer Rückübersetzung* f (vgl. 4.6.) *und ist* $Fl_{S'}^{\sigma'}(\emptyset)$ *unentscheidbar, so ist auch* $Fl_S^{\sigma}(\emptyset)$ *unentscheidbar.* (Alle zu den in 4.6. behandelten Definitionsmethoden gehörigen Rückübersetzungsvorschriften lassen sich so normieren, daß die Zuordnung f, die einem Ausdruck $H \in S'$ den entsprechenden Ausdruck $[H] \in S$ zuordnet, eine berechenbare Wortfunktion wird.)

Beweis. Nach Definition 2 aus 4.6. gilt für solche Ausdrücke $H \in S'$, die eine Rückübersetzung $[H] = f(H)$ besitzen,

$$H \leftrightarrow [H] \in Fl_{S'}^{\sigma'}(\emptyset),$$

also insbesondere $H \in Fl_{S'}^{\sigma'}(\emptyset)$ genau dann, wenn $[H] \in Fl_S^{\sigma}(\emptyset)$. Wäre nun $Fl_S^{\sigma}(\emptyset)$ entscheidbar, so hätte man, um $H \in Fl_{S'}^{\sigma'}(\emptyset)$? zu entscheiden, nur zu einem zu H gleichbedeutenden Ausdruck $\overline{\overline{H}}$ (notfalls $\overline{\overline{H}}$ abgeschlossen) überzugehen, der eine Rückübersetzung besitzt, und dann $[\overline{\overline{H}}] \in Fl_S^{\sigma}(\emptyset)$ zu testen.

Satz 6 (Satz von Church). *Enthält eine einsortige Sprache* S *wenigstens ein zweistelliges Relationssymbol oder ein zweistelliges Operationssymbol, so ist die Menge* $Fl_S(\emptyset)$, *d. h. die Allgemeingültigkeit in* S *bezüglich des elementaren Folgerns, unentscheidbar.*

Der wesentliche Teil des Beweises dieses Satzes steckt in den Sätzen 1 bis 5. Enthält S ein zweistelliges Relationssymbol, so kann man dieses z. B. als Elementrelation der Mengenlehre deuten und jede beliebige Formalisierung der Theorie der Wörter und ihrer rekursiven Berechenbarkeit (wie auch jede andere mathematische Theorie) als definitorische Erweiterung von S auffassen. Enthält S ein zweistelliges Operationssymbol F, so kann man z. B. die Tatsache benutzen, daß sich aus der zunächst inhaltlich durch

$$F(x, y) := \begin{cases} \emptyset, & \textit{falls } x \in y, \\ \{\emptyset\}, & \textit{falls } x \notin y, \end{cases}$$

definierten „charakteristischen Funktion" der Elementrelation einer geeignet formalisierten Mengenlehre die Konstante $\emptyset$ und die $\in$-Relation zurückdefinieren lassen:

$$\emptyset := \iota z\big(\bigvee x F(z, x) = z \wedge \neg \bigvee x F(x, z) = z\big),$$
$$x \in y :\leftrightarrow F(x, y) = \emptyset.$$

Da man mittels eines n-stelligen Relationssymbols R $(n > 2)$ eine zweistellige Relation R′ (z. B. durch $R'(x, y) :\leftrightarrow \bigvee z_1 \ldots z_{n-2} R(z_1, \ldots, z_{n-2}, x, y)$) und mittels eines n-stelligen Operationssymbols F $(n > 2)$ eine zweistellige Operation F′ (z. B. durch $F'(x, y) := F(x, \ldots, x, y)$) definieren kann, folgt aus dem Satz von Church (unter Berücksichtigung von Satz 5) sofort die Unentscheidbarkeit der Allgemeingültigkeit in allen einsortigen elementaren Sprachen mit Ausnahme derjenigen, deren Basis nur aus einstelligen Relationssymbolen und eventuell Konstantensymbolen besteht. In bezug auf mehrsortige Sprachen sei nur bemerkt, daß sich der Satz von CHURCH sofort auf alle Fälle ausdehnen läßt, in denen für eine gewisse Sortennummer j eine Relation vom Typ (j, j) oder eine Operation vom Typ $(j, j; i)$ (i, j beliebige Sortennummern) definierbar ist.

8.3. Entscheidbarkeit der Allgemeingültigkeit von Ausdrücken, die nur einstellige Relationssymbole enthalten

Aus Gründen der einfacheren Formulierung beschränken wir uns hier auf Ausdrücke einsortiger Sprachen, die aus prädikativen Ausdrücken der Formen $x_i = x_j$ und $R_i(\mathrm{x}_j)$ aufgebaut sind. Eventuelle Mehrsortigkeit kann durch Einführung von Sortenprädikaten (vgl. 3.2.) auf den einsortigen Fall zurückgeführt werden. Konstantensymbole c können durch zusätzliche einstellige Relationssymbole $\mathrm{R_c}$ eliminiert werden. Statt H(c) ist dann der semantisch äquivalente Ausdruck

$$\bigvee !!\mathrm{x}\mathrm{R_c}(\mathrm{x}) \wedge \bigvee \mathrm{x}\big(\mathrm{R_c}(\mathrm{x}) \wedge \mathrm{H}(\mathrm{x})\big)$$

zu betrachten. Daher gilt die im folgenden behauptete Entscheidbarkeit der Allgemeingültigkeit im Prinzip für Sprachen S, die durch Basen der Form $(n; \boldsymbol{B}_r, \emptyset, \boldsymbol{B}_c)$ definiert werden, wobei $\boldsymbol{B}_r$ nur aus einstelligen Relationssymbolen besteht.

Zu jedem Ausdruck H kann man durch gebundene Umbenennung und Generalisierung vollfreier Variablen einen Ausdruck $\overline{\mathrm{H}}$ konstruieren, der abgeschlossen und bei einer beliebigen Interpretation ω genau dann wahr ist, wenn H bei ω allgemeingültig ist. Daher können wir uns auf die Untersuchung der Allgemeingültigkeit abgeschlossener Ausdrücke beschränken. Ein abgeschlossener Ausdruck heißt *erfüllbar*, wenn er bei wenigstens einer Interpretation wahr wird. Er ist daher genau dann erfüllbar, wenn seine Negation nicht allgemeingültig ist. Damit folgt

Satz 1. *Für eine elementare Sprache $\boldsymbol{S}$ ist die Menge $Fl_{\boldsymbol{S}}(\emptyset)$ $(= ag_{\boldsymbol{S}})$ genau dann entscheidbar, wenn die Menge der erfüllbaren abgeschlossenen Ausdrücke von $\boldsymbol{S}$ entscheidbar ist.*

Es seien $\mathrm{R}_1, \ldots, \mathrm{R}_m$ einstellige Relationssymbole. Eine *Elementarkonjunktion in* $\mathrm{R}_1, \ldots, \mathrm{R}_m$ ist ein Ausdruck der Form

$$(\neg)\, \mathrm{R}_1(\mathrm{x}) \wedge (\neg)\, \mathrm{R}_2(\mathrm{x}) \wedge \ldots \wedge (\neg)\, \mathrm{R}_m(\mathrm{x}),$$

wobei die eingeklammerten Negationszeichen $(\neg)$ andeuten sollen, daß das i-te Konjunktionsglied unabhängig von den anderen gleich $\mathrm{R}_i(\mathrm{x})$ oder gleich $\neg\, \mathrm{R}_i(\mathrm{x})$ sein kann (vgl. die Definition der aussagenlogischen Elementarkonjunktionen in 2.3.). Demnach gibt es 2^m Elementarkonjunktionen $\mathrm{K}_\mu(\mathrm{x})$ $(\mu = 1, \ldots, 2^m)$ in $\mathrm{R}_1, \ldots, \mathrm{R}_m$.

Eine *Normalkonjunktion bezüglich* $R_1, \ldots, R_m$ ist ein abgeschlossener Ausdruck der Form

$$\bigvee_{k_1} xK_{\mu_1}(x) \wedge \ldots \wedge \bigvee_{k_r} xK_{\mu_r}(x) \wedge \neg \bigvee_{l_1} xK_{\nu_1}(x) \wedge \ldots \wedge \neg \bigvee_{l_s} xK_{\nu_s}(x), \tag{1}$$

wobei $K_{\mu_1}, \ldots, K_{\mu_r}$ und $K_{\nu_1}, \ldots, K_{\nu_s}$ jeweils paarweise verschiedene Elementarkonjunktionen in $R_1, \ldots, R_m$ sind. Eine *kontrapränexe Normalform bezüglich* $R_1, \ldots, R_m$ (hier im folgenden kurz als Nf bezeichnet) ist eine Alternative von endlich vielen Normalkonjunktionen der Form (1).

Satz 2. *Es gibt ein algorithmisches Verfahren, durch das man jeden abgeschlossenen Ausdruck* H, *in dem keine Basissymbole außer* $R_1, \ldots, R_m$ *vorkommen, in eine zu ihm semantisch äquivalente* Nf *überführen kann.*

Das zum Beweis von Satz 2 anzugebende Verfahren ist recht kompliziert. Wir verweisen auf [3, § 9], wo es ausführlich dargestellt wird.

Eine beliebige Interpretation ω der durch die Basissymbole $R_1, \ldots, R_m$ gegebenen Sprache wird durch eine nichtleere Grundmenge M und m (eventuell zum Teil leere) Teilmengen

$$M_i := \{\xi : \xi \in M \wedge \xi \in \omega(R_i)\} \qquad (i = 1, \ldots, m)$$

definiert. Sind M und M_i vorgegeben, so kann man M so in 2^m paarweise disjunkte Mengen $E_{i_1,\ldots,i_m}$ ($i_1, \ldots, i_m \in \{0, 1\}$) zerlegen, daß jede der Mengen M_i als Vereinigung von 2^{m-1} dieser Bausteinmengen $E_{i_1,\ldots,i_m}$ darstellbar ist. Wir definieren induktiv:

$$E_{i_1} := \begin{cases} M_1, & \textit{falls } i_1 = 1, \\ M \setminus M_1, & \textit{falls } i_1 = 0; \end{cases}$$

ist schon $E_{i_1,\ldots,i_k}$ für ein $k < m$ definiert, so sei

$$E_{i_1,\ldots,i_k,i_{k+1}} := \begin{cases} E_{i_1,\ldots,i_k} \cap M_{k+1}, & \textit{falls } i_{k+1} = 1, \\ E_{i_1,\ldots,i_k} \setminus M_{k+1}, & \textit{falls } i_{k+1} = 0. \end{cases}$$

Insbesondere ist also

$$E_{\underbrace{1,\ldots,1}_{m}} = M_1 \cap M_2 \cap \ldots \cap M_m,$$

$$E_{1,0,1,\ldots,1} = (M_1 \cap M_3 \cap \ldots \cap M_m) \setminus M_2$$

usw. und

$$M_j = \bigcup_{i_j=1} E_{i_1,\ldots,i_m}.$$

Man sieht nun, daß jede der 2^m Elementarkonjunktionen gerade die Lage des mit x bezeichneten Dinges in einer ganz bestimmten der Mengen $E_{i_1,\ldots,i_m}$ aussagt und daß daher jede Normalkonjunktion für gewisse dieser Mengen eine endliche Mindestzahl von Elementen und für gewisse (zum Teil andere) dieser Mengen eine endliche Höchstzahl von Elementen aussagt. Da umgekehrt durch die Wahl von 2^m paarweise disjunkten Mengen $E_{i_1,\ldots,i_m}$ mit der Nebenbedingung, daß ihre Vereinigung (d. h. M) nicht leer sein soll, jeweils eine Interpretation ω eindeutig bzw. durch Vorgabe der Mächtigkeiten der Mengen $E_{i_1,\ldots,i_m}$ eine Interpretation ω bis auf Isomorphie eindeutig bestimmt ist, folgt sofort

Satz 3. *Eine Normalkonjunktion* (1) *ist genau dann erfüllbar, wenn sie nicht die Form*

$$\neg \bigvee \mathrm{x}\mathrm{K}_1(\mathrm{x}) \wedge \neg \bigvee \mathrm{x}\mathrm{K}_2(\mathrm{x}) \wedge \ldots \wedge \neg \bigvee \mathrm{x}\mathrm{K}_{2^m}(\mathrm{x})$$

hat und wenn außerdem aus $\mathrm{K}_{\mu_i} = \mathrm{K}_{\nu_j}$ *stets* $k_i < l_j$ *folgt.*

Da eine Alternative genau dann erfüllbar ist, wenn wenigstens ein Alternativglied erfüllbar ist, ist folglich die Erfüllbarkeit einer Nf in einfacher Weise entscheidbar.

Aus den Sätzen 2 und 3 folgt sofort die Entscheidbarkeit der Allgemeingültigkeit für alle Ausdrücke der betrachteten Art. Ferner ergibt sich aus Satz 2 nachträglich, daß man in prädikatenlogischen Sprachen der hier betrachteten Art nur sehr einfache Sachverhalte über die möglichen Interpretationen formulieren kann, nämlich nur gewisse endliche Alternativen über die endlichen Mindest- und Höchstzahlen von Mengen $E_{i_1,\ldots,i_m}$. Es ist also z. B. nicht möglich, in einer solchen Sprache Aussagen des Inhalts

— der Grundbereich ist endlich

— die Anzahl der Elemente des Grundbereichs ist eine Primzahl

zu formulieren. Weiterhin folgt aus Satz 3: Jeder überhaupt erfüllbare Ausdruck der betrachteten Art ist schon in einem endlichen Grundbereich erfüllbar.

Wir wollen nun für den Fall, daß in den betrachteten Ausdrücken das Zeichen „=" nicht vorkommt, ein zweites Entscheidungsverfahren für die Allgemeingültigkeit behandeln, das sich wesentlich einfacher theoretisch rechtfertigen läßt und auch in der praktischen Anwendung häufig schneller zum Ziel führt als das oben beschriebene Verfahren über die Aufstellung einer äquivalenten Nf.

Der (nicht notwendig abgeschlossene) Ausdruck H sei aus prädikativen Ausdrücken der Form $\mathrm{R}_i(\mathrm{x}_j)$ aufgebaut, wobei insgesamt genau m verschiedene Relationssymbole $\mathrm{R}_1, \ldots, \mathrm{R}_m$ vorkommen. Ist ω eine beliebige Interpretation mit dem Grundbereich M, so sei für $\xi, \eta \in M$

$$\xi \simeq \eta :\leftrightarrow \quad \textit{Für } i = 1, \ldots, m \textit{ ist } \xi \in \omega(\mathrm{R}_i) \textit{ genau dann, wenn } \eta \in \omega(\mathrm{R}_i). \quad (2)$$

Wie man leicht nachprüft, ist „$\simeq$" eine Äquivalenzrelation in M, deren Äquivalenzklassen übrigens gerade die nichtleeren unter den vorher betrachteten zu ω gehörigen Mengen $E_{i_1,\ldots,i_m}$ sind. Folglich ist ihre Anzahl $\leq 2^m$. Es sei nun $\bar{\xi}$ die Äquivalenz-klasse von ξ für $\xi \in M$ und $\overline{M}$ die Menge dieser Äquivalenzklassen. Im Grundbereich $\overline{M}$ definieren wir eine neue Interpretation $\overline{\omega}$ durch

$$\overline{\omega}(\mathrm{R}_i) := \{\overline{\xi} : \xi \in \omega(\mathrm{R}_i)\} \qquad (i = 1, \ldots, m).$$

Diese Definition ist wegen (2) repräsentantenunabhängig. Setzen wir noch $\overline{f}(x) := \overline{f(x)}$ für Belegungen f bezüglich ω, so läßt sich durch Induktion über die Kompliziertheit von H leicht zeigen:

$$Wert(\mathrm{H}, \omega, f) = Wert(\mathrm{H}, \overline{\omega}, \overline{f}).$$

Ist daher ω eine Interpretation und f eine Belegung bezüglich ω, so daß $Wert(\mathrm{H}, \omega, f) = \boldsymbol{F}$ ist, so ist auch $Wert(\mathrm{H}, \overline{\omega}, \overline{f}) = \boldsymbol{F}$, d. h., jeder nicht allgemeingültige Ausdruck (der hier betrachteten Art) ist schon bei einer solchen Interpretation nicht allgemeingültig, bei der jede der 2^m Mengen $E_{i_1,\ldots,i_m}$ unabhängig von den anderen leer ist oder aus genau einem Element besteht (wobei der Fall, daß alle diese Mengen leer sind, ausgeschlossen wird). Das sind aber nur $2^{2^m} - 1$ paarweise nichtisomorphe Interpretationen, von denen keine aus mehr als 2^m Dingen besteht. Das sich hieraus ergebende Entscheidungsverfahren für die Allgemeingültigkeit identitätsfreier Ausdrücke ist insbesondere für kleine Zahlen m und relativ lange (komplizierte) Ausdrücke günstiger als das natürlich auch hier anwendbare Verfahren über die Herstellung der kontrapränexen Nf, da die Gewinnung dieser Nf natürlich mit wachsender Länge der Ausdrücke beliebig kompliziert wird.

8.4. Aufzählbarkeit und Axiomatisierbarkeit, Gödelscher Unvollständigkeitssatz

Es liegt nahe, als Ersatz für die in vielen Fällen nicht gegebene Entscheidbarkeit formalisierter Theorien deren Aufzählbarkeit zu verlangen. Hierzu stellen wir einleitend fest, daß die in 6.1., Definitionen 1 und 2, erhobenen Forderungen an Schlußregeln und Beweiskalküle sich nachträglich wie folgt präzisieren lassen:

Eine Schlußregel ist eine berechenbare Wortfunktion mit entscheidbarem Definitionsbereich. Ein Beweiskalkül ist eine aufzählbare Menge von Schlußregeln (genauer: eine aufzählbare Menge von Niederschriften von Algorithmen zur Berechnung der betreffenden Wortfunktionen). Aus der Charakterisierung der Schlußregeln als berechenbare Wortfunktionen folgt nach 7.3., Satz 1, sofort: *Jedes durch eine Schlußregel gegebene Schema logischer Axiome ist eine aufzählbare Menge von Ausdrücken.* Weiter folgt aus 7.3., Satz 5:

Satz 1. *Ist (S, σ) eine nichtelementare (eventuell elementare) Sprache, $X \subseteq S$ ein aufzählbares Axiomensystem und $\mathscr{X}$ eine aufzählbare Menge von für (S, σ) zulässigen Schlußregeln (d. h. ein für (S, σ) zulässiger Beweiskalkül), so ist $Bew_{\mathscr{X}}(X)$ aufzählbar.*

Folgerung. *Ist S eine elementare Sprache und $X \subseteqq S$ aufzählbar, so ist die elementare Theorie $Fl_S(X)$ aufzählbar.* (Dies folgt unmittelbar aus Satz 1 und dem Hauptsatz der mathematischen Logik.)

Es liegt nun sehr nahe, den Begriff *Axiomensystem*, der bisher (insbesondere in 4.3. bis 4.6.) im Sinne einer beliebigen Teilmenge X einer Sprache S gebraucht wurde, auf aufzählbare Mengen X von Ausdrücken einzuschränken. Dadurch werden die in konkreten mathematischen Theorien benutzten Axiomensysteme nicht berührt, da diese (falls sie nicht sogar endlich und daher nach 7.3., Satz 3, aufzählbar sind) ja immer durch irgendwelche Schemata, Erzeugungsregeln o. ä. angegeben werden, was gerade ihre Aufzählbarkeit bewirkt. Insbesondere definieren wir nun:

Definition 1. Eine in einer Sprache (S, σ) gegebene Theorie T (d. h. eine Teilmenge T von S mit $Fl_S{}^\sigma(T) = T$) heißt *axiomatisierbar*, wenn ein Axiomensystem, d. h. eine aufzählbare Menge $X \subseteqq S$ mit $T = Fl_S{}^\sigma(X)$, existiert. Insbesondere heißt T *endlich axiomatisierbar*, wenn eine endliche Menge $E \subseteqq S$ mit $T = Fl_S{}^\sigma(E)$ existiert.

Satz 2. *Eine elementare Theorie ist genau dann axiomatisierbar, wenn sie aufzählbar ist.*

Die Aufzählbarkeit axiomatisierbarer elementarer Theorien folgt dabei, wie bereits bemerkt, aus Satz 1. Ist andererseits eine elementare Theorie $T = Fl_S(T)$ aufzählbar, so kann man offenbar im Sinne der präzisierten Definition T selbst als Axiomensystem für T nehmen.

An Definition 1 knüpft sich natürlich die Frage, ob und wie man denn im konkreten Fall überhaupt eine Theorie definieren kann, deren Axiomatisierbarkeit zunächst noch offen ist. Dazu ein Beispiel: Es sei $(S, \sigma_{(2)})$ die in 4.5. zur Formulierung der Axiome von Peano benutzte nichtelementare Sprache, E das endliche in S formulierte Axiomensystem von Peano, demnach $T = Fl_S{}^{\sigma_{(2)}}(E)$ die Menge aller in S formulierbaren und im Bereich der natürlichen Zahlen allgemeingültigen Ausdrücke. Diese Menge T ist offenbar erst recht bezüglich elementarer Folgerungen abgeschlossen, d. h., es kann $T = Fl_S(T)$ auch als elementare Theorie aufgefaßt und im Rahmen der elementaren Sprache S die Frage nach der Axiomatisierbarkeit von T gestellt werden. Diese Frage werden wir später negativ beantworten. Einstweilen stellen wir in Anlehnung an dieses Beispiel fest:

Theorien, d. h. Satzmengen T einer formalisierten Sprache, die bezüglich eines gewissen (im allgemeinen nichtelementaren) Folgerungsbegriffes abgeschlossen sind, d. h. für die $T = Fl_S{}^\sigma(T)$ gilt, können auf mannigfache Weise (etwa als Menge aller bei einer bestimmten Interpretation gültigen Sätze) semantisch definiert werden, so daß die Frage ihrer Aufzählbarkeit und ihrer Axiomatisierbarkeit zunächst offenbleibt. Demgegenüber bezeichnet man Satzmengen T, die durch ein (aufzählbares) Axiomensystem X und einen (nicht notwendig aus nur elementaren Schlußregeln bestehenden) Beweiskalkül $\mathscr{K}$ als $\boldsymbol{T} = Bew_{\mathscr{K}}(\boldsymbol{X})$ definiert sind, als *Theorien mit syntaktisch definierter Satzmenge*. Es sei bemerkt, daß extreme Formalisten unter den Grundlagenmathematikern nur den letzteren Standpunkt anerkennen. Theorien mit syntaktisch definierter Satzmenge sind auf Grund von Satz 1 stets aufzählbar. Der Hauptsatz der mathematischen Logik läßt sich unter den neuen Aspekten auch so interpretieren, daß zumindest für elementare Sprachen (bzw. für das elementare Folgern) jede axiomatisierbare Theorie mit semantisch definierter Satzmenge auch als Theorie mit syntaktisch definierter Satzmenge aufgefaßt werden kann.

Satz 3. *Jede (elementare oder nichtelementare) aufzählbare und vollständige Theorie ist entscheidbar.*

Beweis. Es sei $\boldsymbol{T} = Fl_{\boldsymbol{S}}{}^{\sigma}(\boldsymbol{X})$ ($\boldsymbol{X} \subseteqq \boldsymbol{S}$ beliebig) vollständig, d. h., für abgeschlossene Ausdrücke $\mathrm{H} \in \bar{\boldsymbol{S}}$ gilt $\mathrm{H} \in \boldsymbol{T}$ oder $\neg\, \mathrm{H} \in \boldsymbol{T}$. Falls $\boldsymbol{T}$ σ-widerspruchsvoll ist, d. h. $\boldsymbol{T} = \boldsymbol{S}$, ist $\boldsymbol{T}$ entscheidbar nach 8.1. Andernfalls gilt für $\mathrm{H} \in \bar{\boldsymbol{S}}$ genau eine der Beziehungen $\mathrm{H} \in \boldsymbol{T}$ und $\neg\, \mathrm{H} \in \boldsymbol{T}$. Ist daher $\boldsymbol{T}$ aufzählbar und f eine Aufzählung von $\boldsymbol{T}$, d. h. eine berechenbare Abbildung von der Menge der natürlichen Zahlen auf $\boldsymbol{T}$, so gewinnt man ein Entscheidungsverfahren für $\boldsymbol{T}$, indem man, ausgehend von einem beliebigen Ausdruck $\mathrm{H} \in \boldsymbol{S}$, zunächst durch gebundene Umbenennungen und Generalisierungen zu einem abgeschlossenen Ausdruck $\bar{\mathrm{H}}$ übergeht und dann gemäß f so lange $\boldsymbol{T}$ aufzählt, bis entweder $\bar{\mathrm{H}}$ oder $\neg\, \bar{\mathrm{H}}$ erreicht wird.

Auf elementare Theorien spezialisiert, bedeutet Satz 3:

Eine axiomatisierbare und vollständige Theorie ist entscheidbar. Hieraus erhält man durch einfache aussagenlogische Umstellung:

a) *Eine* (etwa auf Grund ihrer Ausdrucksfähigkeit, vgl. 8.2.) *nicht entscheidbare axiomatisierbare Theorie ist nicht vollständig.*

b) *Eine nicht entscheidbare vollständige Theorie ist nicht axiomatisierbar.*

Es zeigt sich, daß die in 4.5. erklärte elementare Nachfolgertheorie der natürlichen Zahlen vollständig und entscheidbar ist. Presburger[1]) zeigte, daß dieser Sachverhalt sogar für diejenige elementare Theorie der natürlichen Zahlen zutrifft, die man erhält, wenn man die Sprache durch ein Symbol $+$ für die Addition und das Axiomensystem {(b′), (c′), (d′), (e′)} (vgl. S. 78) durch die bekannte induktive Charakterisierung

$$\mathrm{n} + 0 = \mathrm{n},$$

$$\mathrm{n} + \mathrm{m}' = (\mathrm{n} + \mathrm{m})'$$

[1]) M. Presburger, Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt, Comptes rendus du I Congres des Mathematiciens des Pays Slaves, Warszawa 1930.

der Addition erweitert. Nimmt man jedoch weiterhin die Multiplikation zu den Grundbegriffen und die Ausdrücke

$$\mathrm{n} \cdot 0 = 0,$$

$$\mathrm{n} \cdot \mathrm{m}' = (\mathrm{n} \cdot \mathrm{m}) + \mathrm{n}$$

zum Axiomensystem hinzu, so wird die so erhaltene elementare Theorie der natürlichen Zahlen ausdrucksfähig und folglich unentscheidbar. Durch Spezialisierung von Satz 3 auf diesen Fall erhält man den folgenden Satz:

Gödelscher Unvollständigkeitssatz. *Jedes aufzählbare Axiomensystem für eine elementare Theorie der natürlichen Zahlen, in der Addition und Multiplikation zu den Grundbegriffen gehören bzw. definierbar sind, ist unvollständig.*

Bekanntlich gibt es eine Reihe von elementar formulierbaren Aussagen über natürliche Zahlen, die bis heute weder bewiesen noch widerlegt werden konnten. Hierzu gehören einige berühmte zahlentheoretische Probleme wie z. B. die Fermatsche Vermutung, daß es für $n \geqq 3$ keine natürlichen Zahlen $a, b, c > 0$ mit $a^n + b^n = c^n$ gibt, oder die Goldbachsche Vermutung, daß jede natürliche Zahl $n \geqq 6$ sich als Summe dreier Primzahlen darstellen läßt. Da die Gesamtheit $\boldsymbol{X}$ aller bisher jemals bei zahlentheoretischen Beweisen verwendeten bzw. zugelassenen Axiome gewiß aufzählbar ist und daher nach dem Gödelschen Unvollständigkeitssatz die Theorie $Fl(\boldsymbol{X})$ unvollständig sein muß, ist es durchaus möglich, daß z. B. gerade die Fermatsche Vermutung zu denjenigen abgeschlossenen Ausdrücken H gehört, die mittels $\boldsymbol{X}$ weder bewiesen noch widerlegt werden können, d. h. für die weder $\mathrm{H} \in Fl(\boldsymbol{X})$ noch $\neg\, \mathrm{H} \in Fl(\boldsymbol{X})$ gilt. In diesem Fall ist es nicht ausgeschlossen, daß man eines Tages durch irgendwelche metatheoretischen oder bis heute unbekannten Beweismethoden hinreichende Gewißheit über die Gültigkeit von H im Bereich der natürlichen Zahlen gewinnt. Dann könnte man H einfach den bisher benutzten Axiomen hinzufügen, aber auch das Axiomensystem $\boldsymbol{X} \cup \{\mathrm{H}\}$ wird, da immer noch aufzählbar, wieder unvollständig sein. Allgemein wird zu jedem Zeitpunkt die Menge $Bew_{\mathscr{K}}(\boldsymbol{Y})$ aller zahlentheoretischen Aussagen, die man aus einem Axiomensystem $\boldsymbol{Y}$ mittels eines zulässigen (nicht notwendig elementaren) Kalküls $\mathscr{K}$ beweisen kann, unvollständig sein.

Abschließend sei bemerkt, daß völlig analoge Verhältnisse auch auf jede mittels aufzählbarer Axiomensysteme und Beweiskalküle betriebene formalisierte Mengenlehre (wegen der Ausdrucksfähigkeit dieser Theorien) zutreffen. Hier sind Auswahlaxiom und verallgemeinerte Kontinuumhypothese erste Beispiele für elementar formulierbare Aussagen, die mit den herkömmlichen Axiomen (heute schon nachweisbar) weder bewiesen noch widerlegt werden können. Mag man sich nun entschließen, das bisher benutzte Axiomensystem durch die als unabhängig und widerspruchsfrei erkannten Aussagen oder deren Negation zu verstärken, man wird immer wieder auf Sachverhalte stoßen, deren Gültigkeit auch mit dem neuen Axiomensystem weder beweisbar noch widerlegbar ist.

9. Historische Entwicklung und gegenwärtige Tendenzen der Grundlagen der Mathematik

Die Mathematik entstand in den Anfängen der menschlichen Gesellschaft als ein Instrument zur Bewältigung von Problemen des täglichen Lebens. Demgemäß bestand sie zunächst in der Anhäufung einfacher arithmetischer und geometrischer Begriffe, einfacher Sachverhalte und der rezeptartigen Lösung einfacher Aufgaben. Als erste Beweisform trat in der spätbabylonischen Mathematik die rechnerische Probe numerischer Lösungen von einfachen arithmetischen und algebraischen Aufgaben auf. Etwa in der Zeit zwischen 700 und 400 v. u. Z. verwandelte sich die Mathematik in den griechischen Stadtstaaten in engem Zusammenhang mit der dort entstehenden materialistischen ionischen Naturphilosophie in eine allgemeine Sätze formulierende und beweisende Wissenschaft. Für die weitere Entwicklung wurde zunächst der Einfluß der platonischen Philosophie bestimmend. Platon (429?—348 v. u. Z.) und seine Anhänger negierten die historische Entstehung der mathematischen Begriffe durch Abstraktion aus der Realität, sprachen ihnen vielmehr eine von der materiellen Welt und vom menschlichen Bewußtsein unabhängige metaphysische Existenz zu und degradierten die materiellen Vorbilder der Begriffe zu unvollkommenen und unzuverlässigen „nachträglichen" Hilfsmitteln der Veranschaulichung. Sichere Erkenntnisse konnten nach Platon niemals durch Erfahrung und Erprobung, sondern nur durch reines Denken gewonnen werden.

Als typisches Produkt dieser philosophischen Lehren entstand um 300 v. u. Z. das älteste uns erhaltene Lehrbuch der Mathematik, die *Elemente* des Euklid, zugleich der erste Versuch, eine Wissenschaft axiomatisch-deduktiv aufzubauen. Die vieldiskutierte Frage nach der Rolle der Definitionen, Axiome und Postulate bei Euklid und nach dem Charakter der logischen Beweisführung in einem Stadium, in dem die wahre Natur des Folgerns noch unklar war, beantwortet sich vor dem

Hintergrund der platonischen Philosophie fast von selbst: Es gibt nur eine Interpretation der mathematischen Begriffe Punkt, Gerade, Kreis, rechter Winkel usw., nämlich durch die von Anbeginn und unabhängig von Realität und Bewußtsein existierenden Ideen dieser Begriffe. Daher stehen auch alle Eigenschaften dieser Begriffe und alle Beziehungen zwischen ihnen von vornherein fest und sind sozusagen denknotwendig. Es liegt nur an der Unvollkommenheit des menschlichen Verstandes, daß er nicht alle diese Tatsachen sofort und ohne Unterweisung erkennen kann. Die einfachsten unter ihnen sind jedoch ohne weiteres „evident", und alle weiteren lassen sich nach und nach durch richtiges Denken enthüllen. Dabei wurde übersehen, daß die dem „richtigen Denken" zugrunde liegenden Normen in Wirklichkeit ebenfalls durch Abstraktion aus praktischen Erfahrungen entstanden sind. Vom platonischen Standpunkt aus ist es natürlich keineswegs beliebig, welche Sätze als Axiome genommen werden. Aber das Kriterium für ihre Auswahl ist nicht ihre durch Erfahrung bestätigte (approximative) Gültigkeit in der Realität, sondern ihre „Evidenz". So erklärt sich auch die Kritik der antiken Mathematiker am V. Postulat, das in der euklidischen Formulierung eben nicht im gleichen Maße wie die übrigen Axiome und Postulate „evident" war. Daher wurde mehrfach der Verdacht geäußert, Euklid habe diesen Satz nur deshalb als Postulat bezeichnet, weil es ihm nicht gelungen war, ihn zu beweisen (d. h. im Sinne der damals herrschenden Auffassung: ihn auf „evidentere" Aussagen zurückzuführen). Von diesem Standpunkt aus sind die zahlreichen aus unserer Sicht nicht stichhaltigen „Beweise" des V. Postulats völlig legitim, insofern sie seine Gültigkeit tatsächlich aus der Gültigkeit subjektiv „evidenterer" Sachverhalte begründen. Zu den euklidischen Definitionen (*Ein Punkt ist, was keinen Teil hat* usw.), die natürlich keine Definitionen im Sinne der mathematischen Logik sind, sei bemerkt, daß sie im Grunde auch vom platonischen Standpunkt aus überflüssig sind, da die Begriffe der Geometrie ja in gleicher Weise denknotwendig sind wie die Sätze. Euklid selbst mag diese Definitionen, die im Gegensatz zum größten Teil des mathematischen Inhalts der *Elemente* wahrscheinlich von ihm stammen, aber natürlich nirgends benutzt werden, als literarischen Schmuck empfunden haben. Spätere Generationen haben, da es am eigentlichen mathematischen Aufbau der *Elemente* wenig zu verbessern gab, mit Vorliebe an diesen Definitionen gefeilt.

Die platonische Auffassung vom Wesen der mathematischen Begriffe und des mathematischen Beweises blieb schon in der Antike nicht unwiderlegt, obwohl sie damals dominierte, da der Platonismus zugleich eine ideologische Stütze der Sklavenhaltergesellschaft war. Demokrit von Abdera (460?—370? v. u. Z.), Archimedes von Syracus (287?—212 v. u. Z.), Heron von Alexandria (um 100 u. Z.) und andere antike Mathematiker vertraten mehr oder weniger ausgesprochen materialistische Auffassungen, die sich nicht nur fruchtbar auf ihr mathematisches Schaffen, sondern auch auf ihre Position zur Frage der Anwendung der Mathematik in der Praxis auswirkten. Andererseits spielt der Platonismus speziell in der Mathematik noch bis in die Gegenwart hinein eine gewisse Rolle. Auf den ersten Blick mag es harmlos erscheinen, wenn der Mathematiker bei seiner Arbeit von den von ihm untersuchten Objekten (insbesondere etwa Mengen, reelle Zahlen o. ä.) eine im wesentlichen dem Platonismus entsprechende Vorstellung hegt. Schließlich sind weite Teile klassischer und noch heute gültiger Mathematik auf dieser erkenntnistheoretischen Grundlage entstanden. Philosophische Auffassungen innerhalb der Mathematik werden jedoch nicht unabhängig von den allgemeinen weltanschaulichen Positionen ihrer Träger gebildet. Letzten Endes führt die Leugnung des Widerspiegelungscharakters der mathematischen Begriffe und allgemein das Aufgeben materialistischer Positionen in der Mathematik immer wieder zu einer Entfremdung von der Praxis, zur Verkennung ihrer Rolle in der Gesellschaft und ihres humanistischen Anliegens.

Zugleich mit der axiomatischen Methode entstanden die Anfänge einer formalen Logik als Anleitung zum richtigen Schlußfolgern. ARISTOTELES (384—322 v. u. Z.) gab 19 zulässige Schlußregeln an, im Bereich von Aussagen der vier Formen

(a) Alle A sind B (z. B.: Alle Fische sind Tiere),

(i) Einige A sind B (z. B.: Einige Rechtecke sind Quadrate),

(e) Kein A ist ein B (z. B.: Kein Mensch ist ein Fisch),

(o) Einige A sind nicht B (z. B.: Einige Fische sind keine Fleischfresser),

aus jeweils zwei Prämissen auf eine Conclusio zu schließen. Zu den Aristotelischen Schlußweisen gehören also z. B.

(1) Alle A sind B.
Alle B sind C.
(Folglich) Alle A sind C.

(2) Einige A sind B.
Alle B sind C.
(Folglich) Einige A sind C.

Später hat man für diese Schlußweisen dreisilbige Merkwörter eingeführt, welche die der ersten und zweiten Prämisse und der Conclusio zugeordneten Vokale (a, e, i, o) angeben, so z. B. *barbara* für (1) und *dimatis* für (2). Die Aristotelische Regellogik wird als *Syllogistik* bezeichnet. Die Stoiker (etwa 300—200 v. u. Z.) formulierten dem Sinne nach die moderne Definition der Implikation als Wahrheitsfunktion und gaben eine Reihe von allgemeingültigen aussagenlogischen Schemata im Zusammenhang mit der Implikation an, wie z. B. (in moderner Schreibweise)

$$p \to p \quad \text{und} \quad (p \to q) \to \big((q \to r) \to (p \to r)\big).$$

Diese Fragmente einer Aussagenlogik bildeten zusammen mit der Syllogistik einen Teil derjenigen Errungenschaften der antiken Wissenschaft, die auch während der Zeit des europäischen Feudalismus bewahrt und gepflegt wurden, insbesondere im Rahmen der Scholastik des 12. bis 14. Jahrhunderts. Bis zum Ende des 18. Jahrhunderts wurden jedoch insgesamt nur geringe Fortschritte in der formalen Logik erzielt, und diese wurde meist, losgelöst von der Mathematik, als Teil der Philosophie betrieben. Als Vertreter der mittelalterlichen Logik seien hier die Briten Duns Scotus (1270?—1308) und Roger Bacon (1214—1294) und die Spanier Raimundus Lullus (1234—1315) und Petrus Hispanus (1210—1277) genannt. Letzterer formulierte u. a. die logische Äquivalenz von $\neg \bigwedge x H(x)$ und $\bigvee x \neg H(x)$ und ähnliche Gesetze, freilich nur in Gestalt von Sätzen wie *Non omnis, quidam non* (d. h., nicht für alle [bedeutet so viel wie] für ein gewisses nicht). Bei den meisten dieser mittelalterlichen Philosophen war der rationale Kern der formalen Logik tief verborgen unter einem Wust mystischer und religiöser Spekulationen. (So konstruierte Lullus eine „logische Maschine" in der Absicht, auf diesem Wege die Existenz Gottes zu beweisen.) Es gab jedoch, vor allem bei Bacon, auch erstaunliche Ansätze spontan materialistischen Denkens, z. B. richtige Erkenntnisse über die Entstehung logischer Schlußweisen durch Verallgemeinerung von Erfahrungen.

In allen traditionellen Zweigen der Mathematik wurde die Umgangssprache in der Formulierung von Sätzen und Beweisen und sogar in der Durchführung von Rechnungen erst spät durch die Ansätze einer Formalisierung, d. h. durch fachspezifische Wörter und Symbole für Objekte, Operationen und Relationen und den Gebrauch von Variablen zurückgedrängt. Wo dies geschah, bildeten sich Kalküle und Algorithmen. Die mit den neuen Methoden relativ schnell erzielten großen praktischen Fortschritte waren aber zunächst oft mit einem Mangel an begrifflicher Klarheit verbunden. Als markante Beispiele seien die Entwicklung des logarithmischen Rechnens und die Entstehung der analytischen Geometrie genannt. Im 17. Jahrhundert entstand mit der Differential- und Integralrechnung zum ersten Mal seit Jahrhunderten eine neue mathematische Disziplin, in der der Kalkül von Anfang an die Hauptrolle spielte und der inhaltlichen Bewältigung der neuen Begriffe weit vorauseilte. Unter dem Eindruck der im numerischen Rechnen, in der Algebra, analytischen Geometrie und Analysis rasch erzielten großen Fortschritte konnte der Glaube an die Möglichkeit immer umfassenderer und leistungsfähigerer mathematischer Kalküle entstehen. Gottfried Wilhelm Leibniz (1646—1716), der selbst in der Erfindung zweckmäßiger Symbole und Algorithmen viel Bleibendes geleistet hat, widmete mehrere Arbeiten dem Projekt einer universellen formalisierten Sprache, in der man jeden Sachverhalt formulieren und über seine Wahrheit durch ein rechnerisches Verfahren entscheiden können sollte. Diese unausgeführten (und natürlich in dieser Allgemeinheit undurchführbaren) Projekte von Leibniz wurden erst zu Beginn dieses Jahrhunderts wiederentdeckt, so daß Leibniz nicht zu den direkten Vorläufern der sich im 19. Jahrhundert sprunghaft entwickelnden modernen Logik gezählt werden kann.

Den Grundstein für die Entwicklung der modernen Logik legte George Boole (1815—1864) mit seiner 1847 erschienenen Arbeit *The Mathematical Analysis of Logic, Being an Essay Towards a Calculus of Deductive Reasoning.* Hierin entwickelte er den heute noch aktuellen Begriff der Booleschen Algebra als einer Struktur mit Addition, Subtraktion, Multiplikation und zwei ausgezeichneten Elementen 0 bzw. 1. Die Axiome Booles erweisen sich als aussagenlogisch allgemeingültige

Ausdrücke, wenn man 0 bzw. 1 als Wahrheitswerte *falsch* bzw. *wahr*, die Addition $x + y$ als *entweder x oder y*, die Substraktion $x - y$ als *x und nicht y* und die Multiplikation xy als *x und y* deutet. Interpretiert man andererseits 1 als festen Grundbereich, 0 als leere Menge, die Variablen als Variablen für beliebige Teilmengen des Grundbereichs, die Summe als symmetrische Differenz, die Subtraktion als Mengendifferenz und die Multiplikation als Durchschnitt, so erhält man die aus den Anfängen der Mengenlehre wohlbekannte Boolesche Mengenalgebra. BOOLES in weiteren Arbeiten ausgebaute Theorie gipfelte in der Ermittlung der zu einer vorgegebenen Formel äquivalenten kanonischen Normalformen (vgl. 2.3.). Die spätere große Ausstrahlungskraft seiner Arbeiten mag z. T. darauf beruhen, daß er die den Zeitgenossen vertraute Symbolik der traditionellen Arithmetik benutzte und die weitgehenden Analogien zwischen den dort und in der Aussagenlogik geltenden Rechengesetzen herausschälte. Gleichzeitig mit BOOLE und unabhängig von ihm begann auch AUGUSTUS DE MORGAN (1806—1871) mit Veröffentlichungen zur Grundlegung der modernen Logik. Er bemühte sich um die Klärung der prädikatenlogischen Natur der Aristotelischen Schlüsse und suchte diese zu verallgemeinern. W. ST. JEVONS (1835—1882) ersetzte die Boolesche Addition durch die Alternative (Vereinigung) und die Subtraktion durch die einstellige Negation (Komplementbildung) und gab damit den Booleschen Normalformen die endgültige (heutige) Form. CHARLES SANDERS PEIRCE (1839—1914) und ERNST SCHRÖDER (1841—1902) bezogen die Quantifizierung in den algebraischen (daher als *Algebra der Logik* bezeichneten) Aufbau der Logik ein, indem sie die Generalisierung als unendliches Produkt und die Partikularisierung als unendliche Alternative deuteten. (Bei SCHRÖDER steht $\prod \mathbf{x}$ statt $\bigwedge \mathbf{x}$ und $\sum \mathbf{x}$ statt $\bigvee \mathbf{x}$.) SCHRÖDER beschäftigte sich erstmals systematisch mit dem allgemeinen Relationsbegriff und führte grundlegende Begriffe der Relationentheorie wie Symmetrie, Transitivität, Äquivalenz- und Ordnungsrelationen, Verkettung von Relationen und inverse Relation ein. Sein Hauptwerk *Vorlesungen über die Algebra der Logik* erschien in drei Bänden in den Jahren 1890 bis 1905. Zu den Pionieren der modernen Logik zählte weiterhin GOTTLOB FREGE (1848—1925), der mit seiner *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* (1879), eine formalisierte prädikatenlogische Sprache entwickelte. FREGE wendete seine Begriffsschrift auch auf konkrete mathematische Theorien an, z. B. in *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet* (1893). Die Begriffsschrift konnte sich jedoch wegen ihrer eigenwilligen und nichtlinearen Schreibweise (vgl. die Beilage zu [35]) nicht durchsetzen, und FREGES spätere Arbeiten wurden durch die in der naiven Mengenlehre aufgetretenen Antinomien stark erschüttert.

GRZEGORCZYK legt im Anhang zu seinem Lehrbuch [11] überzeugend dar, wie die Erkenntnis der prädikatenlogischen Struktur der Sprache (insbesondere der in der Mathematik gebrauchten) und der Ausbau der Prädikatenlogik Hand in Hand gingen mit den Bemühungen um die Präzisierung der Grundbegriffe der Arithmetik und Analysis. Er führt aus, wie in den von AUGUSTIN LOUIS CAUCHY (1789—1857), Karl Weierstrass (1815—1897) und anderen Mathematikern des 19. Jahrhunderts formulierten Definitionen solcher Begriffe wie Konvergenz, Stetigkeit und Differenzierbarkeit aussagenlogische Verknüpfungen und Quantifizierungen von bisher unbekannter Kompliziertheit auftraten, während die vorher in der Mathematik überwiegend gebrauchten Formulierungen keine komplizierte Logik erforderten. In der Tat finden sich unter den Vertretern der neuen Strenge in der Mathematik auch einige bedeutende Logiker. In diesem Zusammenhang seien vor allem Bernard Bolzano (1781—1848) und Giuseppe Peano (1858—1932) genannt. Bolzano wirkte in Prag. Sein Name ist jedem Mathematik-Studenten durch den Satz von Bolzano-Weierstrass geläufig. Dieser Satz ist jedoch nur eine kleine Probe von Bolzanos tiefgründigen Untersuchungen über reelle Zahlen und Zahlenmengen, Konvergenz, Stetigkeit und Differenzierbarkeit. Sein Hauptwerk zur Philosophie und Logik der Mathematik erschien 1827 unter dem Titel *Wissenschaftslehre*. Hierin gab er als erster eine der heutigen Definition annähernd entsprechende Erklärung des Folgerns (vgl. hierzu den in [4] abgedruckten Ausschnitt). Bolzanos Arbeiten blieben leider ohne großen Widerhall, zumal er aus politischen Gründen aus dem Universitätsdienst entlassen wurde und Publikationsverbot erhielt. Peano wirkte in Turin. Von seinen mathematischen Leistungen sei hier neben der axiomatischen Charakterisierung der natürlichen Zahlen (1891), die übrigens fast gleichzeitig analog auch von Richard Dedekind (1831—1916) vorgeschlagen wurde, die Kritik des Kurvenbegriffs (1890) und die von ihm erstmals formulierten Existenzsätze für die Lösbarkeit von Differentialgleichungen erwähnt. Peanos Schaffen ist durch die konsequente Anwendung der mathematischen Logik in der Mathematik und sein Bemühen gekennzeichnet, durch geschickte und manchmal schockierende Beispiele die Notwendigkeit exakter Definitionen und Beweise auch in scheinbar anschaulichen Bereichen der Mathematik zu demonstrieren.

Parallel zur Entstehung der modernen Logik vollzieht sich im 19. Jahrhundert eine andere, für die Grundlagen der Mathematik ebenso wichtige Entwicklung. Als N. I. Lobačevskij (1792—1856) und J. Bolyai (1802—1860) um 1830 die Widerspruchsfreiheit der nichteuklidischen Geometrie behaupten, stoßen sie auf Unverständnis und Ablehnung ihrer Zeitgenossen, zumal der Begriff der Widerspruchsfreiheit einer mathematischen Theorie zu diesem Zeitpunkt überhaupt noch nicht klar ist und sie den exakten Beweis ihrer Behauptung (durch Angabe eines Modells) nicht erbringen können. Die These von der Denknotwendigkeit der euklidischen Geometrie ist gerade im 18. Jahrhundert durch die damals erhebliche Autorität von Immanuel Kant (1724—1804) neu bekräftigt worden. Erst die sich hauptsächlich durch und nach Carl Friedrich Gauss (1777—1855) entwickelnde innere Geometrie gekrümmter Flächen verhilft allmählich der Erkenntnis zum Durchbruch, daß die euklidische Geometrie sozusagen nur ein Ausartungsfall unter vielen in der Natur realisierten Geometrien ist und daß die Frage nach den geometrischen Eigenschaften des physikalischen Raumes nicht durch die Philosophie, sondern nur durch Experimente geklärt werden kann. Inzwischen hat „die“ Geometrie sich bereits in eine Anzahl von nach Methode und Inhalt sehr unterschiedlichen Theorien aufgespalten: projektive Geometrie, darstellende Geometrie, sphärische Geometrie, Differentialgeometrie usw. Gegen Ende des 19. Jahrhunderts schält sich nach einem langen schwierigen Entwicklungsprozeß die abstrakte Gruppentheorie als eine erste bewußt nicht kategorische algebraische Theorie heraus. Damit hat sich das Bild der Mathematik gründlich gewandelt. Sie ist nicht mehr nur die Wissenschaft von den Zahl- und Raumgrößen (wie sie noch kurz zuvor etwa von ENGELS durchaus zutreffend charakterisiert werden konnte), sondern eine Anhäufung höchst unterschiedlicher Theorien, denen scheinbar nur eines gemeinsam ist: Aus gewissen Grundannahmen (Axiomen), die sich auf gewisse Grundbegriffe beziehen, werden Schlußfolgerungen gezogen. Spätestens die Gruppentheorie zeigt, daß dabei im allgemeinen die platonische und die Kantsche Vorstellung vom Wesen der Grundbegriffe unhaltbar sind, und DAVID HILBERT (1862—1943) macht um die Jahrhundertwende mit seinem berühmt gewordenen Ausspruch, man müsse statt Punkte, Geraden und Ebenen auch Tische, Bänke und Bierseidel sagen können, klar: Was für die Gruppentheorie gilt, das gilt auch für die euklidische Geometrie. Unter den Grundbegriffen einer axiomatischen Theorie hat man sich beliebige Dinge vorzustellen, die lediglich in den durch die Axiome formulierten Beziehungen zueinander stehen. Ein mathematischer Beweis ist nur dann exakt, wenn er jegliche anschauliche Vorstellung von den Grundbegriffen ausschließt und unter alleiniger Anwendung von ausdrücklich hingeschriebenen Voraussetzungen und durch die Logik sanktionierten Schlußweisen erfolgt. Daß es von der Aufstellung dieses Programms noch ein weiter Weg bis zu seiner konsequenten Ausführung ist, zeigt am besten ein Blick in die verschiedenen Auflagen von HILBERTS *Grundlagen der Geometrie* (1899), in denen viele vom heutigen Standpunkt grobe Formulierungsfehler erst nach und nach mit HILBERTS eigener wachsender Einsicht in die Grundlagenprobleme beseitigt werden.

Eine Disziplin der Mathematik wurde zunächst vom sonst allgemein anerkannten axiomatischen Vorgehen ausgenommen, und zwar war es die um 1870 von GEORG CANTOR (1845—1918) begründete Mengenlehre, die in gewisser Weise den Platz der euklidischen Geometrie einnahm, indem jetzt ihre Grundbegriffe mit einer a priori gegebenen denknotwendigen Interpretation versehen wurden. Anders gesagt: Der in weiten Teilen der Mathematik überwundene Platonismus fand zunächst in der Mengenlehre eine neue Heimstatt. Das Reich der Mengen existiert für CANTOR von Anbeginn, unabhängig von jeder materiellen Realität und vom menschlichen Bewußtsein. Die in diesem Reich geltenden Gesetze sind denknotwendig und lassen sich durch reines Denken nach und nach entdecken. Dieser neue Platonismus, zu dem CANTOR sich wiederholt nachdrücklich bekannte, liegt z. B. der anfangs herrschenden Vorstellung von der Denknotwendigkeit des Auswahlaxioms (wo in Wahrheit lediglich eine im Endlichen gesammelte Erfahrung unbewußt auf unendliche Mengen übertragen wurde) und der Idee CANTORS und seiner Zeitgenossen zugrunde, es müsse im Prinzip feststehen, ob die Kontinuumhypothese gilt oder nicht, und man brauche den wahren Sachverhalt nur zu entdecken. Der naiven Mengenlehre war nur ein kurzes Leben bis zur Entdeckung der ersten Antinomien (Widersprüche) beschieden. Die aus heutiger Sicht fast unvorstellbar niederschmetternde Wirkung dieser Antinomien auf viele damalige Mathematiker (jedoch bei weitem nicht alle!) liegt gerade darin begründet, daß die Mengenlehre nicht als eine beliebige mathematische Theorie, sondern als Bestandteil der Logik aufgefaßt wurde und daß daher die Widersprüche in der Mengenlehre nicht nur Zweifel an der Mengenlehre, sondern am logischen Schließen schlechthin weckten.

Es ist heute üblich geworden, die verschiedenen Reaktionen, die der Zusammenbruch der naiven Mengenlehre auslöste, in drei Richtungen der (methodischen und philosophischen) Bewältigung einzuteilen, die als *Logizismus*, *Formalismus* und *Intuitionismus* bezeichnet werden. Diese Klassifikation mag einer ersten Übersicht förderlich sein. Sie vereinfacht und vergröbert aber die genaue Analyse der vielen tatsächlichen Tendenzen. In Wirklichkeit schwankten gerade die führend an den Grundlagenauseinandersetzungen beteiligten Mathematiker zwischen den durch die drei Richtungen markierten Extremen. Außerdem sollte sorgfältig zwischen den mathematisch-methodischen (und als solche oft sehr fruchtbaren) Ansätzen und den motivierend dahinterstehenden philosophischen Strömungen unterschieden werden.

Der Logizismus beharrte auf der Position, daß die Mengenlehre, und damit im Prinzip die gesamte Mathematik, ein Teil der Logik sei. Ausgehend von dieser These umging man die in der naiven Mengenlehre aufgetretenen Widersprüche durch eine strenge Stufung der Mengen und allgemeiner der Relationen derart, daß der Typ einer beliebigen Relation stets höher (komplizierter) ist als die Typen der Dinge, die in dieser Relation stehen können. Dieser Ansatz führte zur verzweigten Typentheorie. Im Hauptwerk des Logizismus, den von BERTRAND RUSSELL (1872—1969) und ALFRED NORTH WHITEHEAD (1861—1947) verfaßten dreibändigen *Principia Mathematica*, wird der Versuch unternommen, die gesamte Mathematik im Rahmen der verzweigten Typentheorie formalisiert aufzubauen. Dieser Formalismus ist jedoch so schwerfällig, daß sich schon bald die boshafte Legende bildete, kein Mensch habe jemals das gesamte Werk studiert, nicht einmal RUSSELL und WHITEHEAD selbst, denn keiner von beiden habe den vom anderen verfaßten Teil vollständig gelesen. Die Kompliziertheit der verzweigten Typentheorie beruht z. T. darauf, daß es zum Zeitpunkt ihrer Entstehung (1910 bis 1913) noch keine mengentheoretische Definition des allgemeinen Relationsbegriffes und daher keine Möglichkeit gab, beliebige n-stellige Relationen als spezielle einstellige Relationen (Mengen) aufzufassen. Durch die mengentheoretische Definition des n-Tupel-Begriffs vereinfacht sich die verzweigte Typentheorie zur einfachen Typentheorie, deren Gegenstand lediglich die auf einem gewissen System von Urelementen aufgebaute Hierarchie von Mengen endlicher Stufe ist. Dabei bestehen Mengen erster Stufe aus Urelementen, und Mengen $(n + 1)$-ter Stufe enthalten als Elemente nur Mengen n-ter Stufe. Die einfache Typentheorie wird heute meist als eine gewisse Variante der axiomatischen Mengenlehre aufgefaßt, die zwar infolge der strengen Homogenität der Mengen technisch etwas schwerfällig ist, jedoch gewisse methodische Vorzüge besitzt und im

Prinzip für die meisten Bedürfnisse der Mathematik ausreicht. Von dieser Auffassung ist der eigentliche Logizismus deutlich unterschieden durch die platonische Interpretation des Begriffs „Menge n-ter Stufe" (in bezug auf einen gegebenen Bereich von Urelementen). Für den Logizismus ist die Frage der Widerspruchsfreiheit der typentheoretischen Mengenlehre, und damit aller in sie semantisch einbettbaren mathematischen Theorien, durch das als existent angenommene „Standardmodell" gegenstandslos. Als philosophische Plattform ist der Logizismus heute schwer erschüttert durch die auch für den typentheoretischen Aufbau der Mengenlehre nachvollziehbare Konstruktion von Nichtstandardmodellen, und insbesondere auch durch das Skolemsche Paradoxon (vgl. 6.4.).

Der von HILBERT begründete Formalismus verzichtete unter dem Eindruck der Antinomien für die Mengenlehre wie für die gesamte Mathematik überhaupt auf den Interpretationsbegriff und wollte jede mathematische Theorie rein syntaktisch begründen. HILBERT formulierte um 1900[1]) das nach ihm benannte Programm, wenigstens solche Theorien wie die Mengenlehre und die Arithmetik der natürlichen Zahlen, auf deren gesicherte Widerspruchsfreiheit die Widerspruchsfreiheit vieler anderer mathematischer Theorien zurückgeführt werden kann, durch elementare Axiomensysteme zu charakterisieren und deren syntaktische Vollständigkeit und Widerspruchsfreiheit mittels unanfechtbarer „finiter" Methoden zu beweisen. Der realisierbare Teil des Hilbertschen Programms wird als Beweistheorie bezeichnet. Die Beweistheorie hatte neben der genauen Untersuchung der syntaktischen Aspekte der Metamathematik auch einige direkte Teilerfolge in Richtung des Hilbertschen Programms aufzuweisen. So gelang G. GENTZEN 1936 der Nachweis der syntaktischen Widerspruchsfreiheit eines Fragments der elementaren Zahlentheorie mit annähernd „finiten" Mitteln, P. S. NOVIKOV 1941 ein noch weitergehendes Ergebnis, allerdings mit stärkeren Hilfsmitteln. Die prinzipielle Undurchführbarkeit des Hilbertschen Programms bewies jedoch KURT GÖDEL (1906—1978) schon 1930. Über den in 8.4. behandelten Unvollständigkeitssatz hinaus zeigte er, daß der Beweis der Widerspruchsfreiheit einer Theorie, die ein hinreichend ausdrucksfähiges Fragment der Zahlentheorie enthält, mit den von HILBERT zugelassenen Mitteln unmöglich ist. Die Resultate der Beweistheorie sind u. a. in [18] und [30] lehrbuchmäßig dargestellt.

Als Vorläufer des Intuitionismus ist LEOPOLD KRONECKER (1823—1891) anzusehen, der bereits die gesamte Mathematik auf die „intuitiv" einsichtigen Eigenschaften der natürlichen Zahlen zurückführen wollte und insbesondere in scharfem Gegensatz zu CANTOR jeglichen Gebrauch des Aktualunendlichen in der Mathematik ablehnte. Als eigentlicher Begründer des Intuitionismus gilt LUITZEN EGBERTUS JAN BROUWER (1882—1967) neben HENRI POINCARÉ (1854—1912), HENRI LEBESGUE (1875—1941), HERMANN WEYL (1885—1955) und anderen, die in z. T. gemäßigterer

[1]) U. a. in seinem berühmten „Problemvortrag" auf dem 2. Internationalen Mathematikerkongreß als 2. Problem. Siehe hierzu und für das Folgende auch den Kommentar von A. S. ESENIN-VOL'PIN zum 2. Hilbertschen Problem in: Die Hilbertschen Probleme, Leipzig 1971 (Übersetzung aus dem Russischen).

Form ähnliche Ansichten vertraten. Die Intuitionisten suchten die Ursache für die aufgetretenen Widersprüche im unbedachten Gebrauch des Aktualunendlichen und insbesondere im Existenzbegriff der klassischen Mathematik. Existenz ist nach Ansicht der Intuitionisten im Sinne von Konstruierbarkeit zu interpretieren (wobei allerdings niemals völlige Übereinstimmung über die zulässigen Konstruktionsmittel herrschte). Aus der Sicht der Intuitionisten sind viele Grundgesetze der klassischen Logik zu verwerfen, vor allem das „tertium non datur", d. h. die Allgemeingültigkeit von $p \vee \neg p$. Während Brouwer selbst jede Formalisierung und Kalkülisierung seiner Vorstellungen ablehnte, gab A. Heyting 1930 einen aus Axiomen und Schlußregeln bestehenden Beweiskalkül an, der genau die im intuitionistischen Sinne gültigen Ausdrücke einer prädikatenlogischen Sprache aufzählt. Wenig später interpretierte A. N. Kolmogorov diesen Kalkül als eine Logik der Konstruktionsaufgaben. Hieran anknüpfend hat sich in den vergangenen Jahrzehnten die *konstruktive Mathematik* entwickelt, als deren wichtigste Vertreter hier A. A. Markov, P. S. Novikov, N. A. Šanin, A. S. Esenin-Vol'pin und R. L. Goodstein genannt seien. Die Konstruktivisten schränken ihre Betrachtungen auf effektiv kodierbare Objekte ein (vgl. 7.1.) und interpretieren Existenz als algorithmische Berechenbarkeit. Die konstruktive Mathematik spiegelt in gewisser Weise viele Verhältnisse besser wider als die klassische Mathematik. Jedoch würde die Anwendung ihrer Prinzipien auf die gesamte Mathematik diese erheblich komplizieren, inhaltlich verarmen und (wie Hilbert schon in bezug auf den Intuitionismus formulierte) am Streik der Masse der Mathematiker scheitern. Es besteht hier eine gewisse Analogie zum Verhältnis zwischen klassischer und relativistischer Physik. Die letzte spiegelt die Realität genauer wider, ist aber zugleich komplizierter, und es würde niemandem einfallen, sie auf alltägliche Probleme anzuwenden, für die die klassische Physik nach wie vor genau genug ist.

Zusammenfassend kann man sagen, daß jede der drei Hauptströmungen die Mathematik um wertwolle inhaltliche und methodische Gesichtspunkte bereicherte, jedoch jeweils einen Aspekt der Mathematik verabsolutierte: der Logizismus den semantischen Aspekt und die Rolle der Mengenlehre, der Formalismus den syntaktischen Aspekt und die Rolle der axiomatischen Methode, der Konstruktivismus den Aspekt der Effektivität.

Nach den grundlegenden Resultaten von Gödel wandte sich die Mehrheit der an Grundlagenfragen interessierten Mathematiker den mit dem präzisierten Berechenbarkeits- und Entscheidbarkeitsbegriff zusammenhängenden Fragen zu. Dabei standen zunächst im Vordergrund[1])

— die Entwicklung mehrerer Varianten des Algorithmenbegriffs und der Nachweis ihrer Gleichwertigkeit (vgl. 7.5.),

[1]) Zu den folgenden Angaben vgl. A. Mostowski u. a., Der gegenwärtige Stand der Grundlagenforschung in der Mathematik, in: Die Hauptreferate des 8. Polnischen Mathematikerkongresses, DVW, Berlin 1954. Das Studium dieses Berichts als Ergänzung zum vorliegenden Abschnitt ist sehr zu empfehlen.

— der Nachweis der Unentscheidbarkeit zahlreicher Probleme aus verschiedenen Bereichen der Mathematik und Logik (u. a. Unentscheidbarkeit der elementaren Arithmetik der rationalen Zahlen, J. ROBINSON 1949; Unentscheidbarkeit des Wortproblems für Halbgruppen, E. POST 1943 und A. A. MARKOV 1947; Unentscheidbarkeit des Wortproblems für Gruppen, P. S. NOVIKOV 1952),
— der Nachweis der Entscheidbarkeit einiger elementarer Theorien (u. a. elementare Arithmetik der reellen Zahlen und elementare euklidische Geometrie, A. TARSKI 1951; kommutative Gruppentheorie, W. SZMIELEW 1948),
— die Übertragung des Berechenbarkeitsbegriffs auf kodierbare Objekte, wie reelle Zahlen, Ordinal- und Kardinalzahlen u. a. (Übergänge zur konstruktiven Mathematik).

In den letzten Jahren entwickelte sich aus der innerhalb der mathematischen Logik entstandenen Reduktionstheorie des Entscheidungsproblems [37] die Theorie der relativen Berechenbarkeit und Entscheidbarkeit [34] und aus Fragestellungen der konstruktiven Mathematik die Theorie der numerierten (kodierten) Mengen.[1])

Während die bisher genannten neuen Forschungsrichtungen letzten Endes aus syntaktischen Problemen der Logik hervorgegangen sind, entwickelte sich aus den semantischen Aspekten in breitem Maße die Modelltheorie (u. a. [25], [33], [46]) als Zweig der allgemeinen Algebra. Im weiteren Sinne sollten zur Modelltheorie auch die mannigfachen algebraischen Charakterisierungen der Modellklassen geometrischer Axiomensysteme gezählt werden. Hier sei als Beispiel eines in letzter Zeit erzielten Fortschritts die Aufklärung der Rolle des Axioms von PASCH (insbesondere der Nachweis seiner Unabhängigkeit) genannt.[2])

Zusammenfassend kann heute festgestellt werden, daß mathematische Theorien, die historisch aus Grundlagenproblemen entstanden sind, sehr schnell einen hohen Abstraktionsgrad und eine beträchtliche Selbständigkeit erreichen, so daß ihnen die Spuren ihrer Entstehung oft nach kurzer Zeit kaum noch anzusehen sind und sie häufig Anwendungen in vom Ausgangspunkt ihrer Entstehung weit entfernten Bereichen der Mathematik und Technik finden. Die starke Verzweigung und hohe Spezialisierung der Grundlagen der Mathematik, die schon längst von keinem einzelnen mehr überblickt werden kann, geht wohl am deutlichsten aus der gegenwärtig in den internationalen Referatenorganen der Mathematik gebräuchlichen Klassifi-

---

[1]) Der gegenwärtige Stand dieser Theorie ist in drei in deutscher Übersetzung vorliegenden Veröffentlichungen von JU. L. ERŠOV dargestellt (vgl. JU. L. ERŠOV, Theorie der Numerierungen I, DVW, Berlin 1973, oder Zeitschr. f. Math. Logik u. Grdlg. d. Math. 19 (1972), 289—388; II, DVW, Berlin 1976, oder Zeitschr. f. Math. Logik u. Grdlg. d. Math. 21 (1975), 473—584; III, DVW, Berlin 1978, oder Zeitschr. f. Math. Logik u. Grdlg. d. Math. 23 (1977), 289—371).

[2]) L. W. SZCZERBA, Independence of Pasch's axiom, Bull. Acad. Polon. Sci., Sér. Sci. Math. Astron. Phys., 18 (1970), 491—498, und L. W. SZCZERBA and W. SZMIELEW, On the euclidean geometry without the Pasch axiom, Bull. Acad. Polon. Sci., Sér. Sci. Math. Astron. Phys., 18 (1970), 659—666. Weitere Arbeiten zu diesem Problemkreis sind zitiert in: Kolloquium Grundlagen der Geometrie und algebraische Methoden, Potsdamer Forschungen, Reihe B, Heft 3, 1974.

kation des Gebietes Logik und Grundlagen der Mathematik hervor, die im folgenden gekürzt wiedergegeben wird:

Klassische logische Systeme

Aussagenkalkül
Prädikatenkalkül erster Stufe
Prädikatenkalküle höherer Stufe
Unübliche Quantoren
Formalisierte Sprachen mit unendlich langen Ausdrücken
Sonstiges

Nichtklassische formale Systeme

Mehrwertige Logik
Modale Logik u. ä.
Formalisierungen der intuitionistischen Logik u. ä.
Kombinatorische Logik
Sonstiges

Beweistheorie

Konstruktive Mathematik

Intuitionistische Mathematik
Algorithmen
Berechenbare Funktionen
Sonstiges

Rekursionstheorie

Thue- und Post-Systeme u. ä.
Automaten
Turingmaschinen
Klassifikation rekursiver Funktionen
Aufzählbare Mengen
Rekursionstheorie im Bereich von Ordinalzahlen, Mengen und anderen abstrakten Strukturen
Unentscheidbarkeitsgrade
Hierarchien
Rekursive Äquivalenztypen
Formale Systeme der Berechenbarkeit
Kombinatorische Funktionen
Wortprobleme
Anwendungen
Sonstiges

Methodologie der deduktiven Systeme

Entscheidbarkeit und Unentscheidbarkeit
Axiomatisierbarkeit
Endliche Axiomatisierbarkeit
Vollständigkeit, Kategorizität usw.
Sonstiges

Modelltheorie

Modelle elementarer Theorien
Modelle anderer Theorien
Modellkonstruktionen
Anwendung auf Algebra, Zahlentheorie u. a.
Nichtstandardmodelle
Sonstiges

Algebra der Logik

Boolesche Algebren, Verbände, Topologien
Relationenalgebra
Zylindrische und polyadische Algebren
Sonstiges

Mengenlehre

Widerspruchsfreiheit und Unabhängigkeit
Nichtklassische Mengentheorien
Axiomatik
Auswahlaxiom und äquivalente Sätze
Kontinuumhypothese (einschließlich verallgemeinerter)
Deskriptive Mengenlehre
Große Kardinal- und Ordinalzahlen
Sonstiges

Die hier kurz dargestellte historische Entwicklung mathematischer Grundlagenfragen von der Antike bis in unsere Zeit macht eine Tendenz deutlich: Überwogen am Anfang der Entwicklung Fragestellungen deutlich philosophischen Inhalts, so sind die Untersuchungen mathematischer Grundlagenfragen heute — etwa seit 50 Jahren — vornehmlich auf innermathematische und logische bzw. metamathematische und metalogische Fragestellungen gerichtet und werden dementsprechend im allgemeinen mit mathematischen und nicht mit philosophischen Methoden durchgeführt. Die dabei gewonnenen Erkenntnisse haben über zahlreiche Probleme, so über das Verhältnis von Mathematik und Logik, von Mathematik und Sprache, von Semantik und Syntax, über die Tragweite und die Grenzen der axiomatischen und der konstruktiven Methode u. a. m. Klarheit geschaffen oder zumindest vieles erhellt, was vorher im Dunklen war. Nicht wenige dieser Erkenntnisse haben inzwischen auch für andere Wissenschaften Bedeutung erlangt oder sogar zur Herausbildung neuer Wissenschaftszweige, wie der Kybernetik und der Informationswissenschaft, beigetragen. Darin zeigt sich die große Leistungsfähigkeit und Bedeutung der mathematischen Grundlagenforschung in dem in diesem Lehrbuch dargestellten „engeren Sinne".

Zu den Grundlagenfragen jeder Wissenschaft, also auch der Mathematik, gehören darüber hinaus oder sogar in erster Linie philosophische, vor allem erkenntnistheoretische Probleme, Grundlagenfragen im „weiteren Sinne", in die die von der betreffenden Einzelwissenschaft untersuchten Grundlagenprobleme eingebettet sind und mit denen sie in Wechselwirkung stehen. Wir haben das Wechselverhältnis zwischen den Grundlagen der Mathematik und der marxistisch-leninistischen Philosophie in diesem Lehrbuch bereits an verschiedenen Stellen behandelt.

Die im Vergleich zu anderen Wissenschaften sehr weit- und tiefgehende mathematische Grundlagenforschung hat bei nicht wenigen Mathematikern und Philosophen nichtmarxistischer Weltanschauung in den letzten Jahrzehnten die Auffassung hervorgebracht, in der Mathematik gäbe es keine philosophischen Grundlagenprobleme bzw. die Mathematiker vermögen diese Probleme mit innermathematischen Methoden selbst zu lösen bzw. brauchten sich nicht um sie zu kümmern. Mathematik einschließlich mathematischer Grundlagenforschung könne man auch betreiben, ohne über philosophische Fragen der Mathematik Bescheid zu wissen oder über sie nachzudenken. Eine solche Leugnung der Notwendigkeit und Bedeutung philosophischer Grundlagenuntersuchungen in der Mathematik ist nach A. D. Alexandrov „nichts anderes als selbstzufriedene Kulturlosigkeit"[1]).

Es konnte nicht Anliegen dieses Lehrbuches sein, die philosophischen Grundlagen der Mathematik in diesem umfassenden Sinne in die Darstellung einzubeziehen. Es erscheint aber notwendig, auch an dieser Stelle nachdrücklich auf die Existenz und Bedeutung der philosophischen Probleme der Mathematik hinzuweisen.

---

1) A. D. Alexandrov, Mathematik und Dialektik, in: Ideen des exakten Wissens 4 (1971), S. 256 [russ.].

Literatur

Aus dem riesigen Angebot von Büchern über Grundlagen der Mathematik oder gewisse Teilgebiete wurden die folgenden Titel unter den beiden Gesichtspunkten ihrer grundlegenden (bzw. eventuell historischen) Bedeutung und ihrer Erreichbarkeit für den Leser ausgewählt.

[1] Ackermann, W., Solvable Cases of the Decision Problem, North-Holland Publ. Comp., Amsterdam 1954.

[2] Asser, G., Einführung in die mathematische Logik, Teil I. Aussagenkalkül, 6. Aufl., BSB B. G. Teubner, Leipzig 1982.

[3] Asser, G., Einführung in die mathematische Logik, Teil II. Prädikatenkalkül der ersten Stufe, 2. Aufl., BSB B. G. Teubner, Leipzig 1975.

[3'] Asser, G., Einführung in die mathematische Logik, Teil III. Prädikatenlogik höherer Stufe, BSB B. G. Teubner, Leipzig 1981.

[4] Berka, K., und L. Kreiser, Logik-Texte, kommentierte Auswahl zur Geschichte der modernen Logik, Akademie-Verlag, Berlin 1971.

[5] Bernays, P., und A. A. Fraenkel, Axiomatic Set Theory, North-Holland Publ. Comp., Amsterdam 1958.

[6] Church, A., Introduction to Mathematical Logic, Princeton Univ. Press, Princeton, N. J., 1956.

[7] Davis, M., The Undecidable, Raven Press, New York 1965.

[8] Davis, M., Computability and Unsolvability, McGraw-Hill Book Co., Inc., New York—Toronto—London 1958.

[9] Fraenkel, A., and Y. Bar-Hillel, Foundations of Set Theory, North-Holland Publ. Comp., Amsterdam 1958 (russische Übersetzung: Mir, Moskau 1966).

[10] Freudenthal, H., Einführung in die Sprache der Logik, Oldenbourg-Verlag, München—Wien 1965 (russische Übersetzung: Mir, Moskau 1969).

[11] Grzegorczyk, A., An Outline of Mathematical Logic, Reidel, Dordrecht—Boston/PWN, Warschau 1974.

[12] Hasenjaeger, G., Einführung in die Grundbegriffe und Probleme der modernen Logik, Verlag Karl Alber, Freiburg—München 1962.

[13] Henkin, L., P. Suppes and A. Tarski (Ed.), The axiomatie Method, North-Holland Publ. Comp., Amsterdam 1959.

[14] HERMES, H., Introduction to Mathematical Logic, Springer-Verlag, Berlin—Heidelberg—New York 1973 (deutsch: B. G. Teubner, Stuttgart 1963).

[15] HERMES, H., Aufzählbarkeit, Entscheidbarkeit, Berechenbarkeit, 2. Aufl., Springer-Verlag, Berlin—Göttingen—Heidelberg 1971 (ab 1965 im gleichen Verlag auch in englisch).

[16] HILBERT, D., Grundlagen der Geometrie, 9. Aufl., B. G. Teubner, Stuttgart 1962.

[17] HILBERT, D., und W. ACKERMANN, Grundzüge der theoretischen Logik, 4. Aufl., Springer-Verlag, Berlin—Göttingen—Heidelberg 1959.

[18] HILBERT, D., und P. BERNAYS, Grundlagen der Mathematik, Band 1 und 2, 2. Aufl., Springer-Verlag, Berlin—Heidelberg—New York 1968.

[19] JACOBS, K. (Ed.), Selecta Mathematica II (Turing-Maschinen und berechenbare Funktionen), Springer-Verlag, Berlin—Heidelberg—New York 1970.

[20] KLAUA, D., Allgemeine Mengenlehre I und II, Akademie-Verlag, Berlin 1968 bzw. 1969 (beim gleichen Verlag auch als Taschenbücher: Elementare Axiome der Mengenlehre (1971), Grundbegriffe der axiomatischen Mengenlehre 1 und 2 (1973), Kardinal- und Ordinalzahlen 1 und 2 (1974)).

[21] KLAUS, G., Moderne Logik, VEB Deutscher Verlag der Wissenschaften, Berlin 1972.

[22] KLEENE, ST. C., Introduction to Metamathematics, North-Holland Publ. Comp. und Noordhoff, Amsterdam—Groningen 1952.

[23] KLEENE, ST. C., Mathematical Logic, Wiley, New York—London—Sydney 1967 (russische Übersetzung: Mir, Moskau 1973).

[24] LORENZEN, P., Formale Logik, W. de Gruyter, Berlin 1958.

[25] MALCEW, A. I., Algebraic Systems, Akademie-Verlag, Berlin 1973 (Übersetzung aus dem Russischen).

[26] MALCEW, A. I. Algorithmen und rekursive Funktionen, Akademie-Verlag, Berlin 1974 (Übersetzung aus dem Russischen).

[27] Марков, А. А., Теория алгорифмов, Тр. Матем. ин.-та им. Стеклова АН СССР 42 (1954), 1—376.

[28] MOLODSCHI, W. N., Studien zu philosophischen Problemen der Mathematik, VEB Deutscher Verlag der Wissenschaften, Berlin 1977 (Übersetzung aus dem Russischen).

[29] MOSTOWSKI, A., Sentences Undecidable in Formalized Arithmetic, North-Holland Publ. Comp., Amsterdam 1957.

[30] NOVIKOV, P. S., Grundzüge der mathematischen Logik, Vieweg, Braunschweig/VEB Deutscher Verlag der Wissenschaften, Berlin 1973 (Übersetzung aus dem Russischen).

[31] PÉTER, R., Rekursive Funktionen, Akadémiai Kiadó, Budapest 1951/Akademie-Verlag, Berlin 1957.

[32] RASIOWA, H., and R. SIKORSKI, The Mathematics of Metamathematics, PWN, Warschau 1963.

[33] ROBINSON, A., Introduction to Model Theory and to the Metamathematics of Algebra, North-Holland Publ. Comp., Amsterdam 1963 (russische Übersetzung: Mir, Moskau 1967).

[34] ROGERS JR., H., Theory of Recursive Functions and Effective Computability, McGraw-Hill, New York—London—Sydney 1967 (russische Übersetzung: Mir, Moskau 1972).

[35] SEGETH, W., Elementare Logik, 6. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1971.

[36] SINOWJEW, A., und H. WESSEL, Logische Sprachregeln, VEB Deutscher Verlag der Wissenschaften, Berlin 1975.

[37] SURANYI, J., Reduktionstheorie des Entscheidungsproblems im Prädikatenkalkül der ersten Stufe, Akadémiai Kiadó, Budapest/VEB Deutscher Verlag der Wissenschaften, Berlin 1959.

[38] SURMA, ST. J. (Ed.), Studies in the History of Mathematical Logic, Ossolineum, Warschau 1973.

[39] Стяжкин, Н. И., Становлеиие идей математической логики, Наука, Москва 1964.

[40] TARSKI, A., Introduction to Logic and to the Methodology of Deductive Sciences, New York 1965.
[41] TARSKI, A., A. MOSTOWSKI and A. ROBINSON, Undecidable Theories, North-Holland Publ. Comp., Amsterdam 1953.
[42] THIELE, H., Wissenschaftstheoretische Untersuchungen in algorithmischen Sprachen I, VEB Deutscher Verlag der Wissenschaften, Berlin 1966.
[43] VARGA, T., Mathematische Logik für Anfänger (2 Teile), Verlag Volk und Wissen, Berlin 1968 (Aussagenlogik) bzw. 1973 (Prädikatenlogik).
[44] WANG HAO, A Survey of Mathematical Logic, Science Press, Peking 1962.
[45] BORKOWSKI, L., Formale Logik, Akademie-Verlag, Berlin 1976.
[46] CHANG, C. C., and H. J. KEISLER, Model Theory, North-Holland Publ. Comp., Amsterdam—London 1973.
[47] HEITSCH, W., Mathematik und Weltanschauung, Akademie-Verlag, Berlin 1976.
[48] RUSZA, I., Die Begriffswelt der Mathematik (insbesondere Kapitel 5 und 6), Verlag Volk und Wissen, Berlin 1976.
[49] SCHREIBER, P., Die allgemeinste Form der formalisierten Sprachen, Jenaer Frege-Konferenz 1979 (Wiss. Beiträge der Friedrich-Schiller-Universität), S. 414—432.
[50] SCHREIBER, P., Grundlagen der konstruktiven Geometrie, VEB Deutscher Verlag der Wissenschaften, Berlin 1984.

Namen- und Sachverzeichnis