Mathematische Logik

Helmut Schwichtenberg

Mathematisches Institut der Universität München
Wintersemester 1999/2000 und Sommersemester 2000

Vorwort

Dieses Skriptum entstand als Begleitmaterial zu einer Vorlesung über Mathematische Logik am Mathematischen Institut der Universität München, die ich in zwei Teilen im Wintersemester 1999/2000 und im Sommersemester 2000 gehalten habe.

Bedanken möchte ich mich bei den Herren Klaus Aehlig und Dr. Oliver Deiser, die die Übungen zu dieser Vorlesung korrigiert bzw. betreut und in vielen Diskussionen zu ihrem Inhalt wesentlich beigetragen haben.

München, August 2000
Helmut Schwichtenberg

Inhaltsverzeichnis

1. Logik

Untersuchungsgegenstand dieser Vorlesung sind mathematische Beweise. In diesem einführenden Kapitel befassen wir uns mit den Grundlagen der Formalisierung derartiger Beweise. Dazu verwenden wir ein System des "natürlichen Schließens", das 1934 von GENTZEN [9] eingeführt wurde. Es gibt zwei Hauptgründe für diese Wahl. Erstens handelt es sich – wie schon der Name sagt – um ein *natürliches* Beweissystem, in dem Beweise in einer Art dargestellt sind, wie sie ein sorgfältiger Mathematiker formulieren würde, wenn er alle Einzelheiten einer Argumentation aufschreiben möchte. Zweitens sind formale Beweise im natürlichen Schließen eng verknüpft (mittels der sogenannten CURRY-HOWARD Korrespondenz) mit Termen im getypten λ-Kalkül. Dies liefert nicht nur eine kompakte Schreibweise für formale Beweise (die sonst leicht zu unhandlichen Baumstrukturen werden), sondern öffnet auch einen Weg zur Anwendung von im Kontext des λ-Kalküls gebräuchlichen rechnerischen Techniken.

Außer der klassischen Logik werden wir uns auch mit konstruktiven Logiken befassen: mit Minimallogik und mit intuitionistischer Logik. Dies wird einige interessante Aspekte von Beweisn zum Vorschein bringen. So ist es zum Beispiel möglich und sinnvoll, zwischen Existenzbeweisen zu unterscheiden, die das als existent nachgewiesene Objekt tatsächlich liefern, und solchen, die dies nicht tun. Als Beispiel betrachten wir die folgende Aussage.

Es gibt irrationale Zahlen a, b mit a^b rational.

Einen Beweis erhält man wie folgt durch Fallunterscheidung.

Fall $\sqrt{2}^{\sqrt{2}}$ ist rational. Man wähle $a = \sqrt{2}$ und $b = \sqrt{2}$. Dann sind a, b irrational, und nach Annahme ist a^b rational.

Fall $\sqrt{2}^{\sqrt{2}}$ ist irrational. Man wähle $a = \sqrt{2}^{\sqrt{2}}$ und $b = \sqrt{2}$. Dann sind nach Annahme a, b irrational, und

$$a^b = \left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}} = \left(\sqrt{2}\right)^2 = 2$$

ist rational. □

Solange wir nicht entschieden haben, ob $\sqrt{2}^{\sqrt{2}}$ nun rational ist oder nicht, wissen wir nicht, welche Zahlen a, b wir nehmen müssen. Damit haben wir ein Beispiel eines Existenzbeweises, der es nicht erlaubt, das als existent nachgewiesene Objekt tatsächlich anzugeben.

Eine besondere Eigenart von GENTZENs Kalkül des natürlichen Schließens ist es, daß für jede logische Verknüpfung $\wedge$, $\to$ und $\forall$ *Einführungs-* und *Beseitigungsregeln* vorhanden sind. Dies führt auf das System der *Minimallogik*, das von JOHANSSON [15] 1937 eingeführt wurde. Gibt man dann dem speziellen Aussagensymbol $\bot$ (gelesen "falsum", für "Falschheit") einen besonderen Status, so erhält man die intuitionistische Logik (mittels des Prinzips des ex-falso-quodlibet) und die klassische Logik (mittels des Prinzips des indirekten Beweisens)). Für diesen Aufbau der Logik ist es wichtig, daß wir uns auf eine Sprache beschränken, die nur die logischen Verknüpfungen $\to$, $\wedge$ und $\forall$ enthält; die Disjunktion $\vee$ und der Existenzquantor $\exists$ können mittels ihrer DE MORGANschen Definitionen eingeführt werden. Für diese Verknüpfungen lassen sich dann die üblichen Einführungs- und Beseitigungsregeln herleiten. Schließlich erweitern wir die Sprache noch um den starken (konstruktiven) Existenzquantor $\exists^*$. Man kann dann in den Fällen, wo ein Existenzbeweis tatsächlich konstruktiv durch Angabe eines Beispiels geführt wurde, dies auch in der Formelsprache angemessen ausdrücken.

1.1 Formale Systeme

Der Mathematiker studiert Axiomensysteme. Man unterscheidet zwischen “klassischen” Axiomensystemen, die ein intendiertes Modell möglichst vollständig beschreiben sollen (etwa die natürlichen Zahlen oder die ebene Geometrie) und “modernen” Axiomensystemen wie etwa denen für Gruppen, Ringe und Körper, die von vornherein so konzipiert sind, daß sie möglichst viele Modelle besitzen. Die Grundbegriffe der jeweiligen mathematischen Disziplin sind dann durch die Axiome implizit bestimmt. Ausgehend von den Grundbegriffen werden weitere Begriffe definiert, die aber prinzipiell entbehrlich sind. Aus den Axiomen werden dann mittels logischer Schlußregeln die Sätze der Theorie hergeleitet. – Wir werden uns hier in erster Linie mit klassischen Axiomensystemen befassen.

Einen mathematischen Satz kann man unter zwei Aspekten betrachten: einerseits teilt er einen Inhalt mit (semantischer Aspekt), andererseits liegt er als Zeichenreihe vor (syntaktischer Aspekt). Der semantische Aspekt scheint auf den ersten Blick viel wichtiger zu sein. Es ist jedoch nötig und in manchen Fällen ergiebig, sich auch mit dem syntaktischen Aspekt zu befassen.

Die syntaktische Seite eines Axiomensystems fassen wir in dem Begriff eines *formalen Systems* zusammen. Ein formales System ist bestimmt durch

(1) eine (formale) Sprache, und
(2) Axiome und Schlußregeln.

Als Sprache verwendet man am besten eine künstliche, formale Sprache, um Eindeutigkeit und Einfachheit zu gewährleisten. Eine solche Sprache ist bestimmt durch einen Zeichenvorrat (Alphabet), daraus gebildete Terme (die Objekte bedeuten) und Formeln (die Aussagen bedeuten). Eine Aussage ist nach Aristoteles ein *sprachliches Gebilde, von dem es sinnvoll ist zu sagen, es sei wahr oder falsch.* Die Axiome sind dann einfach spezielle Formeln, und die Schlußregeln gestatten den Übergang von gewissen (meist endlich vielen) Prämissen auf eine Konklusion – im allgemeinen unter zusätzlichen Bedingungen.

Wir wollen hier eine Klasse von formalen Systemen beschreiben, die auf der Logik erster Stufe basieren. Sie reichen für die Darstellung der üblichen Axiomensysteme der Mathematik aus. Als erstes wollen wir ihre Sprache einführen. Dazu beginnen wir mit einer informalen Diskussion.

Jede Sprache verwendet Begriffe. In unserem Fall einer Sprache über mathematische Gegenstände können wir es uns relativ einfach machen. Hier kann man zwischen logischen und nichtlogischen Begriffen unterscheiden.

Wir behandeln zunächst die nichtlogischen Begriffe. Wir legen einen sogenannten Individuenbereich zugrunde (etwa den Bereich $\mathbb{N}$ der natürlichen Zahlen oder die Trägermenge G einer Gruppe) und verstehen dann unter einem Begriff einfach eine Teilmenge des Individuenbereichs; ein Beispiel ist der Primzahlbegriff. In der Sprache muß für jeden relevanten Begriff ein Symbol zu seiner Benennung vorhanden sein, etwa P für den Primzahlbegriff; solche Symbole nennt man Relationssymbole. Es ist notwendig, auch mehrstellige Begriffe zuzulassen, etwa die Kleiner-oder-Gleich-Beziehung $n \le m$ zwischen natürlichen Zahlen, oder die dreistellige Relation “A liegt zwischen B und C” für Punkte $A, B, C \in \mathbb{R}^2$. Die zugehörigen Symbole nennt man Relationssymbole.

Ferner brauchen wir in der Mathematik offenbar Funktionen, etwa die Nachfolgerfunktion und die Addition auf $\mathbb{N}$ oder die Gruppenverknüpfung. In der Sprache muß dann für jede vorkommende Funktion ein Symbol zu ihrer Benennung vorhanden sein, etwa S für die Nachfolgerfunktion, $+$ für die Addition und $\circ$ für die Gruppenverknüpfung; solche Symbole nennt man Funktionssymbole. Jedes Funktionssymbol hat eine feste Stellenzahl. Ferner brauchen wir offenbar in der Sprache Konstanten für einzelne Individuen (etwa 5). Man kann Konstante als entartete (genauer: nullstellige) Funktionssymbole auffassen. Aus den Funktionssymbolen und den Konstanten kann man Terme aufbauen, etwa $\mathsf{S}(5)$ oder $5 + 7$. Mit solchen Termen lassen sich dann atomare Formeln bilden, etwa $P(7)$ (mit der Bedeutung “7 ist eine Primzahl”) oder $\mathsf{S}(5) < 9$.

Wir nächstes betrachten wir die logischen Begriffe. Aus atomaren Formeln kann man mittels logischer Verknüpfungen weitere Formeln zusammensetzen, etwa $2 < 3 \wedge P(7)$. Allgemein ist mit A und B auch die Konjunktion $A \wedge B$ (gelesen “A und B”) eine Formel.

Man beachte, daß die Wahrheit oder Falschheit von $A \wedge B$ allein von der Wahrheit oder Falschheit von A und von B abhängt, nicht von der Bedeutung dieser Aussagen. Wir nehmen deshalb mit Frege an, daß zwei verschiedene Wahrheitswerte W und F gegeben sind. Man beachte, daß es nicht darauf ankommt, was genau die Wahrheitswerte W und F sind, solange sie nur verschieden sind; wir wollen hier 1 und 0 wählen. Wir betrachten jetzt Wahrheitsfunktionen $H\colon \{1,0\}^n \to \{1,0\}$. Die Konjunktion $\wedge$ wird dann vollständig durch ihre Wahrheitsfunktion $H_\wedge\colon \{1,0\}^2 \to \{1,0\}$ beschrieben, die durch $H_\wedge(p,q) := \min(p,q)$ bestimmt ist.

Mit A ist auch die Negation $\neg A$ (gelesen "nicht A") eine Formel. Die zugehörige Wahrheitsfunktion $H_\neg\colon \{1,0\} \to \{1,0\}$ ist einstellig und bestimmt durch $H_\neg(p) := 1-p$.

Mit A und B ist auch die Disjunktion $A \vee B$ (gelesen "A oder B") eine Formel. Man beachte, daß das "oder" hier im nicht ausschließenden Sinn gemeint ist; dies ist in der Mathematik üblich. Bei klassischer Auffassung ist die Disjunktion $\vee$ vollständig durch ihre Wahrheitsfunktion $H_\vee\colon \{1,0\}^2 \to \{1,0\}$ beschrieben, die durch $H_\wedge(p,q) := \max(p,q)$ bestimmt ist. Man beachte, daß dann $A \vee B$ dasselbe bedeutet wie $\neg(\neg A \wedge \neg B)$, also $\vee$ auch durch $\wedge$ und $\neg$ definiert werden kann. Es gilt nämlich $H_\vee(p,q) := H_\neg(H_\wedge(H_\neg(p), H_\neg(q)))$, wie man etwa durch Probieren der $2^2 = 4$ Fälle sofort feststellt. Wir werden auch noch eine starke (oder konstruktive) Auffassung des "oder" besprechen, bei der $\vee$ nicht mehr so definiert werden kann.

Mit A und B ist auch die Implikation $A \to B$ (gelesen "wenn A, so B" oder "A impliziert B") eine Formel. Man beachte, daß wenn-so hier im mathematischen Sinn gemeint ist, d.h. es braucht kein kausaler Zusammenhang zwischen A und B zu bestehen. So ist zum Beispiel $2 < 3 \to P(7)$ wahr. Die Wahrheitsfunktion $H_\to\colon \{1,0\}^2 \to \{1,0\}$ ist $H_\to(p,q) := H_\vee(H_\neg(p), q)$, also

$$\begin{aligned} &H_\to(1,1) = H_\to(0,1) = H_\to(0,0) = 1,\\ &H_\to(1,0) = 0. \end{aligned}$$

Schließlich ist mit A und B auch die Äquivalenz $A \leftrightarrow B$ (gelesen "A genau dann, wenn B" oder "A äquivalent B") eine Formel. Die Wahrheitsfunktion $H_\leftrightarrow\colon \{1,0\}^2 \to \{1,0\}$ ist

$$\begin{aligned} &H_\leftrightarrow(1,1) = H_\leftrightarrow(0,0) = 1,\\ &H_\leftrightarrow(1,0) = H_\leftrightarrow(0,1) = 0. \end{aligned}$$

Die Äquivalenz ist offenbar aus $\wedge$ und $\to$ definierbar durch $(A \to B) \wedge (B \to A)$.

Ferner ist es nützlich, noch das Falsum $\perp$ als logische Konstante einzuführen. Es ist bestimmt durch die "nullstellige" Wahrheitsfunktion $H_\perp := 0$. Man kann dann etwa die Negation $\neg$ aus der Implikation $\to$ und dem Falsum $\perp$ definieren durch $\neg A := A \to \perp$, denn es gilt

$$\begin{aligned} &H_\to(1, H_\perp) = H_\to(1,0) = 0 = H_\neg(1),\\ &H_\to(0, H_\perp) = H_\to(0,0) = 1 = H_\neg(0). \end{aligned}$$

Um auch allgemeine Aussagen machen zu können – etwa $\forall x(e \circ x = x)$ –, benötigen wir noch Variablen $x, y, z, \dots$. Dabei stellen wir uns vor, daß die Variablen über die Elemente unseres Individuenbereichs laufen. Der Termbegriff ist dann so zu erweitern, daß auch Variablen darin zugelassen sind.

Mit A ist auch die Allformel $\forall x A$ (gelesen "für alle x gilt A") eine Formel. Hierbei wird im allgemeinen A eine Formel sein, in der die Variable x vorkommt.

Schließlich ist mit A auch die Existenzformel $\exists x A$ (gelesen "es gibt ein x mit A") eine Formel. Man beachte, daß bei der üblichen *klassischen* (oder schwachen) Auffassung des "es gibt" $\exists x A$ dasselbe bedeutet wie $\neg\forall x \neg A$, also $\exists$ auch durch $\forall$ und $\neg$ definiert werden kann. Wir werden auch noch eine konstruktive (oder starke) Auffassung des Existenzquantors besprechen, bei der dies nicht mehr der Fall ist.

Man beachte ferner, daß die Verwendung der Quantoren $\forall x$ und $\exists x$ die Unterscheidung zwischen freien und gebundenen Vorkommen von Variablen notwendig macht. So ist etwa in $\forall x(x \le x + y)$ das zweite und dritte Vorkommen der Variablen x gebunden, aber das einzige Vorkommen der Variablen y frei. Die letzte Formel kann man unter Erhalt ihrer Bedeutung auch umschreiben in $\forall z(z \le z + y)$, denn dies ist nur eine Umbenennung der gebundenen Variablen. Ein ähnliches Phänomen ist aus der Analysis bekannt, wo ebenfalls

$$\int_0^1 \sin(xy)\,dx = \int_0^1 \sin(ty)\,dt$$

ist; x bzw. t kommen hier gebunden und y frei vor.

Wir geben jetzt einige Beispiele zur Formalisierung mathematischer Aussagen.

Als erstes betrachten wir die Axiome der Gruppentheorie. Grundbegriffe sind die Gruppenverknüpfung $\circ$ (ein zweistelliges Funktionssymbol), die Einheit e (eine Konstante), die Inversenbildung $^{-1}$ (ein einstelliges Funktionssymbol) und schließlich die Gleichheit $=$ (ein zweistelliges Relationssymbol). Als Axiome haben wir

$$\begin{array}{ll} \forall x \forall y \forall z [x \circ (y \circ z) = (x \circ y) \circ z] & \text{(Assoziativität)}, \\ \forall x [e \circ x = x] & \text{(Linkseinheit)}, \\ \forall x [x^{-1} \circ x = e] & \text{(Linksinverses)}. \end{array}$$

Zwei Mengen sind gleich, wenn sie dieselben Elemente enthalten:

$\forall$	für alle Mengen
$\exists$	es gibt eine Menge
$x \in y$	x ist ein Element von y
$x = y$	x ist gleich y

$$\forall x \forall y (\forall z. z \in x \leftrightarrow z \in y) \to x = y.$$

f ist stetig im Punkt a:

f	einstelliges Funktionssymbol
a, 0	reelle Zahlen
$\forall$	für alle reelle Zahlen
$\exists$	es gibt eine reelle Zahl
$x > y$ $(x < y)$	x ist größer (kleiner) als y
$d(x, y)$	der Abstand von x und y

$$\forall \varepsilon . \varepsilon > 0 \to \exists \delta . \delta > 0 \land \forall x . d(x, a) < \delta \to d(f(x), f(a)) < \varepsilon;$$

hierfür schreibt man zur Abkürzung meist

$$\forall \varepsilon{>}0 \exists \delta{>}0 \forall x . d(x, a) < \delta \to d(f(x), f(a)) < \varepsilon.$$

Es gibt eine gemeinsame obere Schranke für Summe und Produkt zweier positiver rationaler Zahlen:

$\forall$	für alle positiven rationalen Zahlen
$\exists$	es gibt eine positive rationale Zahl
$+$	Addition
$\cdot$	Multiplikation
$x \leq y$	x ist kleiner oder gleich y
$2xy$	verdoppeltes Produkt von x und y

$$\forall x \forall y \exists z . x + y \leq z \land x \cdot y \leq z.$$

Für positive rationale Zahlen $x \leq 1$ und y ist $y + 1$ eine gemeinsame obere Schranke für ihre Summe und ihr Produkt:

$$\forall x . x \leq 1 \to \forall y . x + y \leq y + 1 \land x \cdot y \leq y + 1$$

oder kürzer

$$\forall x{\leq}1 \forall y . x + y \leq y + 1 \land x \cdot y \leq y + 1.$$

Für positive rationale Zahlen $x, y > 1$ ist $2xy$ eine gemeinsame obere Schranke für ihre Summe und ihr Produkt:

$$\forall x . x \not\leq 1 \to \forall y . y \not\leq 1 \to [x + y \leq 2xy \land x \cdot y \leq 2xy]$$

oder kürzer

$$\forall x{\not\leq}1 \forall y{\not\leq}\, 1 . x + y \leq 2xy \land x \cdot y \leq 2xy.$$

Formale Sprachen

Wir geben jetzt eine genaue Beschreibung der Sprachen der Logik erster Stufe.

Gegeben sei eine abzählbar unendliche Menge $\{ v_i \mid i \in \mathbb{N} \}$ von *Variablen*; Mitteilungszeichen für Variable sind x, y, z. Eine *Sprache* $\mathcal{L}$ *erster Stufe* ist bestimmt durch ihre *Signatur*. Darunter versteht man folgendes.

– Für jede natürliche Zahl $n \geq 0$ eine (eventuell leere) Menge $\mathsf{Rel}_{\mathcal{L}}^{(n)}$ von n-stelligen *Relationssymbolen* (oder auch *Prädikatensymbolen*). 0-stellige Relationssymbole heißen *Aussagensymbole*. Wir nehmen an, daß ein spezielles Aussagensymbol $\bot$ (gelesen "falsum" oder "bottom") vorhanden ist.
– Für jede natürliche Zahl $n \geq 0$ eine (eventuell leere) Menge $\mathsf{Fun}_{\mathcal{L}}^{(n)}$ von n-stelligen *Funktionssymbolen*. Die 0-stelligen Funktionssymbole heißen *Konstante*.

Wir setzen voraus, daß die Menge der Variablen und alle Mengen $\mathsf{Rel}_{\mathcal{L}}^{(n)}$ und $\mathsf{Fun}_{\mathcal{L}}^{(n)}$ paarweise disjunkt sind.

Zum Beispiel ist die Sprache $\mathcal{L}_G$ der Gruppentheorie bestimmt durch die Signatur bestehend aus den Funktionssymbolen $\circ, e, ^{-1}$ (2-, 0- bzw. 1-stellig) und dem 2-stelligen Relationssymbol $=$.

$\mathcal{L}$-*Terme* werden induktiv wie folgt definiert.

– Jede Variable ist ein $\mathcal{L}$-Term.
– Jede Konstante von $\mathcal{L}$ ist ein $\mathcal{L}$-Term.
– Sind $t_1, \dots, t_n$ $\mathcal{L}$-Terme und ist f ein n-stelliges Funktionssymbol von $\mathcal{L}$ mit $n \geq 1$, so ist $f(t_1, \dots, t_n)$ ein $\mathcal{L}$-Term.

Mit $\mathsf{Ter}_{\mathcal{L}}$ bezeichnen wir die Menge der $\mathcal{L}$-Terme. Für jeden $\mathcal{L}$-Term t definieren wir wie folgt die Menge $\mathsf{vars}(t)$ aller Variablen, die in dem Term auftreten.

– Ist t eine Variable, so ist $\mathsf{vars}(t) := \{t\}$.
– Ist t eine Konstante von $\mathcal{L}$, so ist $\mathsf{vars}(t) := \emptyset$.
– Sind $t_1, \dots, t_n$ $\mathcal{L}$-Terme und ist f ein n-stelliges Funktionssymbol von $\mathcal{L}$, so ist

$$\mathsf{vars}(f(t_1, \dots, t_n)) := \mathsf{vars}(t_1) \cup \ldots \cup \mathsf{vars}(t_n).$$

Ist $\mathsf{vars}(t) = \emptyset$, so heißt der Term t *geschlossen* oder *Grundterm*.

Zweistellige Funktionssymbole werden meist in *Infix*-Schreibweise verwendet. Man schreibt dann $x+y$ statt $+(x, y)$. Klammern werden so gesetzt, daß sich die Struktur eines Terms eindeutig ergibt. Statt $f(t_1, \dots, t_n)$ schreiben wir oft auch $ft_1 \dots t_n$. Terme der Sprache der Gruppentheorie, also $\mathcal{L}_G$-Terme, sind dann zum Beispiel e, $x \circ e$ und $(x^{-1} \circ y^{-1})^{-1}$.

Aus $\mathcal{L}$-Termen erhält man jetzt sehr einfach $\mathcal{L}$-*Primformeln* oder *atomare Formeln* von $\mathcal{L}$: Sind $t_1, \dots, t_n$ $\mathcal{L}$-Terme und ist R ein n-stelliges Relationssymbol von $\mathcal{L}$, so ist $R(t_1, \dots, t_n)$ eine $\mathcal{L}$-Primformel; insbesondere ist $\bot$ eine $\mathcal{L}$-Primformel.

Auch zweistellige Relationssymbole werden meist in *Infix*-Schreibweise verwendet. Man schreibt dann $x < y$ statt $<(x, y)$. Klammern werden so gesetzt, daß sich die Struktur einer atomaren Formel eindeutig ergibt. Statt $R(t_1, \dots, t_n)$ schreiben wir oft $Rt_1 \dots t_n$.

$\mathcal{L}$-*Formeln* werden aus $\mathcal{L}$-Primformeln wieder induktiv definiert durch

– Jede $\mathcal{L}$-Primformel ist eine $\mathcal{L}$-Formel.
– Sind A und B $\mathcal{L}$-Formeln, so auch $(A \wedge B)$ und $(A \to B)$.
– Ist A eine $\mathcal{L}$-Formel und x eine Variable, so ist $\forall x A$ eine $\mathcal{L}$-Formel.

Für jede $\mathcal{L}$-Formel A definieren wir wie folgt die Menge $\mathsf{FV}(A)$ aller in A *freien Variablen*.

$$\begin{aligned}
\mathsf{FV}(R(t_1, \dots, t_n)) &:= \mathsf{vars}(t_1) \cup \ldots \cup \mathsf{vars}(t_n), \\
\mathsf{FV}(A \wedge B) &:= \mathsf{FV}(A) \cup \mathsf{FV}(B), \\
\mathsf{FV}(A \to B) &:= \mathsf{FV}(A) \cup \mathsf{FV}(B), \\
\mathsf{FV}(\forall x A) &:= \mathsf{FV}(A) \setminus \{x\}.
\end{aligned}$$

Falls $\mathsf{FV}(A) = \emptyset$, so heißt A *geschlossene Formel* oder *Satz*. Wir schreiben $A[\boldsymbol{x}]$, um mitzuteilen, daß alle freien Variablen von A in der Liste $\boldsymbol{x}$ enthalten sind.

Wir beschränken uns hier der Einfachheit halber auf die logischen Symbole $\wedge$, $\to$ und $\forall$, da Negation, Disjunktion, Äquivalenz und der Existenzquantor daraus definierbar sind, und zwar durch

$$\begin{aligned}
\neg A &:= A \to \bot, \\
A \vee B &:= \neg(\neg A \wedge \neg B), \\
A \leftrightarrow B &:= (A \to B) \wedge (B \to A), \\
\exists x A &:= \neg \forall x \neg A.
\end{aligned}$$

Wir schreiben $t \neq s$ für $\neg(t = s)$ und $t \not< s$ für $\neg(t < s)$. Klammern werden so gesetzt, daß sich die Struktur einer Formel eindeutig ergibt. Um Klammern zu sparen, vereinbaren wir folgendes.

1. Der Wirkungsbereich von $\neg$ und $\forall x$ ist so klein wie möglich. Beispiele:
$$\begin{array}{rlll}
\neg A \wedge B & \text{meint} & (\neg A) \wedge B, & \text{nicht} \quad \neg(A \wedge B). \\
\forall x A \to B & \text{meint} & (\forall x A) \to B, & \text{nicht} \quad \forall x(A \to B).
\end{array}$$
2. Der Wirkungsbereich von $\wedge$ und $\vee$ ist so klein wie möglich, unter Berücksichtigung von (1). Beispiel:
$$\neg A \wedge B \to C \quad \text{meint} \quad ((\neg A) \wedge B) \to C.$$
3. Bei wiederholter Anwendung einer Verknüpfung ist stets Rechtsklammerung gemeint. Beispiel:
$$A \to B \to C \quad \text{meint} \quad A \to (B \to C).$$
4. Schreibt man hinter $\forall x$ einen Punkt, so ist der Wirkungsbereich von $\forall x$ ist so groß wie möglich. Beispiel:
$$\forall x.A \to B \quad \text{meint} \quad \forall x(A \to B), \quad \text{nicht} \quad (\forall x A) \to B.$$

Als Mitteilungszeichen verwenden wir (auch mit Indizes)

r, s, t	für Terme
x, y, z	für Variablen
c	für Konstanten
P, Q, R	für Relationssymbole
f, g, h	für Funktionssymbole
A, B, C, D	für Formeln

Im folgenden sei die Sprache $\mathcal{L}$ fest gewählt. Wir lassen deshalb den Bezug auf $\mathcal{L}$ weg und sprechen kurz von Termen und Formeln statt von $\mathcal{L}$-Termen und $\mathcal{L}$-Formeln.

Der Einfachheit halber identifizieren wir Formeln, die sich nur durch gebundene Umbenennung unterscheiden. Dann läßt sich die Substitution $A[x := t]$ eines Terms t für eine Variable x besonders einfach definieren.

Definition 1.1.1. Für Terme r und t definieren wir das Ergebnis der *Substitution* von t für x in r wie folgt durch Induktion über den Aufbau von r.

$$\begin{aligned}
y[x := t] &:= \begin{cases} t & \text{falls } x = y \\ y & \text{sonst} \end{cases} \\
c[x := t] &:= c \\
f(t_1, \dots, t_n)[x := t] &:= f(t_1[x := t], \dots, t_n[x := t])
\end{aligned}$$

Für Formeln A und Terme t definieren wir das Ergebnis der Substitution von t für x in A wie folgt durch Induktion über den Aufbau von A.

$$\begin{aligned}
R(t_1, \dots, t_n)[x := t] &:= R(t_1[x := t], \dots, t_n[x := t]) \\
(A \wedge B)[x := t] &:= A[x := t] \wedge B[x := t] \\
(A \to B)[x := t] &:= A[x := t] \to B[x := t] \\
(\forall y A)[x := t] &:= \forall y A[x := t], \quad \text{wobei oBdA } y \neq x \text{ und } y \notin \mathsf{vars}(t).
\end{aligned}$$

1.2 Beweise

Wir wollen jetzt den Begriff einer *logischen Herleitung* einer Formel A definieren. Dazu verwenden wir ein System des “natürlichen Schließens”, das von 1934 von Gentzen eingeführt wurde. Seine besondere Eigenart ist es, daß für jede logische Verknüpfung *Einführungs- und Beseitigungsregeln* vorhanden sind.

Wir beginnen mit einigen Beispielen für natürliche Beweise. Gegeben sei also eine Sprache $\mathcal{L}$ erster Stufe. Der Einfachheit halber betrachten wir Beweise in der reinen Logik, d.h. ohne Annahmen über die Funktionen und Relationen.

$$(A \wedge B \to C) \to (A \to (B \to C)). \tag{1.1}$$

Beweis. Gelte $A \wedge B \to C$. Zu zeigen: $A \to (B \to C)$. Gelte also A. Zu zeigen: $B \to C$. Gelte also B. Zu zeigen: C. Wir haben $A \wedge B$, nach den letzten beiden Annahmen. Also auch C, nach der ersten Annahme. □

$$(A \to (B \to C)) \to (A \wedge B \to C). \tag{1.2}$$

Beweis. Gelte $A \to (B \to C)$. Zu zeigen: $A \wedge B \to C$. Gelte also $A \wedge B$. Zu zeigen: C. Wir haben A, nach der letzten Annahme. Also auch $B \to C$, nach der ersten Annahme. Wir haben B, wieder nach der letzten Annahme. Daher auch C, nach den letzten beiden Aussagen. □

$$(\forall x.A \to B) \to (A \to \forall x B), \quad \text{falls } x \notin \mathsf{FV}(A). \tag{1.3}$$

Beweis. Gelte $\forall x.A \to B$. Zu zeigen: $A \to \forall x B$. Gelte also A. Zu zeigen: $\forall x B$. Sei x beliebig; man beachte, daß wir bisher keine Annahmen über x gemacht haben. Zu zeigen: B. Wir haben $A \to B$, nach der ersten Annahme. Also auch B, nach der zweiten Annahme. □

$$(A \to \forall x B) \to \forall x.A \to B, \quad \text{falls } x \notin \mathsf{FV}(A). \tag{1.4}$$

Beweis. Gelte $A \to \forall x B$. Zu zeigen: $\forall x.A \to B$. Sei x beliebig; man beachte, daß wir bisher keine Annahmen über x gemacht haben. Zu zeigen: $A \to B$. Gelte also A. Zu zeigen: B. Wir haben $\forall x B$, nach der ersten und zweiten Annahme. Also auch B. □

Ein Charakteristikum dieser Beweise ist es, daß Annahmen eingeführt und wieder beseitigt werden: Zu jedem Zeitpunkt im Beweis kennt man die jetzt freien Annahmen.

Wir reservieren das Wort *Beweis* für einen informalen Kontext (oder die “Meta-Stufe”); eine formale Darstellung eines Beweises nennen wir *Herleitung* oder *Ableitung*.

Eine anschauliche Art, Herleitungen zu definieren, besteht darin, sie als beschriftete Bäume aufzufassen. Die Beschriftungen der inneren Knoten sind Formeln, die der Blätter sind Formeln oder Terme. Die Beschriftungen der Nachfolger eines Knotens ν sind die *Prämissen* einer Regelanwendung, die Formel am Knoten ν ist ihre *Konklusion*. An der Wurzel des Baums befindet sich die Konklusion der gesamten Herleitung. Im natürlichen Schließen arbeitet man mit *Annahmen*, die an Blättern des Baums stehen; sie können *offen* oder *geschlossen* (man sagt auch: *gestrichen*) sein.

Jede dieser Annahmen ist mit einer *Marke* versehen. Als Marken verwenden wir *Annahmenvariablen* $\square_0, \square_1, \dots$; Mitteilungszeichen für Annahmenvariablen sind $u, v, w, u_0, u_1, \dots$. Die (bisherigen) Variablen nennen wir oft auch *Objektvariablen*, um sie von den Annahmenvariablen zu unterscheiden. Wenn an einer späteren Stellen (d.h. an einem Knoten im Baum unterhalb einer solchen Annahme) die Abhängigkeit von dieser Annahme beseitigt wird, notieren wir dies durch Angabe der Annahmenvariablen. Da wir dieselbe Annahme auch mehrfach verwenden können (dies war etwa im Beispiel 1.2 der Fall), darf in einem Baum eine mit u markierte Annahme A (mitgeteilt durch $u\colon A$) auch mehrfach vorkommen. Wir verlangen jedoch, daß verschiedene Annahmeformeln stets verschiedene Marken bekommen.

Einen inneren Knoten im Baum verstanden wir als Resultat eines Übergangs von *Prämissen* zu einer *Konklusion*. Die zulässigen Übergänge werden durch die *Regeln* bestimmt. Die Beschriftung des Knotens enthält dann neben der Konklusion noch den Namen der verwendeten Regel. In manchen Fällen bindet eine Regel eine Annahmenvariable u (und beseitigt damit die Abhängigkeit von allen darüber stehenden, mit u markierten Annahmen $u\colon A$) oder eine Objektvariable x (und beseitigt damit die Abhängigkeit von x). Dann wird die abgebundene Annahmen- oder Objektvariable der Beschriftung hinzugefügt.

Wir geben jetzt die Regeln des natürlichen Schließens an. Zunächst haben wir eine Annahmeregel, die es gestattet, eine beliebige mit einer Marke u versehene Formel A als Annahme hinzuschreiben:

$$u\colon A \quad \text{Annahme}$$

Die restlichen Regeln des natürlichen Schließens gliedern sich in Einführungs- und Beseitigungsregeln für die logischen Verknüpfungen $\wedge$, $\to$ und $\forall$. Für die Konjuktion $\wedge$ haben wir eine Einführungsregel $\wedge^+$ und zwei Beseitigungsregeln $\wedge_0^-$ und $\wedge_1^-$.

$$\dfrac{\begin{matrix}\mathcal{D}_0\\ A\end{matrix}\qquad\begin{matrix}\mathcal{D}_1\\ B\end{matrix}}{A\wedge B}\,\wedge^+ \qquad\qquad \dfrac{\begin{matrix}\mathcal{D}\\ A\wedge B\end{matrix}}{A}\,\wedge_0^- \qquad\qquad \dfrac{\begin{matrix}\mathcal{D}\\ A\wedge B\end{matrix}}{B}\,\wedge_1^-$$

Für die Implikation $\to$ gibt es eine Einführungsregel $\to^+u$ und eine Beseitigungsregel $\to^-$, die man auch *modus ponens* nennt. Die linke Prämisse $A\to B$ in $\to^-$ nennt man *Hauptprämisse*, die rechte Prämisse A *Nebenprämisse*. Man beachte, daß bei Anwendung einer $\to^+u$-Regel *alle* darüber stehenden mit u markierten Annahmen A gestrichen werden.

$$\dfrac{\begin{matrix}[u\colon A]\\ \mathcal{D}\\ B\end{matrix}}{A\to B}\,\to^+u \qquad\qquad \dfrac{\begin{matrix}\mathcal{D}_0\\ A\to B\end{matrix}\qquad\begin{matrix}\mathcal{D}_1\\ A\end{matrix}}{B}\,\to^-$$

Für den Allquantor $\forall$ gibt es eine Einführungsregel $\forall^+x$ und eine Beseitigungsregel $\forall^-$, die als rechte Prämisse den zu substituierenden Term t hat. Die Regel $\forall^+x$ unterliegt der folgenden *Variablenbedingung*: Die Herleitung $\mathcal{D}$ der Prämisse A darf keine offenen Annahmen enthalten, in denen x frei vorkommt.

$$\dfrac{\begin{matrix}\mathcal{D}\\ A\end{matrix}}{\forall xA}\,\forall^+x \qquad\qquad \dfrac{\begin{matrix}\mathcal{D}\\ \forall xA\end{matrix}\qquad t}{A[x:=t]}\,\forall^-$$

Wir geben Herleitungen für die oben bewiesenen Formeln (1.1) – (1.4) an. Da meist die verwendete Regel durch die an den Knoten stehenden Formeln bestimmt ist, verzichten wir im allgemeinen auf die Angabe der Regel.

$$\dfrac{\dfrac{\dfrac{\dfrac{u\colon A\wedge B\to C \qquad \dfrac{v\colon A \qquad w\colon B}{A\wedge B}}{C}}{B\to C}\,\to^+w}{A\to(B\to C)}\,\to^+v}{(A\wedge B\to C)\to(A\to(B\to C))}\,\to^+u \tag{1.1}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{u\colon A\to(B\to C) \qquad \dfrac{v\colon A\wedge B}{A}}{B\to C} \qquad \dfrac{v\colon A\wedge B}{B}}{C}}{A\wedge B\to C}\,\to^+v}{(A\to(B\to C))\to(A\wedge B\to C)}\,\to^+u \tag{1.2}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{u\colon \forall x.A\to B \qquad x}{A\to B} \qquad v\colon A}{B}}{\forall xB}\,x}{A\to\forall xB}\,\to^+v}{(\forall x.A\to B)\to(A\to\forall xB)}\,\to^+u \tag{1.3}$$

Hier ist zu beachten, daß die Variablenbedingung erfüllt ist: x kommt nicht frei in A (und auch nicht frei in $\forall x.A \to B$) vor.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{u\colon A \to \forall x B \qquad v\colon A}{\forall x B} \qquad x}{B}}{A \to B}\,\to^+ v}{\forall x.A \to B}\,x}{(A \to \forall x B) \to \forall x.A \to B}\,\to^+ u \tag{1.4}$$

Auch hier ist die Variablenbedingung erfüllt: x kommt nicht frei in A vor.

Wir schreiben $\vdash A$ und nennen A *herleitbar* (in der *Minimallogik*), wenn es eine Herleitung von A ohne freie Annahmen gibt. Eine Formel B heißt herleitbar aus den Annahmen $A_1, \dots, A_n$, wenn es eine Herleitung mit freien Annahmen unter $A_1, \dots, A_n$ gibt. Sei Γ eine (endliche oder unendliche) Menge von Formeln. Wir schreiben $\Gamma \vdash B$, wenn die Formel B aus endlich vielen Annahmen $A_1, \dots, A_n \in \Gamma$ herleitbar ist.

Sequenzenformulierung des natürlichen Schließens

Um die an einem Knoten ν eines Herleitungsbaums freien Annahmen zu finden, muß man die Annahmen an den Blättern oberhalb von ν durchsuchen. Diejenigen davon, deren Marke nicht zwischen dem Blatt und ν gestrichen wurde, sind an dem Knoten ν frei. Obwohl schreibtechnisch aufwendig ist es für manche metamathematischen Untersuchungen bequem, die an einem Knoten freien Annahmen explizit mitzuführen. Man nennt die Menge der an einem Knoten freien Annahmen den *Kontext*. Ein Kontext ist also eine Menge $\{u_1\colon A_1, u_2\colon A_2, \dots, u_n\colon A_n\}$ mit paarweise verschiedenen u_i. Die A_i brauchen nicht verschieden zu sein; dies entspricht der Tatsache, daß in einer natürlichen Herleitung dieselbe Annahmeformel mehrmals mit verschiedenen Marken vorkommen kann. Die Herleitungen werden dann Bäume, an denen jeder Knoten beschriftet ist mit einer Sequenz der Form $\Gamma \Rightarrow B$, wobei Γ ein Kontext ist.

Mit Γ, Δ oder auch $\Gamma\Delta$ bezeichnen wir die Vereinigung $\Gamma \cup \Delta$. Die Verwendung dieser Bezeichnung soll stets die Voraussetzung implizieren, daß die Vereinigung *konsistent*, also wieder ein Kontext ist.

$$u\colon A \Rightarrow A \quad \text{Annahme}$$

$$\dfrac{\Gamma \Rightarrow A \qquad \Delta \Rightarrow B}{\Gamma\Delta \Rightarrow A \wedge B}\,\wedge^+ \qquad \dfrac{\Gamma \Rightarrow A \wedge B}{\Gamma \Rightarrow A}\,\wedge_0^- \qquad \dfrac{\Gamma \Rightarrow A \wedge B}{\Gamma \Rightarrow B}\,\wedge_1^-$$

$$\dfrac{\Gamma \Rightarrow B}{\Gamma \setminus \{u\colon A\} \Rightarrow A \to B}\,\to^+ u \qquad \dfrac{\Gamma \Rightarrow A \to B \qquad \Delta \Rightarrow A}{\Gamma\Delta \Rightarrow B}\,\to^-$$

$$\dfrac{\Gamma \Rightarrow A}{\Gamma \Rightarrow \forall x A}\,\forall^+ \qquad \dfrac{\Gamma \Rightarrow \forall x A \qquad t}{\Gamma \Rightarrow A[x := t]}\,\forall^-$$

Die Variablenbedingung bei der Regel $\forall^+$ ist jetzt einfach, daß x nicht frei in einer der Formeln des Kontexts Γ vorkommt.

Intuitionistischen und klassische Logik

In unserer $\to\wedge\forall$-Sprache erhalten wir die *intuitionistische Logik*, indem wir gewisse zusätzliche Annahmen verwenden, und zwar die sogenannten *Ex-Falso-Quodlibet*-Formeln (oder “Axiome”) Efq_R für jedes von $\bot$ verschiedene Relationssymbol R

$$\forall \boldsymbol{x}.\bot \to R\boldsymbol{x} \tag{Efq_R}$$

Ähnlich erhält man die *klassische Logik*: wir nehmen für jedes jedes von $\bot$ verschiedene Relationssymbol R das *Prinzip des indirekten Beweisens* für R als zusätzliche Annahme hinzu, also die Formel

$$\forall \boldsymbol{x}.\neg\neg R\boldsymbol{x} \to R\boldsymbol{x}; \tag{Stab_R}$$

diese Formel bezeichnet man auch als *Stabilität* von R.

Man beachte, daß mit $\bot$ für R beide Formeln trivialerweise herleitbar sind; z.B. für die Stabilität haben wir $\neg\neg\bot \to \bot = ((\bot \to \bot) \to \bot) \to \bot$. Die gesuchte Herleitung ist

$$\dfrac{v\colon (\bot \to \bot) \to \bot \qquad \dfrac{u\colon \bot}{\bot \to \bot}\,\to^+ u}{\bot}$$

Sei

$$\begin{aligned}\mathsf{Efq} &:= \{\, \mathsf{Efq}_R \mid R \text{ Relationssymbol} \neq \bot \,\},\\ \mathsf{Stab} &:= \{\, \mathsf{Stab}_R \mid R \text{ Relationssymbol} \neq \bot \,\}.\end{aligned}$$

Wir nennen die Formel A *klassisch (intuitionistisch) herleitbar* und schreiben $\vdash_c A$ ($\vdash_i A$), wenn es eine Herleitung von A aus Stabilitätsannahmen Stab_R (Ex-Falso-Quodlibet Annahmen Efq_R) gibt. Ebenso definieren wir klassische (intuitionistische) Herleitbarkeit aus Γ und schreiben $\Gamma \vdash_c A$ ($\Gamma \vdash_i A$), also

$$\begin{aligned}\Gamma \vdash_i A \quad &:\Longleftrightarrow \quad \Gamma \cup \mathsf{Efq} \vdash A,\\ \Gamma \vdash_c A \quad &:\Longleftrightarrow \quad \Gamma \cup \mathsf{Stab} \vdash A.\end{aligned}$$

Lemma 1.2.1. *(Ex-falso-quodlibet). Für jede Formel A gilt $\vdash_i \bot \to A$.*

Beweis. Durch Induktion über A konstruieren wir für jede Formel A eine Herleitung $\mathcal{D}_A$ von $\bot \to A$. *Fall $R\boldsymbol{t}$.* Mit Efq_R. *Fall $A \land B$.*

$$\dfrac{\dfrac{\dfrac{\begin{matrix}\mathcal{D}_A\\ \bot \to A\end{matrix} \qquad u\colon \bot}{A} \qquad \dfrac{\begin{matrix}\mathcal{D}_B\\ \bot \to B\end{matrix} \qquad u\colon \bot}{B}}{A \land B}}{\bot \to A \land B}\,\to^+ u$$

Fall $A \to B$.

$$\dfrac{\dfrac{\dfrac{\begin{matrix}\mathcal{D}_B\\ \bot \to B\end{matrix} \qquad u\colon \bot}{B}}{A \to B}}{\bot \to A \to B}\,\to^+ u$$

Fall $\forall x A$.

$$\dfrac{\dfrac{\dfrac{\begin{matrix}\mathcal{D}_A\\ \bot \to A\end{matrix} \qquad u\colon \bot}{A}}{\forall x A}}{\bot \to \forall x A}\,\to^+ u$$

□

Lemma 1.2.2. *(Stabilität). Für jede Formel A (unserer $\to\land\forall$-Sprache) gilt $\vdash_c \neg\neg A \to A$.*

Beweis. Induktion über A. In den konstruierten Herleitungen lassen wir der Kürze halber Anwendungen von $\to^+$ am Schluß fort. *Fall $R\boldsymbol{t}$.* Mit Stab_R. *Fall $A \land B$.* Mit $\vdash (\neg\neg A \to A) \to (\neg\neg B \to B) \to \neg\neg(A \land B) \to A \land B$, was leicht aus $\vdash \neg\neg(A \land B) \leftrightarrow (\neg\neg A \land \neg\neg B)$ folgt. *Fall $A \to B$.* Mit $\vdash (\neg\neg B \to B) \to \neg\neg(A \to B) \to A \to B$. Eine Herleitung ist

$$\dfrac{u\colon \neg\neg B \to B \qquad \dfrac{\dfrac{v\colon \neg\neg(A \to B) \qquad \dfrac{\dfrac{u_1\colon \neg B \qquad \dfrac{u_2\colon A \to B \qquad w\colon A}{B}}{\bot}}{\neg(A \to B)}\,\to^+ u_2}{\bot}}{\neg\neg B}\,\to^+ u_1}{B}$$

Fall $\forall x A$. Offenbar genügt es zu zeigen, daß $\vdash (\neg\neg A \to A) \to \neg\neg\forall x A \to A$. Eine Herleitung ist

$$\dfrac{u\colon \neg\neg A \to A \qquad \dfrac{\dfrac{v\colon \neg\neg\forall x A \qquad \dfrac{\dfrac{u_1\colon \neg A \qquad \dfrac{u_2\colon \forall x A \qquad x}{A}}{\bot}}{\neg\forall x A}\,\to^+ u_2}{\bot}}{\neg\neg A}\,\to^+ u_1}{A}$$ □

Lemma 1.2.3. *$\Gamma \vdash A \implies \Gamma \vdash_i A$ und $\Gamma \vdash_i A \implies \Gamma \vdash_c A$.*

Beweis. Es genügt zu zeigen, daß $\vdash_c \mathsf{Efq}_R$. Dies sieht man wie folgt; R sei etwa einstellig.

$$\dfrac{\dfrac{\dfrac{\dfrac{\forall x.\neg\neg Rx \to Rx \qquad x}{\neg\neg Rx \to Rx} \qquad \dfrac{u\colon \bot}{\neg\neg Rx}\,\to^+ v^{\neg Rx}}{Rx}}{\bot \to Rx}\,\to^+ u}{\forall x.\bot \to Rx}\,\forall^+$$ □

Die Umkehrungen gelten jedoch nicht; Gegenbeispiele sind:

$$\begin{array}{ll} \nvdash \bot \to P, & \text{aber } \vdash_i \bot \to P, \\ \nvdash_i ((P \to Q) \to P) \to P, & \text{aber } \vdash_c ((P \to Q) \to P) \to P \end{array}$$

$\vdash_i \bot \to P$ folgt aus Lemma 1.2.1, und die Peirce-Formel $((P \to Q) \to P) \to P$ läßt sich leicht klassisch herleiten. Die negative Aussagen erfordern ein genaueres Studium der Herleitbarkeit. Wir werden in Abschnitt 1.4 einen Beweis geben.

Lemma 1.2.4. *(Fallunterscheidung).* $\vdash_c (A \to B) \to (\neg A \to B) \to B$.

Beweis.

$$\dfrac{\begin{matrix}\mathcal{D}_{\mathsf{Stab}} \\ \neg\neg B \to B\end{matrix} \qquad \dfrac{\dfrac{w\colon \neg B \qquad \dfrac{u_2\colon \neg A \to B \qquad \dfrac{\dfrac{w\colon \neg B \qquad \dfrac{u_1\colon A \to B \qquad v\colon A}{B}}{\bot}}{\neg A}\,\to^+ v}{B}}{\bot}}{\neg\neg B}\,\to^+ w}{B}$$

wobei $\mathcal{D}_{\mathsf{Stab}}$ gemäß dem Stabilitätslemma 1.2.2 gewählt ist. □

Wir nennen zwei Formeln A und B *äquivalent* in der Minimallogik bzw. in der klassischen oder intuitionistischen Logik, wenn $\vdash A \leftrightarrow B$ bzw. $\vdash_c A \leftrightarrow B$ oder $\vdash_i A \leftrightarrow B$.

Lemma 1.2.5. *(Äquivalenzlemma). Für $\vdash_{mic} \in \{\vdash, \vdash_i, \vdash_c\}$ gilt folgendes. Ist $\vdash_{mic} A_1 \leftrightarrow A_2$ und entsteht B_2 aus B_1 durch Ersetzen eines Teils A_1 von B_1 durch A_2, so gilt auch $\vdash_{mic} B_1 \leftrightarrow B_2$.*

Beweis. Induktion über B_1. Falls ganz B_1 ersetzt wird, so ist die Behauptung klar. Andernfalls muß B_1 eine zusammengesetzte Formel sein.

Fall $C_1 \wedge D_1$. Die Ersetzung finde etwa in C_1 statt. Zu zeigen ist dann $\vdash_{mic} C_1 \wedge D_1 \leftrightarrow C_2 \wedge D_1$. $\to$:

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D} \\ C_1 \to C_2\end{matrix} \qquad \dfrac{C_1 \wedge D_1}{C_1}}{C_2} \qquad \dfrac{C_1 \wedge D_1}{D_1}}{C_2 \wedge D_1}$$

wobei $\mathcal{D}$ nach IH bekannt ist. $\leftarrow$ beweist man ähnlich.

Fall $C_1 \to D_1$. Falls die Ersetzung in C_1 stattfindet, ist zu zeigen $\vdash_{mic} (C_1 \to D_1) \leftrightarrow (C_2 \to D_1)$. $\to$:

$$\dfrac{\dfrac{C_1 \to D_1 \qquad \dfrac{\begin{matrix}\mathcal{D}\\ C_2 \to C_1\end{matrix} \qquad u\colon C_2}{C_1}}{D_1}}{C_2 \to D_1}\,\to^+ u$$

wobei wieder $\mathcal{D}$ nach IH bekannt ist. $\leftarrow$ beweist man ähnlich. Falls die Ersetzung in D_1 stattfindet, ist zu zeigen $\vdash_{mic} (C_1 \to D_1) \leftrightarrow (C_1 \to D_2)$. $\to$:

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}\\ D_1 \to D_2\end{matrix} \qquad \dfrac{C_1 \to D_1 \qquad u\colon C_1}{D_1}}{D_2}}{C_1 \to D_2}\,\to^+ u$$

wobei wieder $\mathcal{D}$ nach IH bekannt ist. $\leftarrow$ beweist man ähnlich.

Fall $\forall x C_1$. Zu zeigen ist $\vdash_{mic} \forall x C_1 \leftrightarrow \forall x C_2$. $\to$:

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}\\ C_1 \to C_2\end{matrix} \qquad \dfrac{\forall x C_1 \qquad x}{C_1}}{C_2}}{\forall x C_2}$$

wobei wieder $\mathcal{D}$ nach IH bekannt ist. Man beachte, daß $\mathcal{D}$ keine freien Annahmen enthält. $\leftarrow$ beweist man ähnlich. □

Einbettung der intuitionistischen und klassischen Logik in die Minimallogik

Nachdem wir die klassische und die intuitionistische Logik definiert haben, wollen wir jetzt zeigen, daß beide Logiken in die Minimallogik eingebettet werden können. Dies mag verwunderlich erscheinen; es folgt wesentlich aus der Tatsache, daß wir uns auf eine Sprache beschränkt haben, die nur die Verknüpfungen $\{\to, \land, \forall\}$ enthält.

Eine Formel A (unserer $\to\land\forall$-Sprache) heißt *negativ*, wenn jede atomare Formel $\neq \bot$ in A negiert vorkommt, d.h. in einem Kontext $R\boldsymbol{t} \to \bot$.

Lemma 1.2.6. *Für negative A gilt $\vdash \neg\neg A \to A$.*

Beweis. Dies beweist man wie das Stabilitätslemma 1.2.2 durch Induktion über A, wobei man anstelle der Stabilitätsannahmen verwendet $\vdash \neg\neg\neg R\boldsymbol{t} \to \neg R\boldsymbol{t}$. □

Definition 1.2.7. (Negative Übersetzung ^{g} nach Gödel-Gentzen).

$$\begin{aligned} R\boldsymbol{t}^g &:= \neg\neg R\boldsymbol{t} \quad \text{für } R \neq \bot, \\ \bot^g &:= \bot, \\ (A \land B)^g &:= A^g \land B^g, \\ (A \to B)^g &:= A^g \to B^g, \\ (\forall x A)^g &:= \forall x A^g. \end{aligned}$$

Satz 1.2.8. *Für alle Formeln A gilt*

1. $\vdash_c A \leftrightarrow A^g$,
2. *$\Gamma \vdash_c A$ genau dann, wenn $\Gamma^g \vdash A^g$, wobei $\Gamma^g := \{ B^g \mid B \in \Gamma \}$.*

Beweis. Der erste Teil folgt sofort aus dem Äquivalenzlemma 1.2.5. Für den zweiten Teil ist die Richtung von rechts nach links klar. Für die andere Richtung argumentieren wir durch Induktion nach der klassischen Herleitung. Für eine Stabilitätsannahme $\neg\neg R\boldsymbol{t} \to R\boldsymbol{t}$ gilt $(\neg\neg R\boldsymbol{t} \to R\boldsymbol{t})^g = \neg\neg\neg\neg R\boldsymbol{t} \to \neg\neg R\boldsymbol{t}$, und dies ist leicht herleitbar. *Fall* $\to^+$. Gelte

$$\dfrac{\begin{matrix}[u\colon A]\\ \mathcal{D}\\ B\end{matrix}}{A \to B}\to^+ u$$

Dann haben wir nach IH

$$\begin{matrix}u\colon A^g\\ \mathcal{D}^g\\ B^g\end{matrix} \qquad \text{also} \qquad \dfrac{\begin{matrix}[u\colon A^g]\\ \mathcal{D}^g\\ B^g\end{matrix}}{A^g \to B^g}\to^+ u$$

Fall $\to^-$. Gelte

$$\dfrac{\begin{matrix}\mathcal{D}_0\\ A \to B\end{matrix} \qquad \begin{matrix}\mathcal{D}_1\\ A\end{matrix}}{B}$$

Dann haben wir nach IH

$$\begin{matrix}\mathcal{D}_0^g\\ A^g \to B^g\end{matrix} \qquad \begin{matrix}\mathcal{D}_1^g\\ A^g\end{matrix} \qquad \text{also} \qquad \dfrac{\begin{matrix}\mathcal{D}_0^g\\ A^g \to B^g\end{matrix} \qquad \begin{matrix}\mathcal{D}_1^g\\ A^g\end{matrix}}{B^g}$$

Die restlichen Fälle behandelt man ähnlich. □

Korollar 1.2.9. *(Einbettung der klassischen Logik in die Minimallogik). Für negative A gilt $\vdash_c A$ genau dann, wenn $\vdash A$.*

Beweis. Nach dem Satz haben wir $\vdash_c A$ genau dann, wenn $\vdash A^g$. Da A negativ ist, muß jedes Atom $\neq \bot$ in A negiert vorkommen, ist also in A^g dreifach negiert (als $\neg\neg\neg R\boldsymbol{t}$). Die Behauptung folgt aus $\vdash \neg\neg\neg R\boldsymbol{t} \leftrightarrow \neg R\boldsymbol{t}$. □

Da jede Formel klassisch äquivalent zu einer negativen Formel ist, haben wir damit eine Einbettung der klassischen in die Minimallogik erreicht.

Man beachte, daß $\nvdash \neg\neg P \to P$ (wie wir in Abschnitt 1.4 zeigen werden). Das Korollar gilt also nicht für alle Formeln A.

Abgeleitete Regeln für Disjunktion und Existenz

Disjunktion und Existenz hatten wir definiert durch

$$\begin{aligned} A \vee B &:= \neg(\neg A \wedge \neg B),\\ \exists x A &:= \neg\forall x \neg A.\end{aligned}$$

Herleitungen von Formeln mit $\vee, \exists$ lassen sich jedoch oft leichter finden, wenn man nicht auf diese Definitionen zurückgeht, sondern direkt mit "*abgeleiteten Regeln*" für $\vee, \exists$ arbeitet. Darunter verstehen wir die folgenden Regeln.

Für die Disjunktion $\vee$ gibt es zwei Einführungsregeln $\vee_0^+, \vee_1^+$ und eine Beseitigungsregel $\vee^- uv$.

$$\dfrac{\begin{matrix}\mathcal{D}\\ A\end{matrix}}{A \vee B}\vee_1^+ \qquad \dfrac{\begin{matrix}\mathcal{D}\\ B\end{matrix}}{A \vee B}\vee_0^+ \qquad \dfrac{\begin{matrix}\mathcal{D}\\ A \vee B\end{matrix} \qquad \begin{matrix}[u\colon A]\\ \mathcal{D}_0\\ C\end{matrix} \qquad \begin{matrix}[v\colon B]\\ \mathcal{D}_1\\ C\end{matrix}}{C}\vee^- uv$$

Für den Existenzquantor $\exists$ gibt es eine Einführungsregel $\exists^+$, die als rechte Prämisse den zu substituierenden Term t hat, und eine Beseitigungsregel $\exists^- u$. Die Regel $\exists^- u$ unterliegt der folgenden *Variablenbedingung*: Die Herleitung $\mathcal{D}_0$ darf keine offenen Annahmen außer $u\colon A$ enthalten, in denen x frei vorkommt, und ferner darf B die Variable x nicht frei enthalten.

$$\dfrac{\begin{matrix}\mathcal{D}\\ A[x := t]\end{matrix} \qquad t}{\exists x A}\exists^+ \qquad \dfrac{\begin{matrix}\mathcal{D}\\ \exists x A\end{matrix} \qquad \begin{matrix}[u\colon A]\\ \mathcal{D}_0\\ B\end{matrix}}{B}\exists^- u$$

Der doppelte Strich deutet hierbei an, daß es sich nicht um eine primitive Regel unserer Logik handelt, sondern daß man an dieser Stelle einen Herleitungsteil so einsetzen kann, daß eine korrekt gebildete Herleitung entsteht. Wir geben im folgenden für jede der abgeleiteten Regeln den einzusetzenden Herleitungsteil an. Man beachte, daß es sich um Herleitungen in der klassischen Logik handelt, da Stabilitätsannahmen verwendet werden.

$\lor_0^+$ ($\lor_1^+$ wird entsprechend behandelt):

$$\dfrac{\dfrac{\dfrac{u\colon \neg A \land \neg B}{\neg A} \qquad \begin{matrix}\mathcal{D}\\ A\end{matrix}}{\bot}}{\neg(\neg A \land \neg B)}\to^+ u$$

$\lor^- uv$:

$$\dfrac{\begin{matrix}\mathcal{D}_{\mathsf{Stab}}\\ \neg\neg C \to C\end{matrix} \qquad \dfrac{\dfrac{\begin{matrix}\mathcal{D}\\ \neg(\neg A \land \neg B)\end{matrix} \qquad \dfrac{\dfrac{\dfrac{w\colon \neg C \qquad \begin{matrix}[u\colon A]\\ \mathcal{D}_0\\ C\end{matrix}}{\bot}}{\neg A}\to^+ u \qquad \dfrac{\dfrac{w\colon \neg C \qquad \begin{matrix}[v\colon B]\\ \mathcal{D}_1\\ C\end{matrix}}{\bot}}{\neg B}\to^+ v}{\neg A \land \neg B}}{\bot}}{\neg\neg C}\to^+ w}{C}$$

$\exists^+$:

$$\dfrac{\dfrac{\dfrac{u\colon \forall x \neg A \qquad t}{\neg A[x := t]} \qquad \begin{matrix}\mathcal{D}\\ A[x := t]\end{matrix}}{\bot}}{\neg\forall x \neg A}\to^+ u$$

$\exists^- u$:

$$\dfrac{\begin{matrix}\mathcal{D}_{\mathsf{Stab}}\\ \neg\neg B \to B\end{matrix} \qquad \dfrac{\dfrac{\begin{matrix}\mathcal{D}\\ \neg\forall x \neg A\end{matrix} \qquad \dfrac{\dfrac{\dfrac{v\colon \neg B \qquad \begin{matrix}[u\colon A]\\ \mathcal{D}_0\\ B\end{matrix}}{\bot}}{\neg A}\to^+ u}{\forall x \neg A}}{\bot}}{\neg\neg B}\to^+ v}{B}$$

Als Alternative kann man ebenfalls den Komfort der abgeleiteten Regeln für $\lor, \exists$ erhalten, indem man die folgenden (ableitbaren) Schemata als Axiome benutzt.

$$\begin{aligned}
&A \to A \lor B, \quad B \to A \lor B.\\
&A \lor B \to (A \to C) \to (B \to C) \to C.\\
&A \to \exists x A.\\
&\exists x A \to (\forall x. A \to B) \to B, \quad \text{falls } x \notin \mathsf{FV}(B).
\end{aligned}$$

Konjunktive und disjunktive Normalform

Lemma 1.2.10. *(Assoziativität, Kommutativität und Distributivität von $\land$ und $\lor$).*

1. $\vdash_c A \land (B \land C) \leftrightarrow (A \land B) \land C$.
2. $\vdash_c A \lor (B \lor C) \leftrightarrow (A \lor B) \lor C$.
3. $\vdash_c A \land B \leftrightarrow B \land A$.
4. $\vdash_c A \lor B \leftrightarrow B \lor A$.
5. $\vdash_c A \land (B \lor C) \leftrightarrow (A \land B) \lor (A \land C)$.
6. $\vdash_c A \lor (B \land C) \leftrightarrow (A \lor B) \land (A \lor C)$.

Beweis. Übung. □

Lemma 1.2.11. *1. (DE MORGAN-Äquivalenzen).*

$$\vdash_c \neg(A\wedge B)\leftrightarrow\neg A\vee\neg B,$$
$$\vdash_c \neg(A\vee B)\leftrightarrow\neg A\wedge\neg B.$$

2. ($\to$ aus $\neg,\wedge$ bzw. $\neg,\vee$).

$$\vdash_c (A\to B)\leftrightarrow\neg(A\wedge\neg B),$$
$$\vdash_c (A\to B)\leftrightarrow\neg A\vee B.$$

3. (Tertium non datur). $\vdash_c A\vee\neg A$.

Beweis. 1. Die Behauptung ergibt sich aus der Definition von $\vee$, dem Stabilitätslemma 1.2.2 und dem Äquivalenzlemma 1.2.5.

2. $\vdash_c (A\to B)\leftrightarrow\neg(A\wedge\neg B)$. $\to$.

$$\dfrac{\dfrac{v\colon A\wedge\neg B}{\neg B}\qquad\dfrac{u\colon A\to B\qquad\dfrac{v\colon A\wedge\neg B}{A}}{B}}{\bot}$$

$\leftarrow$.

$$\dfrac{\dfrac{\overset{\mathcal{D}_{\mathsf{Stab}}}{\neg\neg B\to B}\qquad\dfrac{\dfrac{u\colon\neg(A\wedge\neg B)\qquad\dfrac{v\colon A\qquad w\colon\neg B}{A\wedge\neg B}}{\bot}}{\neg\neg B}\to^+w}{B}}{A\to B}\to^+v$$

$\vdash_c (A\to B)\leftrightarrow\neg A\vee B$. $\to$.

$$\dfrac{\dfrac{v\colon\neg\neg A\wedge\neg B}{\neg B}\qquad\dfrac{u\colon A\to B\qquad\dfrac{\overset{\mathcal{D}_{\mathsf{Stab}}}{\neg\neg A\to A}\qquad\dfrac{v\colon\neg\neg A\wedge\neg B}{\neg\neg A}}{A}}{B}}{\bot}$$

$\leftarrow$. Hier kann man die abgeleitete Regel $\vee^-$ verwenden; ferner benötigt man das Ex-Falso-Quodlibet-Lemma 1.2.1.

3. Dies ergibt sich sofort aus der Definition von $\vee$:

$$\dfrac{\dfrac{u\colon\neg A\wedge\neg\neg A}{\neg\neg A}\qquad\dfrac{u\colon\neg A\wedge\neg\neg A}{\neg A}}{\bot}$$ □

Zur Formulierung des nächsten Satzes benötigen wir den Begriff der *Quantorentiefe* $\mathsf{qt}(A)$ einer Formel A; er ist durch Rekursion über A definiert wie folgt.

$$\begin{aligned}\mathsf{qt}(R\boldsymbol{t})&:=0,\\ \mathsf{qt}(A\wedge B)&:=\max(\mathsf{qt}(A),\mathsf{qt}(B))+1,\\ \mathsf{qt}(A\to B)&:=\max(\mathsf{qt}(A),\mathsf{qt}(B))+1,\\ \mathsf{qt}(\forall xA)&:=\mathsf{qt}(A)+1.\end{aligned}$$

Satz 1.2.12. *(Konjunktive und disjunktive Normalform).*

1. Jede Formel A ist klassisch äquivalent zu einer Formel der Gestalt

$$\bigwedge_{i=1}^{n}\bigvee_{j=1}^{n_i}A_{ij}\qquad\text{(konjunktive Normalform)},$$

wobei die A_{ij} aussagenlogisch unzerlegbare Formeln (d.h. atomare Formeln oder von der Form $\forall xA$) oder Negationen davon sind, und $\mathsf{qt}(A_{ij})\le\mathsf{qt}(A)$.

2. Jede Formel ist A klassisch äquivalent zu einer Formel der Gestalt

$$\bigvee_{i=1}^{n} \bigwedge_{j=1}^{n_i} A_{ij} \qquad \text{(disjunktive Normalform)},$$

wobei wieder die A_{ij} aussagenlogisch unzerlegbare Formeln oder Negationen davon sind, und $\mathsf{qt}(A_{ij}) \leq \mathsf{qt}(A)$.

Beweis. Wir zeigen (1) und (2) gemeinsam durch Induktion über die gegebene Formel, unter Verwendung des Äquivalenzlemmas 1.2.5.

Fall Atomare Formel, $\forall x A$. In diesem Fall ist nichts zu zeigen.

Fall $A \wedge B$. Für (1) verwendet man die konjunktiven Normalformen von A und von B. (2) ergibt sich aus den disjunktiven Normalformen von A und von B durch Ausdistribuieren.

Fall $A \to B$. Die Behauptung ergibt sich aus $\vdash_c (A \to B) \leftrightarrow \neg A \vee B$ und den de Morgan-Äquivalenzen, und zwar für (1) unter Verwendung der disjunktiven Normalform von A und der konjunktiven Normalform von B und für (2) unter Verwendung der konjunktiven Normalform von A und der disjunktiven Normalform von B, wobei in (1) noch auszudistribuieren ist. □

Eine konjunktive Normalform von A (die offenbar nicht eindeutig bestimmt ist) nennt man auch eine *Klauselform* von A. Da eine Konjunktion genau dann herleitbar ist, wenn dies für jedes Konjunktionsglied gilt, kann man sich also bei der Suche nach Herleitungen von Formeln auf sogenannte Klauseln, also Disjunktionen von aussagenlogisch unzerlegbaren Formeln beschränken. Man beachte, daß die Herstellung etwa der konjunktiven Normalform einer gegebenen Formel im allgemeinen exponentiell viele (in der Länge der Formel) Rechenschritte erfordert, da bei den Distributivgesetzen Formeln dupliziert werden.

Pränexe Normalform

Lemma 1.2.13. *(Herausziehen von Quantoren). Sei $x \notin \mathsf{FV}(A)$. Dann gilt*

1. $\vdash_c A \wedge \forall xB \leftrightarrow \forall x.A \wedge B$.
2. $\vdash_c A \wedge \exists xB \leftrightarrow \exists x.A \wedge B$.
3. $\vdash_c (A \to \forall xB) \leftrightarrow \forall x.A \to B$.
4. $\vdash_c (A \to \exists xB) \leftrightarrow \exists x.A \to B$.
5. $\vdash_c (\forall xB \to A) \leftrightarrow \exists x.B \to A$.
6. $\vdash_c (\exists xB \to A) \leftrightarrow \forall x.B \to A$.

Beweis. Wir behandeln hier und im folgenden die abgeleiteten Regeln für $\exists$ und $\vee$ wie gewöhnliche Regeln.

(1) ist sehr leicht zu zeigen. (2) $\to$.

$$\dfrac{v\colon \exists xB \qquad \dfrac{\dfrac{u\colon A \qquad u_0\colon B}{A \wedge B} \qquad x}{\exists x.A \wedge B}}{\exists x.A \wedge B}\,\exists^- u_0$$

$\leftarrow$.

$$\dfrac{\dfrac{u\colon \exists x.A \wedge B \qquad \dfrac{u_0\colon A \wedge B}{A}}{A}\,\exists^- u_0 \qquad \dfrac{u\colon \exists x.A \wedge B \qquad \dfrac{\dfrac{u_1\colon A \wedge B}{B} \qquad x}{\exists xB}}{\exists xB}\,\exists^- u_1}{A \wedge \exists xB}$$

(3) Dies hatten wir in (1.3) und (1.4) gezeigt.

(4) $\to$. Hier helfen die abgeleiteten Regeln für $\exists$ nicht; wir müssen auf die Definition von $\exists$ zurückgehen. Ferner ist es nützlich, sich an die (klassische) Äquivalenz von $\neg(A \to B)$ mit $A \wedge \neg B$ zu erinnern. Wir führen den Beweis informal, und zwar so daß klar ist, wie man daraus (unter Zuhilfenahme des Äquivalenzlemmas 1.2.5) eine formale Herleitung gewinnen kann. Gelte also (*) $A \to \neg\forall x \neg B$ und (**) $\forall x.A \wedge \neg B$. Aus (**) erhält man A, also aus (*) $\neg\forall x \neg B$. Aus (**) erhält man auch $\forall x \neg B$, also insgesamt $\perp$, wie gewünscht.

(4) $\leftarrow$.

$$\begin{prooftree}
\AxiomC{$u\colon \exists x.A\to B$}
\AxiomC{$u_0\colon A\to B$}
\AxiomC{$u_1\colon A$}
\BinaryInfC{B}
\AxiomC{x}
\BinaryInfC{$\exists xB$}
\RightLabel{$\exists^- u_0$}
\BinaryInfC{$\exists xB$}
\RightLabel{$\to^+ u_1$}
\UnaryInfC{$A\to\exists xB$}
\end{prooftree}$$

(5) $\to$. Hier helfen wieder die abgeleiteten Regeln für $\exists$ nicht; wir müssen auf die Definition von $\exists$ zurückgehen. Ferner ist es wieder nützlich, die (klassische) Äquivalenz von $\neg(B\to A)$ mit $B\wedge\neg A$ zu verwenden. Wir führen den Beweis wieder informal. Gelte also (*) $\forall xB\to A$ und (**) $\forall x.B\wedge\neg A$. Aus (**) erhält man $\forall xB$, also aus (*) A. Aus (**) erhält man auch $\neg A$, also insgesamt $\bot$, wie gewünscht.

(5) $\leftarrow$.

$$\begin{prooftree}
\AxiomC{$u\colon \exists x.B\to A$}
\AxiomC{$u_0\colon B\to A$}
\AxiomC{$u_1\colon \forall xB$}
\AxiomC{x}
\BinaryInfC{B}
\BinaryInfC{A}
\RightLabel{$\exists^- u_0$}
\BinaryInfC{A}
\RightLabel{$\to^+ u_1$}
\UnaryInfC{$\forall xB\to A$}
\end{prooftree}$$

(6) $\to$.

$$\begin{prooftree}
\AxiomC{$u\colon \exists xB\to A$}
\AxiomC{$u_0\colon B$}
\AxiomC{x}
\BinaryInfC{$\exists xB$}
\BinaryInfC{A}
\RightLabel{$\to^+ u_0$}
\UnaryInfC{$B\to A$}
\RightLabel{$\forall^+$}
\UnaryInfC{$\forall x.B\to A$}
\end{prooftree}$$

(6) $\leftarrow$.

$$\begin{prooftree}
\AxiomC{$v\colon \exists xB$}
\AxiomC{$u\colon \forall x.B\to A$}
\AxiomC{x}
\BinaryInfC{$B\to A$}
\AxiomC{$u_1\colon B$}
\BinaryInfC{A}
\RightLabel{$\exists^- u_1$}
\BinaryInfC{A}
\end{prooftree}$$

$\square$

Eine Formel A heißt *pränex*, wenn $A=Q_1x_1Q_2x_2\dots Q_nx_nB$ mit $Q_i\in\{\forall,\exists\}$ und B quantorenfrei. B heißt der *Kern* von A und $Q_1x_1Q_2x_2\dots Q_nx_n$ das *Präfix* von A. Wir zeigen jetzt, daß sich jede Formel in eine *pränexe Normalform* bringen läßt.

Satz 1.2.14. *(Pränexe Normalform). Zu jeder Formel A findet man eine klassisch äquivalente pränexe Formel A'.*

Beweis. Induktion über A.

Fall Atomare Formel. In diesem Fall ist nichts zu zeigen.

Fall $A\wedge B$. Nach IH haben wir $\vdash_c A\leftrightarrow A'$ und $\vdash_c B\leftrightarrow B'$ mit

$$\begin{aligned}
A' &= Q_1x_1\dots Q_nx_nA_0, && A_0 \text{ quantorenfrei},\\
B' &= Q_1^*y_1\dots Q_m^*y_mB_0, && B_0 \text{ quantorenfrei}.
\end{aligned}$$

Wir können oBdA annehmen, daß kein x_i in B' und kein y_j in A' vorkommt. Nach dem Lemma 1.2.13 über das Herausziehen von Quantoren und dem Äquivalenzlemma 1.2.5 folgt

$$\vdash_c A\wedge B\leftrightarrow Q_1x_1\dots Q_nx_nQ_1^*y_1\dots Q_m^*y_m.A_0\wedge B_0.$$

Fall $A\to B$. Wie eben erhält man

$$\vdash_c (A\to B)\leftrightarrow \overline{Q}_1x_1\dots\overline{Q}_nx_nQ_1^*y_1\dots Q_m^*y_m.A_0\wedge B_0.$$

wobei $\overline{Q}_i=\forall$ falls $Q_i=\exists$, und $\overline{Q}_i=\exists$ falls $Q_i=\forall$.

Fall $\forall xA$. Nach IH haben wir $\vdash_c A\leftrightarrow A'$ mit einem pränexen A'. Nach dem Äquivalenzlemma 1.2.5 folgt

$$\vdash_c \forall xA\leftrightarrow\forall xA'.$$

$\square$

Zur Berechnung der pränexen Normalform ist es nützlich, die folgenden leicht zu beweisenden Äquivalenzen zu verwenden.

Lemma 1.2.15. *1.* $\vdash_c \neg\forall x B \leftrightarrow \exists x \neg B$.
2. $\vdash_c \neg\exists x B \leftrightarrow \forall x \neg B$.
3. $\vdash_c A \vee \forall x B \leftrightarrow \forall x. A \vee B$, *falls* $x \notin \mathsf{FV}(A)$.
4. $\vdash_c A \vee \exists x B \leftrightarrow \exists x. A \vee B$, *falls* $x \notin \mathsf{FV}(A)$.

Beweis. Übung. □

Starke Disjunktion und Existenz

Wenn man für Zwecke der Programmextraktion einen Existenzbeweis führen will, so ist es vorteilhaft, neben dem bisher behandelten *schwachen* oder *klassischen* Existenzquantor $\exists$ (der durch $\neg\forall\neg$ definiert war) auch noch einen *starken* oder *konstruktiven* Existenzquantor $\exists^*$ zuzulassen. Man kann dann in den Fällen, wo ein Existenzbeweis tatsächlich konstruktiv durch Angabe eines Beispiels geführt wurde, dies auch in der Formelsprache angemessen ausdrücken.

Entsprechend könnte man auch neben der bisher behandelten *schwachen* oder *klassischen* Disjunktion $\vee$ (die durch $\neg \wedge \neg$ definiert war) auch noch eine *starken* oder *konstruktive* Disjunktion $\vee^*$ zulassen. In Anwesenheit des Grundtyps ι der natürlichen Zahlen ist dies jedoch entbehrlich: Wir definieren

$$A \vee^* B \quad := \quad \exists^* n.(n = 0 \to A) \wedge (n \neq 0 \to B).$$

Wir wollen kurz diskutieren, welchen Effekt die Hinzunahme von $\vee^*, \exists^*$ auf unsere bisherigen Untersuchungen hat. Wir erweitern also den Formelbegriff um die Klauseln $A \vee^* B$ und $\exists^* x A$. Die Definitionen der Mengen der freien Variablen und der Substitution werden in der offensichtlichen Weise erweitert, ebenso unsere Konventionen betreffend Klammersetzung. Die Logik von $\vee^*$ und $\exists^*$ kann dann durch die abgeleiteten Regeln für den Existenzquantor und die Disjunktion axiomatisiert werden. Alternativ kann man auch die Regeln der Minimallogik beibehalten und den starken Existenzquantor durch die Generalisierung der folgenden Formeln axiomatisieren; diesen Weg wollen wir hier gehen.

$$\begin{aligned}
&A \to A \vee^* B, \quad B \to A \vee^* B. \\
&A \vee^* B \to (A \to C) \to (B \to C) \to C. \\
&A \to \exists^* x A. \\
&\exists^* x A \to (\forall x. A \to B) \to B, \quad \text{falls } x \notin \mathsf{FV}(B).
\end{aligned}$$

Diese Axiomenschemata bezeichnen wir durch $\vee_0^{*+}$, $\vee_1^{*+}$, $\vee^{*-}$, $\exists^{*+}$ und $\exists^{*-}$. Für Formeln A in der durch $\vee^*$ und $\exists^*$ erweiterten Sprache schreiben wir $\vdash A$ und nennen A *herleitbar* (in der *Minimallogik*), wenn es eine Herleitung von A aus diesen Axiomenschemata gibt.

Lemma 1.2.16. *(Ex-falso-quodlibet). Für jede Formel A in der Sprache mit den Verknüpfungen $\to$, $\wedge$, $\vee^*$, $\forall$ und $\exists^*$ gilt*

$$\vdash_i \bot \to A.$$

Beweis. Wir müssen nur die Fälle $A \vee^* B$ und $\exists^* x A$ zusätzlich behandeln, was aber trivial ist. □

Die Einbettung der intuitionistischen Logik läßt sich also wie bisher durchführen.

1.3 Modelle

Es ist eine offensichtliche Frage, ob unsere logischen Regeln ausreichen, d.h. ob wir notwendige Regeln vergessen haben. Um diese Frage beantworten zu können, müssen wir die *Bedeutung* einer Formel kennen, d.h. wir müssen eine *Semantik* angegeben haben. Zunächst wollen wir deshalb den Begriff der Bedeutung eines Terms und einer Formel präzisieren. Dazu definieren wir den Begriff einer Struktur (genauer $\mathcal{L}$-Struktur) und erklären dann, was der Wert eines Terms und die Bedeutung einer Formel in einer solchen

Struktur sein soll. Wir zeigen dann den Korrektheitssatz: er sagt aus, daß jede in der klassischen Logik herleitbare Formel in einer beliebigen Struktur gültig ist.

Ferner definieren wir Strukturbegriffe von BETH und KRIPKE, die beide zur Minimallogik und zur intuitionistischen Logik passen, und zeigen Korrektkeitssätze für beide Logiken. Im nächsten Anschnitt beweisen wir dann die Vollständigkeit unserer Regeln bzgl. dieser Strukturbegriffe, und als Folgerung erhalten wir im übernächsten Abschnitt die Vollständigkeit der klassischen Logik, bezogen auf den üblichen Strukturbegriff.

Strukturen

Definition 1.3.1. $\mathcal{M} = (D, I)$ heißt *Prästruktur* (genauer $\mathcal{L}$-Prästruktur), wenn D eine nichtleere Menge ist (die *Trägermenge* oder der *Individuenbereich* von $\mathcal{M}$) und I eine Abbildung (*Interpretation*), die jedem n-stelligen Funktionssymbol f von $\mathcal{L}$ eine Funktion

$$I(f)\colon D^n \to D$$

zuordnet. Ist $n = 0$, so ist $I(f)$ ein Element von D. $\mathcal{M} = (D, I_0, I_1)$ heißt *Struktur* (genauer $\mathcal{L}$-Struktur), wenn (D, I_0) eine Prästruktur ist und I_1 eine Abbildung, die jedem n-stelligen Relationssymbol R von $\mathcal{L}$ eine n-stellige Relation

$$I_1(R) \subseteq D^n$$

zuordnet. Ist $n = 0$, so ist $I_1(R)$ einer der Wahrheitswerte 1 und 0; insbesondere soll $I_1(\bot) = 0$ sein.

Ist $\mathcal{M} = (D, I)$ bzw. (D, I_0, I_1), so schreiben wir oft $|\mathcal{M}|$ für die Trägermenge D von $\mathcal{M}$ sowie $f^{\mathcal{M}}$ und $R^{\mathcal{M}}$ für die Interpretationen $I_0(f)$ und $I_1(R)$ der Funktions- und Relationssymbole.

Eine (Variablen-) *Belegung* oder *Umgebung* in D ist eine Abbildung, die jeder Variablen $x \in \mathsf{dom}(\eta)$ einen Wert $\eta(x) \in D$ zuordnet. Endliche Belegungen schreiben wir als $[x_1 := a_1, \dots, x_n := a_n]$ (oder auch $[a_1/x_1, \dots, a_n/x_n]$) mit verschiedenen $x_1, \dots, x_n$. Ist η eine Belegung in D und $a \in D$, so sei η_x^a die Belegung in D, die x auf a abbildet und sonst mit η übereinstimmt, also

$$\eta_x^a(y) := \begin{cases} \eta(y), & \text{falls } y \neq x \\ a, & \text{falls } y = x. \end{cases}$$

Seien eine Prästruktur $\mathcal{M}$ und eine Belegung η in $|\mathcal{M}|$ gegeben. Wir definieren eine homomorphe Fortsetzung von η (ebenfalls bezeichnet durch η) auf die Menge $\mathsf{Ter}_{\mathcal{L}}$ der $\mathcal{L}$-Terme t mit $\mathsf{vars}(t) \subseteq \mathsf{dom}(\eta)$ durch

$$\begin{aligned} &\eta(c) := c^{\mathcal{M}}, \\ &\eta(f(t_1, \dots, t_n)) := f^{\mathcal{M}}(\eta(t_1), \dots, \eta(t_n)). \end{aligned}$$

Man beachte, daß die Erweiterung von η von $\mathcal{M}$ abhängt; wir schreiben deshalb auch $t^{\mathcal{M}}[\eta]$ für $\eta(t)$.

Für jede Struktur $\mathcal{M}$, Belegung η in $|\mathcal{M}|$ und Formel A mit $\mathsf{FV}(A) \subseteq \mathsf{dom}(\eta)$ definieren wir $\mathcal{M} \models A[\eta]$ (gelesen: A ist *gültig* in $\mathcal{M}$ unter der Belegung η) durch Rekursion über A. Diese Definition wurde zuerst von TARSKI angegeben.

$$\begin{aligned}
&\mathcal{M} \models R(t_1, \dots, t_n)[\eta] \quad :\Longleftrightarrow \quad (t_1^{\mathcal{M}}[\eta], \dots, t_n^{\mathcal{M}}[\eta]) \in I_1(R) \quad \text{für } R \text{ nicht nullstellig.} \\
&\mathcal{M} \models R[\eta] \quad :\Longleftrightarrow \quad I_1(R) = 1 \quad \text{für } R \text{ nullstellig.} \\
&\mathcal{M} \models (A \wedge B)[\eta] \quad :\Longleftrightarrow \quad \mathcal{M} \models A[\eta] \text{ und } \mathcal{M} \models B[\eta]. \\
&\mathcal{M} \models (A \to B)[\eta] \quad :\Longleftrightarrow \quad \text{wenn } \mathcal{M} \models A[\eta], \text{ so } \mathcal{M} \models B[\eta]. \\
&\mathcal{M} \models (\forall x A)[\eta] \quad :\Longleftrightarrow \quad \text{für alle } a \in |\mathcal{M}| \text{ gilt } \mathcal{M} \models A[\eta_x^a].
\end{aligned}$$

Wegen $I_1(\bot) = 0$ folgt insbesondere $\mathcal{M} \not\models \bot[\eta]$.

Ist Γ eine Menge von Formeln, so schreiben wir $\mathcal{M} \models \Gamma[\eta]$, wenn für alle $A \in \Gamma$ gilt $\mathcal{M} \models A[\eta]$. Gilt $\mathcal{M} \models A[\eta]$ für alle Belegungen η in $|\mathcal{M}|$, so schreiben wir $\mathcal{M} \models A$.

Lemma 1.3.2. *(Koinzidenzlemma). Sei $\mathcal{M}$ eine Struktur, t ein Term, A eine Formel und η, ξ Belegungen in $|\mathcal{M}|$.*

1. *Gilt $\eta(x) = \xi(x)$ für alle $x \in \mathsf{vars}(t)$, so ist $\eta(t) = \xi(t)$.*
2. *Gilt $\eta(x) = \xi(x)$ für alle $x \in \mathsf{FV}(A)$, so gilt $\mathcal{M} \models A[\eta]$ genau dann, wenn $\mathcal{M} \models A[\xi]$.*

Beweis. Induktion über Terme bzw. Formeln. □

Lemma 1.3.3. *(Substitutionslemma). Sei $\mathcal{M}$ eine $\mathcal{L}$-Struktur, t, r $\mathcal{L}$-Terme, A eine $\mathcal{L}$-Formel und η eine Belegung in $|\mathcal{M}|$. Dann gilt*

1. *$\eta(r[x := t]) = \eta_x^{\eta(t)}(r)$.*
2. *$\mathcal{M} \models A[x := t][\eta] \iff \mathcal{M} \models A[\eta_x^{\eta(t)}]$.*

Beweis. 1. Induktion über r. 2. Induktion über A. Wir beschränken uns auf die Fälle einer atomaren Formel und einer Allformel; die restlichen Fälle sind sehr einfach.

Fall $R(s_1, \dots, s_n)$. Zur Vereinfachung nehmen wir $n = 1$ an. Dann gilt

$$\begin{aligned} \mathcal{M} \models R(s)[x := t][\eta] &\iff \mathcal{M} \models R(s[x := t])[\eta] \\ &\iff \eta(s[x := t]) \in R^{\mathcal{M}} \\ &\iff \eta_x^{\eta(t)}(s) \in R^{\mathcal{M}} \quad \text{nach (1)} \\ &\iff \mathcal{M} \models R(s)[\eta_x^{\eta(t)}]. \end{aligned}$$

Fall $\forall y A$. Wir können oBdA $y \neq x$ und $y \notin \mathsf{vars}(t)$ annehmen.

$$\begin{aligned} &\mathcal{M} \models (\forall y A)[x := t][\eta] \\ &\iff \mathcal{M} \models (\forall y A[x := t])[\eta] \\ &\iff \text{für alle } a \in |\mathcal{M}| \text{ gilt } \mathcal{M} \models A[x := t][\eta_y^a] \\ &\iff \text{für alle } a \in |\mathcal{M}| \text{ gilt } \mathcal{M} \models A[(\eta_y^a)_x^b] \text{ mit } b = \eta_y^a(t) = \eta(t) \text{ (IH und Lemma 1.3.2)} \\ &\iff \text{für alle } a \in |\mathcal{M}| \text{ gilt } \mathcal{M} \models A[(\eta_x^b)_y^a] \text{ da } x \neq y \\ &\iff \mathcal{M} \models (\forall y A)[\eta_x^b] \end{aligned}$$

□

Für beliebige (oBdA geschlossene) $\mathcal{L}$-Formeln A werden wir zeigen, daß A genau dann in der klassischen Logik herleitbar ist, wenn A in allen $\mathcal{L}$-Strukturen gültig ist. Wir zeigen zunächst die einfache Richtung.

Satz 1.3.4. *(Korrektheit). Sei $\Gamma \vdash_c B$. Ist dann $\mathcal{M}$ eine Struktur und η eine Belegung in $|\mathcal{M}|$, so folgt aus $\mathcal{M} \models \Gamma[\eta]$ stets $\mathcal{M} \models B[\eta]$.*

Beweis. Induktion über Herleitungen. Da die gegebene Herleitung von B aus Γ nur endlich viele Annahmen frei enthält, können wir oBdA $\Gamma = \{A_1, \dots, A_n\}$ voraussetzen.

Fall $u\colon B$. Dann ist $B \in \Gamma$ und die Behauptung ist klar.

Fall $\mathsf{Stab}_R\colon \forall \boldsymbol{x}. \neg\neg R\boldsymbol{x} \to R\boldsymbol{x}$. Die Behauptung ist wieder klar, da $\mathcal{M} \models \neg\neg A[\eta]$ gleichbedeutend ist mit $\mathcal{M} \models A[\eta]$.

Fall $\wedge^+$. Gelte $\mathcal{M} \models \Gamma[\eta]$. Zu zeigen ist $\mathcal{M} \models (A \wedge B)[\eta]$. Nach IH haben wir $\mathcal{M} \models A[\eta]$ und $\mathcal{M} \models B[\eta]$. Die Behauptung folgt aus der Definition von $\models$. Die Fälle $\wedge_0^-$, $\wedge_1^-$ und $\to^-$ beweist man genauso.

Fall $\to^+$. Gelte $\mathcal{M} \models \Gamma[\eta]$. Zu zeigen ist $\mathcal{M} \models (A \to B)[\eta]$. Gelte also noch $\mathcal{M} \models A[\eta]$. Zu zeigen ist $\mathcal{M} \models B[\eta]$. Nach IH (mit $\Gamma \cup \{A\}$ statt Γ) ist dies richtig.

Fall $\forall^+$. Gelte $\mathcal{M} \models \Gamma[\eta]$. OBdA mögen alle Formeln $A_1, \dots, A_n$ in der gegebenen Herleitung von A als freie Annahmen vorkommen. Zu zeigen ist $\mathcal{M} \models (\forall x A)[\eta]$. Sei also $a \in |\mathcal{M}|$. Zu zeigen ist $\mathcal{M} \models A[\eta_x^a]$. Da aufgrund der Variablenbedingung für $\forall^+$ die Variable x in keiner der Formeln $A_1, \dots, A_n$ frei vorkommt, gilt nach dem Koinzidenzlemma $\mathcal{M} \models \Gamma[\eta_x^a]$. Die IH (mit η_x^a statt η) liefert $\mathcal{M} \models A[\eta_x^a]$.

Fall $\forall^-$. Gelte $\mathcal{M} \models \Gamma[\eta]$. Zu zeigen ist $\mathcal{M} \models A[x := t][\eta]$, d.h. nach dem Substitutionslemma $\mathcal{M} \models A[\eta_x^b]$ mit $b = \eta(t)$. Nach IH haben wir $\mathcal{M} \models (\forall x A)[\eta]$, d.h. für alle $a \in |\mathcal{M}|$ gilt $\mathcal{M} \models A[\eta_x^a]$. Mit $\eta(t)$ für a folgt die Behauptung. □

Beth-Strukturen

Ein zur Minimallogik und zur intuitionistischen Logik passende Strukturbegriff wurde zuerst von Beth [1] konzipiert; er basiert auf einer Vorstellung von “sich entwickelnden möglichen Welten”, die durch Knoten k eines endlich verzweigten Baums indiziert sind. Kenntnisse können sich nur vergrößern, d.h. gilt $\boldsymbol{Rt}$ in einer Welt k, so gilt $\boldsymbol{Rt}$ auch in allen künftigen möglichen Welten.

Jede Beth-Struktur basiert also auf einem endlich verzweigten Baum T. Wir führen zunächst die hierbei notwendigen Begriffe ein. Ein *Knoten* über einer nichtleeren Menge S ist eine endliche Folge $k = \langle a_0, a_1, \dots, a_{n-1} \rangle$ von Elementen $a_i \in S$; n ist die *Länge* $\mathsf{lh}(k)$ von k. Wir schreiben $k \preceq k'$ wenn k ein Anfangsstück von k' ist. Ein *Baum* über S ist eine Menge von Knoten (über S), die gegen Bildung von Anfangsstücken abgeschlossen ist. Ein Baum T ist *endlich verzweigt*, wenn jedes $k \in T$ höchstens endlich viele unmittelbare Nachfolger in T hat. Ein Baum T ist *unbeschränkt*, wenn es für jedes $n \in \mathbb{N}$ einen Knoten $k \in T$ gibt so daß $\mathsf{lh}(k) = n$. Ein *Ast* in einem Baum T ist ein linear geordneter (durch $\preceq$) Teilbaum von T. Ein *Blatt* in T ist ein Knoten k in T ohne echte Fortsetzungen in T.

Für den Vollständigkeitssatz wird es genügen, Beth-Strukturen über dem vollen binären Baum zu betrachten, d.h. der Menge T_{01} aller endlichen 0-1-Folgen (Knoten) k. Die leere Folge wird mit $\langle\rangle$ bezeichnet, und $k0$, $k1$ bezeichnen Erweiterungen der Folge k durch 0 oder 1.

Definition 1.3.5. Sei $(T, \preceq)$ ein endlich verzweigter Baum. $\mathcal{B} = (M, I_0, I_1)$ ist eine $\mathcal{L}$-Beth-*Struktur* über T, wenn (M, I_0) eine $\mathcal{L}$-Prästruktur und I_1 jedem n-stelligen Relationssymbol R von $\mathcal{L}$ und jedem Knoten $k \in T$ eine n-stellige Relation

$$I_1(R, k) \subseteq M^n$$

zuordnet so daß Monotonie gilt, d.h.

$$k \preceq k' \implies I_1(R, k) \subseteq I_1(R, k').$$

Ist $n = 0$, so ist $I_1(R, k)$ wahr oder falsch und die Monotonie sagt aus, daß für $k \preceq k'$ aus $I_1(R, k)$ stets $I_1(R, k')$ folgt.

Von $I_1(\bot, k)$ wird also *nichts* verlangt; das Falsum spielt in der Minimallogik die Rolle eines gewöhnlichen Aussagensymbols.

$t^{\mathcal{B}}[\eta]$ für eine Belegung η wird wie bei klassischen Modellen erklärt. An die Stelle der Modellbeziehung $\mathcal{M} \models A[\eta]$ tritt bei Beth-Strukturen jedoch die *Erzwingungsbeziehung*. Für deren Definition ist es bequem, den zugrunde liegenden Baum T zunächst zu *vervollständigen* zu einem Baum $\bar{T}$ ohne Blätter, indem wir zu jedem Blatt $k \in T$ alle Fortsetzungen $k0$, $k00$, $k000, \dots$ zu T hinzunehmen. Für jeden hinzugekommenen Knoten $k0\dots0$ setzen wir $I_1(R, k0\dots0) := I_1(R, k)$.

Definition 1.3.6. $\mathcal{B}, k \Vdash A[\eta]$ ($\mathcal{B}$ erzwingt A im Knoten k für die Belegung η) wird induktiv wie folgt definiert. Wir schreiben $k \Vdash A[\eta]$ wenn die unterliegende Struktur $\mathcal{B}$ klar ist, und $\forall k' {\succeq_n} k\, A$ für $\forall k' {\succeq} k. \mathsf{lh}(k') = \mathsf{lh}(k) + n \to A$.

$$k \Vdash R(t_1, \dots, t_p)[\eta] \;:\Longleftrightarrow\; \exists n \forall k' {\succeq_n} k\, (t_1^{\mathcal{B}}[\eta], \dots, t_p^{\mathcal{B}}[\eta]) \in I_1(R, k') \quad \text{für } R \text{ nicht nullstellig.}$$
$$k \Vdash R[\eta] \;:\Longleftrightarrow\; \exists n \forall k' {\succeq_n} k\, I_1(R, k') = 1 \quad \text{für } R \text{ nullstellig.}$$
$$k \Vdash (A \vee^* B)[\eta] \;:\Longleftrightarrow\; \exists n \forall k' {\succeq_n} k. k' \Vdash A[\eta] \;\text{ oder }\; k' \Vdash B[\eta].$$
$$k \Vdash (\exists^* x A)[\eta] \;:\Longleftrightarrow\; \exists n \forall k' {\succeq_n} k \exists a {\in} |\mathcal{B}|\, k' \Vdash A[\eta_x^a].$$
$$k \Vdash (A \to B)[\eta] \;:\Longleftrightarrow\; \forall k' {\succeq} k. k' \Vdash A[\eta] \implies k' \Vdash B[\eta].$$
$$k \Vdash (A \wedge B)[\eta] \;:\Longleftrightarrow\; k \Vdash A[\eta] \;\text{ und }\; k \Vdash B[\eta].$$
$$k \Vdash (\forall x A)[\eta] \;:\Longleftrightarrow\; \forall a {\in} |\mathcal{B}|\, k \Vdash A[\eta_x^a].$$

In den Klauseln für Atome, Disjunktionen und Existenzformeln beziehen wir uns also auf eine “Schranke” (“bar”) in $\bar{T}$. Für Atome wäre dies jedoch nicht nötig; es ist jedoch bequem für die Konstruktion von Beth-Strukturen. Man kann eine gegebene Beth-Struktur leicht “vervollständigen” durch “Herunterziehen der Schranke” für Atome. Dies bedeutet, daß wir die ursprünglich gegebene Interpretation I_1 der Relationssymbol ersetzen durch

$$\overline{I}_1(R,k) := \{\, \boldsymbol{a} \in |\mathcal{B}|^p \mid \exists n \forall k' {\succeq_n} k\, \boldsymbol{a} \in I_1(R,k')\,\}.$$

Dann gilt offensichtlich wieder die Monotonie, d.h. $\overline{I}_1(R,k) \subseteq \overline{I}_1(R,k')$ für $k \preceq k'$. In einer solchen vervollständigten BETH-Struktur können wir die Klausel für Atome in der Definition der Erzwingungsbeziehung vereinfachen zu

$$k \Vdash R\boldsymbol{t}[\eta] \iff \boldsymbol{t}^{\mathcal{B}}[\eta] \in \overline{I}_1(R,k).$$

Aus der Definition ergibt sich leicht, daß die Monotonie sich auf Formeln überträgt, d.h. daß aus $k \Vdash A[\eta]$ stets $k' \Vdash A[\eta]$ folgt für $k \preceq k'$. Auch die Umkehrung ist richtig:

Lemma 1.3.7. *(Überdeckungseigenschaft).*

$$\forall k' {\succeq_n} k\, k' \Vdash A[\eta] \implies k \Vdash A[\eta].$$

Beweis. Induktion über A. Wir schreiben $k \Vdash A$ für $k \Vdash A[\eta]$.

Fall $R\boldsymbol{t}$. Gelte

$$\exists n \forall k' {\succeq_n} k\, k' \Vdash R\boldsymbol{t},$$

also nach Definition

$$\exists n \forall k' {\succeq_n} k \exists m \forall k'' {\succeq_m} k' \boldsymbol{t}^{\mathcal{B}}[\eta] \in I_1(R,k'').$$

Da T ein endlich verzweigter Baum ist, haben wir

$$\exists m \forall k' {\succeq_m} k\, \boldsymbol{t}^{\mathcal{B}}[\eta] \in I_1(R,k'),$$

also $k \Vdash R\boldsymbol{t}$.

Die Fälle $A \vee^* B$ und $\exists^* x A$ behandelt man ähnlich.

Fall $A \to B$. Gelte $k' \Vdash A \to B$ für alle $k' \succeq k$ mit $\mathsf{lh}(k') = \mathsf{lh}(k) + n$. Wir müssen zeigen

$$\forall l {\succeq} k. l \Vdash A \implies l \Vdash B.$$

Gelte also $l \succeq k$ und $l \Vdash A$. Zu zeigen ist $l \Vdash B$. Wir verwenden die IH für B mit $m := \max(\mathsf{lh}(k){+}n, \mathsf{lh}(l))$. Gelte also $l' \succeq l$ und $\mathsf{lh}(l') = m$. Es genügt zu zeigen $l' \Vdash B$. Ist $\mathsf{lh}(l') = \mathsf{lh}(l)$, so gilt $l' = l$ und wir sind fertig. Ist $\mathsf{lh}(l') = \mathsf{lh}(k) + n > \mathsf{lh}(l)$, so ist l' eine Erweiterung von l und auch von k mit Länge $\mathsf{lh}(k) + n$, und wir haben $l' \Vdash A \to B$ nach Annahme. Ferner gilt $l' \Vdash A$, da $l' \succeq l$ und $l \Vdash A$. Dies wiederum impliziert $l' \Vdash B$.

Die Fälle $A \wedge B$ und $\forall x\, A$ sind klar. □

Das Koinzidenzlemma und das Substitutionslemma lassen sich wie erwartet auf BETH-Strukturen übertragen.

Lemma 1.3.8. *(Koinzidenzlemma). Sei $\mathcal{B}$ eine* BETH-*Struktur, t ein Term, A eine Formel und η, ξ Belegungen in $|\mathcal{B}|$.*

1. *Gilt $\eta(x) = \xi(x)$ für alle $x \in \mathsf{vars}(t)$, so ist $\eta(t) = \xi(t)$.*
2. *Gilt $\eta(x) = \xi(x)$ für alle $x \in \mathsf{FV}(A)$, so folgt $\mathcal{B}, k \Vdash A[\eta] \iff \mathcal{B}, k \Vdash A[\xi]$.*

Beweis. Induktion über Terme und Formeln. □

Lemma 1.3.9. *(Substitutionslemma). Sei $\mathcal{B}$ eine* BETH-*Struktur, t, r Terme, A eine Formel und η eine Belegung in $|\mathcal{B}|$. Dann gilt*

1. *$\eta(r[x := t]) = \eta_x^{\eta(t)}(r)$.*
2. *$\mathcal{B}, k \Vdash A[x := t][\eta] \iff \mathcal{B}, k \Vdash A[\eta_x^{\eta(t)}]$.*

Beweis. Induktion über Terme und Formeln. □

Hieraus erhalten wir wie üblich den Korrektheitssatz.

Satz 1.3.10. *(Korrektheit). Sei $\Gamma \cup \{A\}$ eine Formelmenge und es gelte $\Gamma \vdash A$. Ist dann $\mathcal{B}$ eine* BETH-*Struktur, k ein Knoten und η eine Belegung in $|\mathcal{B}|$, so folgt aus $\mathcal{B}, k \Vdash \Gamma[\eta]$ stets $\mathcal{B}, k \Vdash A[\eta]$.*

Beweis. Induktion über Herleitungen. Wir beginnen mit den Axiomenschemata $\vee_0^{*+}$, $\vee_1^{*+}$, $\vee^{*-}$, $\exists^{*+}$ und $\exists^{*-}$. Statt $k \Vdash C[\eta]$ schreiben wir kurz $k \Vdash C$, falls η aus dem Kontext bekannt ist.

Fall $\vee_0^{*+}$: $A \to A \vee^* B$. Zu zeigen ist $k \Vdash A \to A \vee^* B$. Sei also $k' \succeq k$ mit $k' \Vdash A$ gegeben. Zu zeigen ist $k' \Vdash A \vee^* B$; dies folgt aber nach Definition aus $k' \Vdash A$. Den Fall $\vee_1^{*+}$: $B \to A \vee^* B$ behandelt man ähnlich.

Fall $\vee^{*-}$: $A \vee^* B \to (A \to C) \to (B \to C) \to C$. Zu zeigen ist $k \Vdash A \vee^* B \to (A \to C) \to (B \to C) \to C$. Sei also $k' \succeq k$ und gelte $k' \Vdash A \vee^* B$, $k' \Vdash A \to C$ und $k' \Vdash B \to C$ (offenbar kann man annehmen, daß dasselbe k' für alle drei Prämissen gewählt ist). Zu zeigen ist $k' \Vdash C$. Nach Definition gibt es ein n so daß für alle $k'' \succeq_n k'$ gilt $k'' \Vdash A$ oder $k'' \Vdash B$. In beiden Fällen folgt $k'' \Vdash C$, da $k' \Vdash A \to C$ und $k' \Vdash B \to C$. Aus der Überdeckungseigenschaft 1.3.7 ergibt sich die Behauptung $k' \Vdash C$.

Fall $\exists^{*+}$: $A \to \exists^* xA$. Zu zeigen ist $k \Vdash (A \to \exists^* xA)[\eta]$. Sei also $k' \succeq k$ mit $k' \Vdash A[\eta]$ gegeben. Zu zeigen ist $k' \Vdash (\exists^* xA)[\eta]$. Wegen $\eta = \eta_x^{\eta(x)}$ gibt es ein $a \in |\mathcal{B}|$ (nämlich $a := \eta(x)$) mit $k' \Vdash A[\eta_x^a]$. Also haben wir $k' \Vdash (\exists^* xA)[\eta]$.

Fall $\exists^{*-}$: $\exists^* xA \to (\forall x.A \to B) \to B$ mit $x \notin \mathsf{FV}(B)$. Zu zeigen ist $k \Vdash (\exists^* xA \to (\forall x.A \to B) \to B)[\eta]$. Sei also $k' \succeq k$ mit $k' \Vdash (\exists^* xA)[\eta]$ und $k' \Vdash (\forall x.A \to B)[\eta]$ gegeben. Zu zeigen ist $k' \Vdash B[\eta]$. Nach Definition gibt es ein n so daß für alle $k'' \succeq_n k'$ es ein $a \in |\mathcal{B}|$ gibt mit $k'' \Vdash A[\eta_x^a]$. Mit $k' \Vdash (\forall x.A \to B)[\eta]$ folgt $k'' \Vdash B[\eta_x^a]$, also wegen $x \notin \mathsf{FV}(B)$ nach dem Koinzidenzlemma auch $k'' \Vdash B[\eta]$. Aus der Überdeckungseigenschaft 1.3.7 ergibt sich die Behauptung $k' \Vdash B[\eta]$.

Fall $\to^+$. Gelte $k \Vdash \Gamma$. Zu zeigen ist $k \Vdash A \to B$. Sei also $k' \succeq k$ mit $k' \Vdash A[\eta]$ gegeben. Zu zeigen ist $k' \Vdash B$. Es gilt $k' \Vdash \Gamma \cup \{A\}$, also nach IH $k' \Vdash B$.

Fall $\to^-$. Gelte $k \Vdash \Gamma$. Nach IH haben wir $k \Vdash A \to B$ und $k \Vdash A$, also auch $k \Vdash B$.

Fall $\forall^+$. Gelte $k \Vdash \Gamma[\eta]$ und $x \notin \mathsf{FV}(\Gamma)$. Wir müssen zeigen $k \Vdash (\forall xA)[\eta]$, d.h. $k \Vdash A[\eta_x^a]$ für ein beliebiges $a \in |\mathcal{B}|$. Wir haben

$$\begin{array}{ll} k \Vdash \Gamma[\eta_x^a] & \text{nach dem Koinzidenzlemma, da } x \notin \mathsf{FV}(\Gamma) \\ k \Vdash A[\eta_x^a] & \text{nach IH.} \end{array}$$

Fall $\forall^-$. Gelte $k \Vdash \Gamma[\eta]$. Wir müssen zeigen $k \Vdash A[x := t][\eta]$. Wir haben

$$\begin{array}{ll} k \Vdash (\forall xA)[\eta] & \text{nach IH} \\ k \Vdash A[\eta_x^{\eta(t)}] & \text{nach Definition} \\ k \Vdash A[x := t][\eta] & \text{nach dem Substitutionslemma.} \end{array}$$

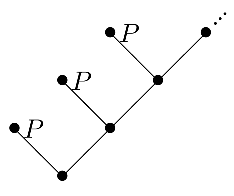
□

Gegenmodelle

Mit Hilfe des Korrektheitssatzes ist es einfach, Gegenmodelle zu finden, aus denen sich die Nicht-Herleitbarkeit in der Minimallogik bzw. in der intuitionistischen Logik ergibt. Unter einer Beth-*Struktur für die intuitionistische Logik* verstehen wir eine Beth-Struktur $\mathcal{B} = (M, I_0, I_1)$, in der $\bot$ nicht erzwungen wird, d.h. $I_1(\bot, k) = 0$ für alle k. In Beth-Strukturen für die intuitionistische Logik haben wir deshalb

$$\begin{aligned} k \Vdash \neg A &\iff \forall k' {\succeq} k\, k' \not\Vdash A, \\ k \Vdash \neg\neg A &\iff \forall k' {\succeq} k\, k' \not\Vdash \neg A \\ &\iff \forall k' {\succeq} k \exists k'' {\succeq} k'\, k'' \Vdash A. \end{aligned}$$

Als Beispiel zeigen wir $\not\vdash_i \neg\neg P \to P$. Um eine Beth-Struktur zu beschreiben, stellen wir den zugrunde liegenden Baum durch ein Diagramm dar, wo wir neben jeden Knoten die Aussagensymbole schreiben, die in diesem Knoten erzwungen werden. Man betrachte die durch das folgende Diagramm gegebene Beth-Struktur.



Offenbar haben wir

$$\langle\rangle \not\Vdash P,$$
$$\langle\rangle \Vdash \neg\neg P.$$

Also gilt $\langle\rangle \not\Vdash \neg\neg P \to P$ und deshalb $\not\vdash \neg\neg P \to P$. Da offenbar $\Vdash \mathsf{Efq}_R$ für jedes R gilt, haben wir auch $\not\vdash_i \neg\neg P \to P$. Das Modell zeigt ebenfalls die Nicht-Herleitbarkeit der PEIRCE-Formel $((P \to Q) \to P) \to P$ in der intuitionistischen Logik.

1.4 Vollständigkeit der Minimallogik und der intuitionistischen Logik

Wir zeigen jetzt die Umkehrung der Korrektheitssatzes.

Satz 1.4.1. *(Vollständigkeit). Sei $\Gamma \cup \{A\}$ eine Formelmenge. Dann sind folgende Aussagen äquivalent.*

1. *$\Gamma \vdash A$.*
2. *$\Gamma \Vdash A$, d.h. für alle BETH-Strukturen $\mathcal{B}$, Knoten k und Belegungen η*

$$\mathcal{B}, k \Vdash \Gamma[\eta] \implies \mathcal{B}, k \Vdash A[\eta].$$

Beweis. Eine Richtung ist der Korrektheitssatz. Für die andere Richtung verwenden wir einen Ansatz von Harvey FRIEDMAN [8] und konstruieren eine BETH-Struktur $\mathcal{B}$ (über der Menge T_{01} aller endlichen 0-1-Folgen k geordnet durch die Anfangsstück-Relation $k \preceq k'$) mit der Eigenschaft, daß $\Gamma \vdash B$ äquivalent ist zu $\mathcal{B}, \langle\rangle \Vdash B[\mathsf{id}]$.

Zur Definition von $\mathcal{B}$ gehen wir aus von einer festen Aufzählung $A_0, A_1, A_2, \dots$ aller $\mathcal{L}$-Formeln, in der jede Formel unendlich oft vorkommen möge; wir fixieren auch eine Aufzählung $x_0, x_1, \dots$ aller Variablen. Sei $\Gamma = \bigcup_n \Gamma_n$ mit endlichen Mengen Γ_n so daß $\Gamma_n \subseteq \Gamma_{n+1}$. Jedem Knoten $k \in T_{01}$ ordnen wir eine endliche Menge Δ_k von Formeln zu, durch Induktion über die Länge von k.

Sei $\Delta_{\langle\rangle} := \emptyset$. Nehmen wir jetzt an, daß für einen Knoten k mit $\mathsf{lh}(k) = n$ die Menge Δ_k schon definiert ist. $\Delta \vdash_n B$ bedeute, daß es eine Herleitung von B aus Δ gibt mit Länge (:= Gesamtzahl der Symbole) $\leq n$. Wir definieren Δ_{k0} und Δ_{k1} wie folgt.

Fall 1. $\Gamma_n, \Delta_k \not\vdash_n A_n$. Dann sei

$$\Delta_{k0} := \Delta_k \quad \text{und} \quad \Delta_{k1} := \Delta_k \cup \{A_n\}.$$

Fall 2. $\Gamma_n, \Delta_k \vdash_n A_n = A'_n \vee^* A''_n$. Dann sei

$$\Delta_{k0} := \Delta_k \cup \{A_n, A'_n\} \quad \text{und} \quad \Delta_{k1} := \Delta_k \cup \{A_n, A''_n\}.$$

Fall 3. $\Gamma_n, \Delta_k \vdash_n A_n = \exists^* x A'_n$. Dann sei

$$\Delta_{k0} := \Delta_{k1} := \Delta_k \cup \{A_n, A'_n[x := x_i]\},$$

mit x_i die erste Variable $\notin \mathsf{FV}(\Gamma_n, A_n, \Delta_k)$.

Fall 4. $\Gamma_n, \Delta_k \vdash_n A_n$, mit A_n weder Disjunktion noch Existenzformel. Dann sei

$$\Delta_{k0} := \Delta_{k1} := \Delta_k \cup \{A_n\}.$$

Offenbar impliziert $k \preceq k'$ stets $\Delta_k \subseteq \Delta_{k'}$. Man beachte zunächst, daß

$$\forall k' {\succeq_n} k\, \Gamma, \Delta_{k'} \vdash B \implies \Gamma, \Delta_k \vdash B. \tag{1.5}$$

Um dies zu sehen, genügt es zu zeigen

$$\Gamma, \Delta_{k0} \vdash B \quad \text{und} \quad \Gamma, \Delta_{k1} \vdash B \implies \Gamma, \Delta_k \vdash B.$$

Dies ist klar in den Fällen 1 und 4, und für die Fälle 2 und 3 folgt es leicht aus den Schemata $\vee^{*-}$ und $\exists^{*-}$ (vgl. Abschnitt 1.2).

Wir zeigen jetzt

$$\Gamma, \Delta_k \vdash B \implies \exists n \forall k' {\succeq_n} k \, B \in \Delta_{k'}. \tag{1.6}$$

Um dies einzusehen, wähle man ein $n \geq \mathsf{lh}(k)$ mit $B = A_n$ und $\Gamma_n, \Delta_k \vdash_n A_n$. Für alle $k' \succeq k$ mit $\mathsf{lh}(k') = n + 1$ gilt dann $A_n \in \Delta_{k'}$ (vgl. die Fälle 2-4).

Mit Hilfe der Mengen Δ_k können wir jetzt eine $\mathcal{L}$-Beth-Struktur $\mathcal{B}$ definieren als $(\mathsf{Ter}_{\mathcal{L}}, I_0, I_1)$ mit den kanonischen $I_0(f)\boldsymbol{t} := f\boldsymbol{t}$ und

$$\boldsymbol{t} \in I_1(R, k) \quad :\iff \quad R\boldsymbol{t} \in \Delta_k.$$

Offenbar ist $t^{\mathcal{B}}[\mathsf{id}] = t$ für alle $\mathcal{L}$-Terme t.

Wir zeigen schließlich, daß

$$\Gamma, \Delta_k \vdash B \iff \mathcal{B}, k \Vdash B[\mathsf{id}], \tag{1.7}$$

und zwar durch Induktion über die logische Komplexität von B. Für $\mathcal{B}, k \Vdash B[\mathsf{id}]$ schreiben wir $k \Vdash B$.

Fall $R\boldsymbol{t}$. Die folgenden Aussagen sind äquivalent.

$$\begin{array}{ll} \Gamma, \Delta_k \vdash R\boldsymbol{t} & \\ \exists n \forall k' {\succeq_n} k \, R\boldsymbol{t} \in \Delta_{k'} & \text{nach (1.6) und (1.5)} \\ \exists n \forall k' {\succeq_n} k \, \boldsymbol{t} \in I_1(R, k') & \text{nach Definition von } \mathcal{B} \\ k \Vdash R\boldsymbol{t} & \text{nach Definition von } \Vdash, \text{ da } t^{\mathcal{B}}[\mathsf{id}] = t. \end{array}$$

Fall $B \vee^* C$. $\Longrightarrow$. Gelte $\Gamma, \Delta_k \vdash B \vee^* C$. Man wähle ein $n \geq \mathsf{lh}(k)$ mit $\Gamma_n, \Delta_k \vdash_n A_n = B \vee^* C$. Für alle $k' \succeq k$ mit $\mathsf{lh}(k') = n$ gilt dann

$$\Delta_{k'0} = \Delta_{k'} \cup \{B \vee^* C, B\} \quad \text{und} \quad \Delta_{k'1} = \Delta_{k'} \cup \{B \vee^* C, C\},$$

also nach IH

$$k'0 \Vdash B \quad \text{und} \quad k'1 \Vdash C.$$

Nach Definition impliziert dies $k \Vdash B \vee^* C$. $\Longleftarrow$.

$$\begin{array}{ll} k \Vdash B \vee^* C & \\ \exists n \forall k' {\succeq_n} k \,. k' \Vdash B \text{ oder } k' \Vdash C & \\ \exists n \forall k' {\succeq_n} k \,. \Gamma, \Delta_{k'} \vdash B \text{ oder } \Gamma, \Delta_{k'} \vdash C & \text{nach IH} \\ \exists n \forall k' {\succeq_n} k \, \Gamma, \Delta_{k'} \vdash B \vee^* C & \\ \Gamma, \Delta_k \vdash B \vee^* C & \text{nach (1.5).} \end{array}$$

Der Fall $B \wedge C$ ist klar.

Fall $B \to C$. $\Longrightarrow$. Gelte $\Gamma, \Delta_k \vdash B \to C$. Wir müssen zeigen $k \Vdash B \to C$, d.h.

$$\forall k' {\succeq} k . k' \Vdash B \implies k' \Vdash C.$$

Sei also $k' \succeq k$ und gelte $k' \Vdash B$. Nach IH habe wir $\Gamma, \Delta_{k'} \vdash B$, also $\Gamma, \Delta_{k'} \vdash C$ nach Annahme. Wieder die IH liefert $k' \Vdash C$.

$\Longleftarrow$. Gelte $k \Vdash B \to C$, d.h. $\forall k' {\succeq} k . k' \Vdash B \implies k' \Vdash C$. Wir müssen zeigen $\Gamma, \Delta_k \vdash B \to C$. Dafür wollen wir (1.5) verwenden. Man wähle ein $n \geq \mathsf{lh}(k)$ mit $B = A_n$. Sei $k' \succeq_m k$ beliebig, wobei $m := n - \mathsf{lh}(k)$. Wir müssen zeigen $\Gamma, \Delta_{k'} \vdash B \to C$.

Im Fall $\Gamma, \Delta_{k'} \vdash_n A_n$ haben wir $k' \Vdash B$ nach IH, also $k' \Vdash C$ nach Annahme, also $\Gamma, \Delta_{k'} \vdash C$ wieder nach IH und deshalb $\Gamma, \Delta_{k'} \vdash B \to C$.

Im Fall $\Gamma, \Delta_{k'} \not\vdash_n A_n$ haben wir nach Definition $\Delta_{k'1} = \Delta_{k'} \cup \{B\}$. Dies liefert $\Gamma, \Delta_{k'1} \vdash B$, also $k'1 \Vdash B$ nach IH, also $k'1 \Vdash C$ nach Annahme, also $\Gamma, \Delta_{k'1} \vdash C$ wieder nach IH. Wegen $\Delta_{k'1} = \Delta_{k'} \cup \{B\}$ folgt $\Gamma, \Delta_{k'} \vdash B \to C$.

Fall $\forall x B$. Die folgenden Aussagen sind äquivalent.

$$\begin{array}{ll}
\Gamma, \Delta_k \vdash \forall x B & \\
\forall t{\in}\mathsf{Ter}_{\mathcal{L}}\, \Gamma, \Delta_k \vdash B[x := t] & \\
\forall t{\in}\mathsf{Ter}_{\mathcal{L}}\, k \Vdash B[x := t] & \text{nach IH} \\
\forall t{\in}\mathsf{Ter}_{\mathcal{L}}\, k \Vdash B[\mathsf{id}_x^t] & \text{nach dem Substitutionslemma, da } t^{\mathcal{B}}[\mathsf{id}] = t, \\
k \Vdash \forall x B & \text{nach Definition von } \Vdash.
\end{array}$$

Fall $\exists^* xB$. Dies ist ähnlich zum Fall $\vee^*$. Im einzelnen verläuft der Beweis wie folgt. $\Longrightarrow$. Gelte $\Gamma, \Delta_k \vdash \exists^* xB$. Man wähle $n \geq \mathsf{lh}(k)$ mit $\Gamma_n, \Delta_k \vdash_n A_n = \exists^* xB$. Für alle $k' \succeq k$ mit $\mathsf{lh}(k') = n$ gilt dann

$$\Delta_{k'0} = \Delta_{k'1} = \Delta_k \cup \{\exists^* xB, B[x := x_i]\}$$

mit x_i nicht frei in $\Delta_k \cup \{\exists^* xB\}$, also nach IH

$$k'0 \Vdash B[x := x_i] \quad \text{und} \quad k'1 \Vdash B[x := x_i].$$

Nach Definition impliziert dies $k \Vdash \exists^* xB$. $\Longleftarrow$.

$$\begin{array}{ll}
k \Vdash \exists^* xB & \\
\exists n \forall k' {\succeq_n} k \exists t{\in}\mathsf{Ter}_{\mathcal{L}}\, k' \Vdash B[\mathsf{id}_x^t] & \\
\exists n \forall k' {\succeq_n} k \exists t{\in}\mathsf{Ter}_{\mathcal{L}}\, k' \Vdash B[x := t] & \\
\exists n \forall k' {\succeq_n} k \exists t{\in}\mathsf{Ter}_{\mathcal{L}}\, \Gamma, \Delta_{k'} \vdash B[x := t] & \text{nach IH} \\
\exists n \forall k' {\succeq_n} k\, \Gamma, \Delta_{k'} \vdash \exists^* xB & \\
\Gamma, \Delta_k \vdash \exists^* xB & \text{nach (1.5).}
\end{array}$$

Jetzt können wir den Beweis des Vollständigkeitssatzes abschließen. Nach (1.7) haben wir $\mathcal{B}, \langle\rangle \Vdash \Gamma[\mathsf{id}]$. Nach Annahme impliziert dies $\mathcal{B}, \langle\rangle \Vdash A[\mathsf{id}]$, also $\Gamma \vdash A$ wieder nach (1.7). □

Als ein unmittelbares Korollar erhalten wir den Vollständigkeitssatz für die intuitionistische Logik.

Korollar 1.4.2. *Sei $\Gamma \cup \{A\}$ eine Formelmenge. Die folgenden Aussagen sind äquivalent.*

1. $\Gamma \vdash_i A$.
2. $\Gamma, \mathsf{Efq} \Vdash A$, *d.h. für alle* BETH-*Strukturen $\mathcal{B}$ für die intuitionistische Logik, Knoten k und Belegungen η*

$$\mathcal{B}, k \Vdash \Gamma[\eta] \implies \mathcal{B}, k \Vdash A[\eta].$$ □

Kripke-Strukturen

KRIPKE-Strukturen wurden von KRIPKE in [17] eingeführt als eine Alternative zu BETH-Strukturen. Sie sind einfacher in dem Sinn, daß die Klauseln für atomare Formeln, $\vee^*$ und $\exists^*$ sich nicht auf eine "Schranke" beziehen, sondern nur auf den gegenwärtigen Knoten. Sie sind jedoch auch komplizierter, da die Bereiche mit den Knoten k anwachsen und nicht mehr konstant sind. Besonders geeignet sind KRIPKE-Strukturen für die Konstruktion endlicher Gegenmodelle. Der Einfachheit halber wollen wir hier annehmen, daß keine Funktionssymbole vorkommen.

Definition 1.4.3. Sei $(T, \preceq)$ ein endlich verzweigter Baum. $\mathcal{K} = (D, I_0, I_1)$ ist eine $\mathcal{L}$-KRIPKE-*Struktur*, wenn D jedem Knoten $k \in T$ eine nichtleere Menge $D(k)$ zuordnet so daß aus $k \preceq k'$ folgt daß $(D(k), I_0(k))$ eine $\mathcal{L}$-sub-Prästruktur von $(D(k'), I_0(k'))$ ist, und I_1 jedem n-stelligen Relationssymbol R von $\mathcal{L}$ und jedem Knoten $k \in T$ eine n-stellige Relation

$$I_1(R, k) \subseteq D(k)^n$$

zuordnet so daß Monotonie gilt, d.h.

$$k \preceq k' \implies I_1(R, k) \subseteq I_1(R, k').$$

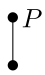
Im Fall $n = 0$ ist $I_1(R,k)$ wahr oder falsch und die Monotonie besagt, daß für $k \preceq k'$ aus $I_1(R,k)$ folgt $I_1(R,k')$.

Eine Belegung η für eine Kripke-Struktur $\mathcal{K}$ ist eine Abbildung, die jeder Variablen ein Element aus $\bigcup_{k\in T} D(k)$ zuordnet. Der Wert $t^{\mathcal{K}}[\eta]$ eines Terms t unter der Belegung η ist an allen Knoten k mit $\eta[\mathsf{vars}(t)] \subseteq D(k)$ erklärt.

Jetzt können wir die Erzwingungsbeziehung definieren. $\mathcal{K}, k \Vdash A[\eta]$ ($\mathcal{K}$ am Knoten k erzwingt A für die Belegung η) wird induktiv wie folgt definiert. Wir schreiben $k \Vdash A[\eta]$ wenn sich $\mathcal{K}$ aus dem Zusammenhang ergibt.

$$k \Vdash R(t_1,\dots,t_p)[\eta] \quad :\Longleftrightarrow$$
$$\eta[\mathsf{vars}(t_1,\dots,t_p)] \subseteq D(k) \text{ und } (t_1^{\mathcal{K}}[\eta],\dots,t_p^{\mathcal{K}}[\eta]) \in I_1(R,k) \quad \text{für } R \text{ nicht nullstellig.}$$
$$k \Vdash R[\eta] \quad :\Longleftrightarrow \quad I_1(R,k) = 1 \quad \text{für } R \text{ nullstellig.}$$
$$k \Vdash (A \vee^* B)[\eta] \quad :\Longleftrightarrow \quad k \Vdash A[\eta] \text{ oder } k' \Vdash B[\eta].$$
$$k \Vdash (\exists^* x A)[\eta] \quad :\Longleftrightarrow \quad \exists a{\in}D(k)\, k \Vdash A[\eta_x^a].$$
$$k \Vdash (A \to B)[\eta] \quad :\Longleftrightarrow \quad \forall k'{\succeq}k.k' \Vdash A[\eta] \implies k' \Vdash B[\eta].$$
$$k \Vdash (A \wedge B)[\eta] \quad :\Longleftrightarrow \quad k \Vdash A[\eta] \text{ und } k \Vdash B[\eta].$$
$$k \Vdash (\forall x A)[\eta] \quad :\Longleftrightarrow \quad \forall k'{\succeq}k \forall a{\in}D(k')\, k' \Vdash A[\eta_x^a].$$

Die Vervollständigung von Beth-Strukturen macht einen Vergleich mit Kripke-Strukturen einfacher. Eine *vervollständigte* Beth-Struktur ist eine Kripke-Struktur mit konstanten Bereichen, und beide Erzwingungsbegriffe stimmen für Formeln ohne $\vee^*$, $\exists^*$ überein. Man beachte aber, daß Kripke-Strukturen etwas flexibler sind, wenn man (endliche) Gegenmodelle konstruieren will. Ein Grund liegt darin, daß nicht jede Kripke-Struktur mit konstanten Bereichen eine vervollständigte Beth-Struktur ist. Ein offensichtliches Beispiel ist



Kripke-Strukturen sind jedoch weniger allgemein als Beth-Strukturen, in dem Sinn daß jede abzählbare Kripke-Struktur in einem angemessenen Sinn in eine Beth-Struktur mit konstantem Bereich $\mathbb{N}$ "transformiert" werden kann; dies wurde bereits von Kripke in [17] bemerkt. Wir beschreiben jetzt diese Transformation, wobei wir uns auf Troelstra und van Dalen [44] stützen.

Sei $\mathcal{K} = (D, I_0, I_1)$ eine abzählbare Kripke-Struktur mit zugrunde liegendem endlich verzweigtem Baum $(T, \preceq)$. Wir transformieren $\mathcal{K}$ in eine Beth-Struktur $\mathcal{B}$ über dem endlich verzweigten Baum T', der aus allen endlichen aufsteigenden oder gleichbleibenden Folgen über $(T, \preceq)$ der Länge ≥ 0 besteht und geordnet ist durch $\sigma \preceq' \tau$ gdw σ ein Anfangsstück von τ ist.

Man zerlege $\mathbb{N}$ in abzählbar unendlich viele abzählbar unendliche disjunkte Mengen N_i, d.h. $\mathbb{N} = \bigcup_{i\in\mathbb{N}} N_i$ und $i \neq j \to N_i \cap N_j = \emptyset$, und setze $M_i := \bigcup_{j\leq i} N_i$. Offenbar können wir für jedes $\sigma = \langle k_0,\dots,k_n\rangle \in T'$ eine Surjektion ψ_σ von M_n auf $D(k_n)$ definieren mit

$\psi_{\sigma k} a = \psi_\sigma a$ für $a \in M_{\mathsf{lh}(\sigma)-1}$,
$\psi_{\sigma k}$ bildet $M_{\mathsf{lh}(\sigma)}$ auf $D(k)$ ab.

Wir definieren jetzt unsere neue Beth-Struktur wie folgt. Für den leeren Knoten $\sigma = \varepsilon$ sei $I_1(R,\sigma)$ leer. Sei jetzt $\sigma = \langle k_0,\dots,k_n\rangle$.

$$(a_1,\dots,a_p) \in I_1(R,\sigma) \quad :\Longleftrightarrow \quad a_1,\dots,a_p \in M_n \text{ und } k_n \Vdash R(\psi_\sigma a_1,\dots,\psi_\sigma a_p)$$

(oder genauer, $k_n \Vdash R(x_1,\dots,x_p)[\psi_\sigma a_1/x_1,\dots,\psi_\sigma a_p/x_p]$; allgemein schreiben wir $k \Vdash A(a_1,\dots,a_p)$ für $k \Vdash A[\psi_\sigma a_1/x_1,\dots,\psi_\sigma a_p/x_p]$, falls $x_1,\dots,x_p$ die freien Variablen von A enthalten). Damit ist offenbar eine Beth-Struktur $\mathcal{B}$ definiert, deren Erzwingungsrelation wir mit $\Vdash'$ bezeichnen.

Lemma 1.4.4. *Sei $\sigma = \langle k_0,\dots,k_n\rangle$. Dann gilt für alle $a_1,\dots,a_p \in M_n$*

$$\sigma \Vdash' A(a_1,\dots,a_p) \iff k_n \Vdash A(\psi_\sigma a_1,\dots,\psi_\sigma a_p).$$

Beweis. Durch Induktion über A. *Fall* $\forall x B$. $\Longrightarrow$. Gelte $\sigma \Vdash' \forall x B(x, a')$ mit $a' \in M_n$. Wir haben zu zeigen $k_n \Vdash \forall x B(x, \psi_\sigma a')$. Sei also $k' \succeq k_n$ und $d \in D(k')$ gegeben; wir müssen zeigen $k' \Vdash B(d, \psi_\sigma a')$. Da $d = \psi_{\sigma k'} b$ für ein $b \in M_{n+1}$ und $\psi_\sigma a' = \psi_{\sigma k'} a'$, heißt dies $k' \Vdash B(\psi_{\sigma k'} b, \psi_{\sigma k'} a')$. Nach IH genügt es zu zeigen $\sigma k' \Vdash' B(b, a')$. Dies folgt aber aus unserer Annahme.

$\Longleftarrow$. Gelte $k_n \Vdash \forall x B(x, \psi_\sigma a')$ mit $a' \in M_n$. Wir müssen zeigen $\sigma \Vdash' \forall x B(x, a')$. Sei also $a \in \mathbb{N}$ gegeben; wir müssen dann zeigen $\sigma \Vdash' B(a, a')$. Man beachte, daß $\sigma = \langle k_0, \dots, k_n \rangle$. Wähle ein hinreichend langes $\tau = \sigma * \langle k_{n+1}, \dots, k_{n+i}, k' \rangle$ mit $k_n \preceq k'$ so daß $a \in M_{\mathsf{lh}(\tau)-1}$. Man beachte, daß ψ_τ die Menge $M_{\mathsf{lh}(\tau)-1}$ auf $D(k')$ abbildet und $\psi_\tau a' = \psi_\sigma a' \in D(k_n)$. Aus $k_n \Vdash \forall x B(x, \psi_\sigma a')$ und $k_n \preceq k'$ ergibt sich $k' \Vdash \forall x B(x, \psi_\tau a')$, also $k' \Vdash B(\psi_\tau a, \psi_\tau a')$, also nach IH $\tau \Vdash' B(a, a')$. Dies gilt für alle Fortsetzungen τ von σ mit $a \in M_{\mathsf{lh}(\tau)-1}$ (d.h. mit mindestens dieser Länge). Nach der Überdeckungseigenschaft 1.3.7 gilt also $\sigma \Vdash' B(a, a')$, wie gewünscht.

Die anderen Fälle sind einfacher und seien als Übungsaufgaben gestellt. $\square$

1.5 Vollständigkeit der klassischen Logik

Wir geben jetzt einen Beweis der Vollständigkeit der klassischen Logik, und zwar unter Zuhilfenahme der Vollständigkeit der Minimallogik.

Zur Vereinfachung zeigen wir zunächst, daß man auf $\wedge$ verzichten kann.

Lemma 1.5.1. *(Elimination von $\wedge$). Für jede Formel A in der auf $\{\to, \wedge, \forall\}$ aufgebauten Sprache findet man Formeln $A_1, \dots, A_n$ ohne $\wedge$ so daß $\vdash A \leftrightarrow \bigwedge\!\!\!\bigwedge_{i=1}^n A_i$.*

Beweis. Induktion über A. *Fall* $\boldsymbol{Rt}$. Wähle $n = 1$ und $A_1 := \boldsymbol{Rt}$. *Fall* $A \wedge B$. Nach IH haben wir $A_1, \dots, A_n$ und $B_1, \dots, B_m$. Wähle $A_1, \dots, A_n, B_1, \dots, B_m$. *Fall* $A \to B$. Nach IH haben wir wieder $A_1, \dots, A_n$ und $B_1, \dots, B_m$. Zur Vereinfachung der Schreibweise nehmen wir $n = 2$ und $m = 3$ an. Dann gilt

$$\begin{aligned} &\vdash (A_1 \wedge A_2 \to B_1 \wedge B_2 \wedge B_3) \\ &\leftrightarrow (A_1 \to A_2 \to B_1) \wedge (A_1 \to A_2 \to B_2) \wedge (A_1 \to A_2 \to B_3). \end{aligned}$$

Fall $\forall x A$. Nach IH für A haben wir $A_1, \dots, A_n$. Wähle $\forall x A_1, \dots, \forall x A_n$, für

$$\vdash \forall x \bigwedge\!\!\!\bigwedge_{i=1}^{n} A_i \leftrightarrow \bigwedge\!\!\!\bigwedge_{i=1}^{n} \forall x A_i. \qquad \square$$

Satz 1.5.2. *(Vollständigkeit). Sei $\Gamma \cup \{A\}$ eine Menge von Formeln (in unserer abzählbaren Sprache $\mathcal{L}$). Die folgenden Aussagen sind äquivalent.*

1. *$\Gamma \vdash_c A$.*
2. *$\Gamma \models A$, d.h. für alle Strukturen $\mathcal{M}$ and Belegungen η gilt*

$$\mathcal{M} \models \Gamma[\eta] \implies \mathcal{M} \models A[\eta].$$

Beweis. Eine Richtung ist der Korrektheitssatz. Für die andere Richtung verwenden wir den Vollständigkeitssatz für die Minimallogik.

Offenbar genügt es, Formeln ohne $\vee^*$, $\exists^*$ zu betrachten, und (nach Lemma 1.5.1) auch ohne $\wedge$.

Gelte $\Gamma \nvdash_c A$, d.h. $\Gamma, \mathsf{Stab} \nvdash A$. Nach dem Vollständigkeitssatz für die Minimallogik haben wir eine abzählbare BETH-Struktur $\mathcal{B} = (\mathsf{Ter}_\mathcal{L}, I_0, I_1)$ über dem vollen binären Baum T_{01} und einen Knoten l_0 mit $l_0 \Vdash \Gamma, \mathsf{Stab}$ und $l_0 \nVdash A$ (wir schreiben $k \Vdash B$ für $\mathcal{B}, k \Vdash B[\mathsf{id}]$).

Ein Knoten k heisse *konsistent* wenn $k \nVdash \bot$, und *stabil* wenn $k \Vdash \mathsf{Stab}$. Sei k ein stabiler Knoten und B eine Formel (ohne $\vee^*, \exists^*$). Wir haben die Stabilität $\mathsf{Stab} \vdash \neg\neg B \to B$ nach Lemma 1.2.2, also $k \Vdash \neg\neg B \to B$, also

$$\begin{aligned} k \nVdash B &\iff k \nVdash \neg\neg B \\ &\iff \exists k' {\succeq} k . k' \text{ konsistent und } k' \Vdash \neg B. \end{aligned} \tag{1.8}$$

Sei α ein Ast im zugrunde liegende Baum T_{01}. Wir definieren

$$\begin{aligned} \alpha \Vdash A \quad &:\Longleftrightarrow \quad \exists k{\in}\alpha\, k \Vdash A, \\ \alpha \text{ is konsistent} \quad &:\Longleftrightarrow \quad \alpha \nVdash \bot, \\ \alpha \text{ is stabil} \quad &:\Longleftrightarrow \quad \exists k{\in}\alpha\, k \Vdash \mathsf{Stab}. \end{aligned}$$

Man beachte

$$\text{Aus } \alpha \Vdash \boldsymbol{A} \text{ und } \vdash \boldsymbol{A} \to B \text{ folgt } \alpha \Vdash B. \tag{1.9}$$

Um dies zu sehen, nehmen wir $\alpha \Vdash \boldsymbol{A}$ an. Dann gilt $k \Vdash \boldsymbol{A}$ für ein $k \in \alpha$, da α linear geordnet ist. Wegen $\vdash \boldsymbol{A} \to B$ liefert der Korrektheitssatz $k \Vdash B$, d.h. $\alpha \Vdash B$.

Ein Ast α heißt *generisch* (in dem Sinn, daß er ein klassisches Modell erzeugt) wenn er konsistent und stabil ist, für alle Formeln B gilt

$$\alpha \Vdash B \text{ oder } \alpha \Vdash \neg B, \tag{1.10}$$

und für alle Formeln $\forall \boldsymbol{y} B$ (wobei $\boldsymbol{y}$ nicht leer ist) mit B keine Allformel

$$\forall \boldsymbol{s}{\in}\mathsf{Ter}_{\mathcal{L}}\, \alpha \Vdash B[\boldsymbol{y} := \boldsymbol{s}] \implies \alpha \Vdash \forall \boldsymbol{y} B \tag{1.11}$$

Schließlich definieren wir für einen Ast α eine klassische Struktur $\mathcal{M}^\alpha = (\mathsf{Ter}_{\mathcal{L}}, I_0, I_1^\alpha)$ durch

$$I_1^\alpha(R) := \bigcup_{k \in \alpha} I_1(R, k) \quad \text{für } R \neq \bot.$$

Wir zeigen, daß für jeden generischen Ast α und jede Formel B in der auf $\{\to, \forall\}$ aufbauenden Sprache gilt

$$\alpha \Vdash B \iff \mathcal{M}^\alpha \models B. \tag{1.12}$$

Der Beweis erfolgt durch Induktion über die logische Komplexität von B.

Fall $R\boldsymbol{t}$, $R \neq \bot$. Dann gilt die Behauptung für alle α.

Fall $\bot$. Es gilt $\alpha \nVdash \bot$ für konsistentes α.

Fall $B \to C$. $\Longrightarrow$. Gelte $\alpha \Vdash B \to C$ und $\mathcal{M}^\alpha \models B$. Wir müssen zeigen $\mathcal{M}^\alpha \models C$. Nun ist $\alpha \Vdash B$ nach IH, also $\alpha \Vdash C$, also $\mathcal{M}^\alpha \models C$ wieder nach IH. $\Longleftarrow$. Gelte $\mathcal{M}^\alpha \models B \to C$. Gilt $\mathcal{M}^\alpha \models B$, so ist $\mathcal{M}^\alpha \models C$, also $\alpha \Vdash C$ nach IH und deshalb $\alpha \Vdash B \to C$. Gilt $\mathcal{M}^\alpha \not\models B$, so ist $\alpha \nVdash B$ nach IH, also $\alpha \Vdash \neg B$ nach (1.10) und deshalb $\alpha \Vdash B \to C$, da α stabil ist (und $\vdash (\neg\neg C \to C) \to \bot \to C$).

Fall $\forall \boldsymbol{y} B$ (wobei $\boldsymbol{y}$ nicht leer ist) mit B keine Allformel. Die folgenden Aussagen sind äquivalent.

$$\begin{array}{ll} \alpha \Vdash \forall \boldsymbol{y} B & \\ \forall \boldsymbol{s}{\in}\mathsf{Ter}_{\mathcal{L}}\, \alpha \Vdash B[\boldsymbol{y} := \boldsymbol{s}] & \text{nach (1.11)} \\ \forall \boldsymbol{s}{\in}\mathsf{Ter}_{\mathcal{L}}\, \mathcal{M}^\alpha \models B[\boldsymbol{y} := \boldsymbol{s}] & \text{nach IH} \\ \mathcal{M}^\alpha \models \forall \boldsymbol{y} B. & \end{array}$$

Wir zeigen schließlich, daß es für jeden konsistenten stabilen Knoten k einen generischen Ast gibt, der k enthält. Zum Beweis sei $A_0, A_1, \dots$ eine Aufzählung aller Formeln. Wir definieren induktiv eine Folge $k = k_0 \preceq k_1 \preceq k_2 \dots$ von konsistenten stabilen Knoten. Sei $k_0 := k$. Nehmen wir jetzt an, daß k_n bereits konstruiert ist. Wir schreiben A_n in der Form $\forall \boldsymbol{y} B$ (wobei $\boldsymbol{y}$ leer sein kann) mit B keine Allformel. Im Fall $k_n \Vdash \forall \boldsymbol{y} B$ sei $k_{n+1} := k_n$. Andernfalls gilt $k_n \nVdash B[\boldsymbol{y} := \boldsymbol{s}]$ für ein $\boldsymbol{s}$, und nach (1.8) gibt es einen konsistenten Knoten $k' \succeq k_n$ mit $k' \Vdash \neg B[\boldsymbol{y} := \boldsymbol{s}]$. Sei $k_{n+1} := k'$. Wegen $k_n \preceq k_{n+1}$ ist auch k_{n+1} stabil.

Sei $\alpha := \{\, l \mid \exists n\, l \preceq k_n \,\}$, also $k \in \alpha$. Wir zeigen, daß α generisch ist. Offenbar ist α konsistent und stabil. Die Aussagen (1.10) und (1.11) können simultan bewiesen werden. Sei $C = \forall \boldsymbol{y} B$ mit B keine Allformel, und man wähle n mit $C = A_n$. Im Fall $k_n \Vdash \forall \boldsymbol{y} B$ ist nichts zu zeigen. Andernfalls gilt $k_n \nVdash B[\boldsymbol{y} := \boldsymbol{s}]$ für ein $\boldsymbol{s}$, und nach Konstruktion $k_{n+1} \Vdash \neg B[\boldsymbol{y} := \boldsymbol{s}]$. Für (1.10) erhalten wir $k_{n+1} \Vdash \neg \forall \boldsymbol{y} B$ (da $\vdash \forall \boldsymbol{y} B \to B[\boldsymbol{y} := \boldsymbol{s}]$), und (1.11) ergibt sich aus der Konsistenz von α.

Wir können jetzt den Beweis des Vollständigkeitssatzes abschließen. Da $l_0 \nVdash A$ und l_0 stabil ist, liefert (1.8) einen konsistenten Knoten $k \succeq l_0$ mit $k \Vdash \neg A$. Offenbar ist auch k stabil. Nach dem, was wir eben bewiesen haben, gibt es einen generischen Ast α mit $k \in \alpha$. Wegen $k \Vdash \neg A$ haben wir $\alpha \Vdash \neg A$, also $\mathcal{M}^\alpha \models \neg A$ nach (1.12). Ferner gilt $\alpha \Vdash \Gamma$, also $\mathcal{M}^\alpha \models \Gamma$ wieder nach (1.12). Also $\Gamma \not\models A$. $\square$

Auf den ersten Blick scheint es, als ob der Vollständigkeitssatz sich auf den Berich “aller Mengen” bezieht, und zwar in der Definition von $\mathcal{M} \models A$. Eine genauere Betrachtung des Beweises zeigt jedoch, daß es genügt, sich auf spezielle abzählbare Mengen zu beschränken. Der Bereich des Modells ist die abzählbare Menge $\mathsf{Ter}_{\mathcal{L}}$ (die mittels einer geeigneten Kodierung durch $\mathbb{N}$ ersetzt werden kann), und man überlegt sich leicht, daß alle verwendeten Begriffe arithmetisch definierbar sind.

Der Vollständigkeitssatz hat viele wichtige Korollare; wir erwähnen nur einige davon. Eine Menge Γ von $\mathcal{L}$-Formeln heisse *konsistent* wenn $\Gamma \nvdash_c \bot$, und *erfüllbar* wenn es eine $\mathcal{L}$-Struktur $\mathcal{M}$ und eine Belegung η in $|\mathcal{M}|$ gibt mit $\mathcal{M} \models B[\eta]$ für alle $B \in \Gamma$.

Korollar 1.5.3. *Sei Γ eine Menge von $\mathcal{L}$-Formeln.*

1. *Wenn Γ konsistent ist, so ist Γ auch erfüllbar.*
2. *(Kompaktheitssatz). Wenn jede endliche Teilmenge von Γ erfüllbar ist, so auch Γ.*

Beweis. 1. Aus $\Gamma \nvdash_c \bot$ erhält man $\Gamma \not\models \bot$ nach dem Vollständigkeitssatz, und dies impliziert die Erfüllbarkeit von Γ.

2. Andernfalls gilt $\Gamma \models \bot$, also $\Gamma \vdash_c \bot$ nach dem Vollständigkeitssatz, also auch $\Gamma_0 \vdash_c \bot$ für eine endliche Teilmenge $\Gamma_0 \subseteq \Gamma$, also $\Gamma_0 \models \bot$ im Widerspruch zu unserer Annahme, daß Γ_0 ein Modell besitzt. □

Korollar 1.5.4. *(*LÖWENHEIM*,* SKOLEM*). Sei Γ eine Menge von $\mathcal{L}$-Formeln (wir hatten angenommen, daß $\mathcal{L}$ abzählbar ist). Ist Γ erfüllbar, so ist Γ auch erfüllbar durch eine $\mathcal{L}$-Struktur mit abzählbarer Trägermenge.*

Beweis. Wir verwenden den Beweis des Vollständigkeitssatzes mit $A = \bot$. Er liefert entweder $\Gamma \vdash_c \bot$ oder aber ein Modell von $\Gamma \cup \{\neg\bot\}$, dessen Trägermenge die abzählbare Menge $\mathsf{Ter}_{\mathcal{L}}$ ist. $\Gamma \vdash_c \bot$ kann jedoch aufgrund der Annahme nicht gelten. □

Der Vollständigkeitssatz für überabzählbare Sprachen

Wir geben noch einen zweiten Beweis des Vollständigkeitssatzes, der das Resultat auch im Fall überabzählbarer Sprachen liefert. Dieser Beweis verwendet als mengentheoretisches Hilfsmittel das Auswahlaxiom (in der Form des ZORNschen Lemmas).

Sei $M \neq \emptyset$ eine Menge. $F \subseteq \mathcal{P}(M)$ heißt *Filter* über M, wenn

1. $M \in F$ und $\emptyset \notin F$;
2. Aus $X \in F$ und $X \subseteq Y \subseteq M$ folgt $Y \in F$;
3. Sind $X, Y \in F$, so ist auch $X \cap Y \in F$.

F heißt *Ultrafilter*, wenn für alle $X \in \mathcal{P}(M)$ gilt

$$X \in F \text{ oder } M \setminus X \in F.$$

Die Intuition hier ist, daß die Elemente X eines Filters F in einem gewissen Sinn “groß” sind. Zum Beispiel ist für unendliches M die Menge $F = \{\, X \subseteq M \mid M \setminus X \text{ endlich} \,\}$ ein Filter.

Lemma 1.5.5. *Ist F ein Ultrafilter und $X \cup Y \in F$, so folgt $X \in F$ oder $Y \in F$.*

Beweis. Übung. □

Lemma 1.5.6. *Sei $M \neq \emptyset$ eine Menge und $S \subseteq \mathcal{P}(M)$. S hat die* endliche Durchschnittseigenschaft*, wenn $X_1 \cap \dots \cap X_n \neq \emptyset$ für alle $X_1, \dots, X_n \in S$ und alle $n \in \mathbb{N}$. Dann gilt: Hat S die endliche Durchschnittseigenschaft, so existiert ein Filter über M mit $F \supseteq S$.*

Beweis. Übung. □

Lemma 1.5.7. *Sei $M \neq \emptyset$ eine Menge und F ein Filter über M. Dann gibt es einen Ultrafilter U über M mit $U \supseteq M$.*

Beweis. Dies ist eine unmittelbare Folgerung aus dem ZORNschen Lemma, das wir im zweiten Teil dieser Vorlesung aus dem Auswahlaxiom beweisen werden. □

Sei $M \neq \emptyset$ eine Menge und seien $A_i \neq \emptyset$ Mengen für $i \in M$. Man setzt

$$\prod_{i \in M} A_i := \{ \alpha \mid \alpha \text{ ist Funktion, } \mathsf{dom}(\alpha) = M \text{ und } \alpha(i) \in A_i \text{ für alle } i \in M \}.$$

Man beachte hierbei, daß nach dem Auswahlaxiom $\prod_{i \in M} A_i \neq \emptyset$ ist. Wir schreiben $\alpha \in \prod_{i \in M} A_i$ als $\langle \alpha(i) \mid i \in M \rangle$.

Seien nun $M \neq \emptyset$ eine Menge, F ein Filter über M und $\mathcal{A}_i$ Strukturen für $i \in M$. Dann ist die *F-Produktstruktur* $\mathcal{A} = \prod_{i \in M}^F \mathcal{A}_i$ definiert durch

1. $|\mathcal{A}| := \prod_{i \in M} |\mathcal{A}_i|$ (man beachte $|\mathcal{A}| \neq \emptyset$).
2. Für ein n-stelliges Relationssymbol R und $\alpha_1, \dots, \alpha_n \in |\mathcal{A}|$ sei
$$R^{\mathcal{A}}(\alpha_1, \dots, \alpha_n) \quad :\Longleftrightarrow \quad \{ i \in M \mid R^{\mathcal{A}_i}(\alpha_1(i), \dots, \alpha_n(i)) \} \in F.$$
3. Für ein n-stelliges Funktionssymbol f und $\alpha_1, \dots, \alpha_n \in |\mathcal{A}|$ sei
$$f^{\mathcal{A}}(\alpha_1, \dots, \alpha_n) := \langle f^{\mathcal{A}_i}(\alpha_1(i), \dots, \alpha_n(i)) \mid i \in M \rangle.$$

Ist F ein Ultrafilter, so heißt $\mathcal{A} = \prod_{i \in M}^F \mathcal{A}_i$ das *F-Ultraprodukt* der $\mathcal{A}_i$ für $i \in M$.

Die Eigenschaften von Ultrafiltern spiegeln in gewisser Weise die Definition der Folgerungsbeziehung $\models$ wieder. Zum Beispiel gilt

$$\begin{aligned} \mathcal{M} \models (A \wedge B)[\eta] &\iff \mathcal{M} \models A[\eta] \text{ und } \mathcal{M} \models B[\eta] \\ X \cap Y \in F &\iff X \in F \text{ und } Y \in F \end{aligned}$$

und

$$\begin{aligned} \mathcal{M} \models \neg A[\eta] &\iff \mathcal{M} \not\models A[\eta] \\ X \notin F &\iff M \setminus X \in F. \end{aligned}$$

Dies ist der Hintergrund des folgenden Satzes.

Satz 1.5.8. *(Fundamentalsatz über Ultraprodukte,* ŁOŚ *1955). Sei $\mathcal{A} = \prod_{i \in M}^F \mathcal{A}_i$ ein F-Ultraprodukt, A eine Formel und η eine Belegung in $|\mathcal{A}|$. Dann gilt*

$$\mathcal{A} \models A[\eta] \iff \{ i \in M \mid \mathcal{A}_i \models A[\eta_i] \} \in F,$$

wobei η_i die durch $\eta_i(x) = \eta(x)(i)$ für $i \in M$ definieren induzierten Belegungen sind.

Beweis. Wir beweisen zunächst eine entsprechende Aussage über Terme.

$$t^{\mathcal{A}}[\eta] = \langle t^{\mathcal{A}_i}[\eta_i] \mid i \in M \rangle. \tag{1.13}$$

Den Beweis führen wir durch Induktion über t. Im Fall einer Variablen folgt die Behauptung aus der Definition. *Fall* $ft_1 \dots t_n$. Zur Vereinfachung der Schreibweise nehmen wir $n = 1$ an, betrachten also ft. Man erhält

$$\begin{aligned} (ft)^{\mathcal{A}}[\eta] &= f^{\mathcal{A}}(t^{\mathcal{A}}[\eta]) \\ &= f^{\mathcal{A}}(\langle t^{\mathcal{A}_i}[\eta_i] \mid i \in M \rangle) \quad \text{nach IH} \\ &= \langle (ft)^{\mathcal{A}_i}[\eta_i] \mid i \in M \rangle. \end{aligned}$$

Fall $Rt_1 \dots t_n$. Zur Vereinfachung der Schreibweise nehmen wir wieder $n = 1$ an, betrachten also Rt. Man erhält

$$\begin{aligned}
\mathcal{A} \models Rt[\eta] &\iff R^{\mathcal{A}}(t^{\mathcal{A}}[\eta]) \\
&\iff \{\, i \in M \mid R^{\mathcal{A}_i}(t^{\mathcal{A}}[\eta](i)) \,\} \in F \\
&\iff \{\, i \in M \mid R^{\mathcal{A}_i}(t^{\mathcal{A}_i}[\eta_i]) \,\} \in F \qquad \text{nach (1.13)} \\
&\iff \{\, i \in M \mid \mathcal{A}_i \models Rt[\eta_i] \,\} \in F.
\end{aligned}$$

Fall $A \wedge B$.

$$\begin{aligned}
\mathcal{A} \models (A \wedge B)[\eta] &\iff \mathcal{A} \models A[\eta] \text{ und } \mathcal{A} \models B[\eta] \\
&\iff \{\, i \in M \mid \mathcal{A}_i \models A[\eta_i] \,\} \in F \text{ und } \{\, i \in M \mid \mathcal{A}_i \models B[\eta_i] \,\} \in F \quad \text{nach IH} \\
&\iff \{\, i \in M \mid \mathcal{A}_i \models A[\eta_i] \,\} \cap \{\, i \in M \mid \mathcal{A}_i \models B[\eta_i] \,\} \in F \\
&\iff \{\, i \in M \mid \mathcal{A}_i \models (A \wedge B)[\eta_i] \,\} \in F.
\end{aligned}$$

Fall $A \to B$.

$$\begin{aligned}
&\mathcal{A} \models (A \to B)[\eta] \\
&\iff \text{wenn } \mathcal{A} \models A[\eta], \text{ so } \mathcal{A} \models B[\eta] \\
&\iff \text{wenn } \{\, i \in M \mid \mathcal{A}_i \models A[\eta_i] \,\} \in F, \text{ so } \{\, i \in M \mid \mathcal{A}_i \models B[\eta_i] \,\} \in F \quad \text{nach IH} \\
&\iff \{\, i \in M \mid \mathcal{A}_i \models A[\eta_i] \,\} \notin F \text{ oder } \{\, i \in M \mid \mathcal{A}_i \models B[\eta_i] \,\} \in F \\
&\iff \{\, i \in M \mid \mathcal{A}_i \models \neg A[\eta_i] \,\} \in F \text{ oder } \{\, i \in M \mid \mathcal{A}_i \models B[\eta_i] \,\} \in F \qquad \text{da } F \text{ Ultrafilter ist} \\
&\iff \{\, i \in M \mid \mathcal{A}_i \models (A \to B)[\eta_i] \,\} \in F.
\end{aligned}$$

Fall $\forall x A$.

$$\begin{aligned}
\mathcal{A} \models (\forall x A)[\eta] &\iff \text{für alle } \alpha \in |\mathcal{A}| \text{ gilt } \mathcal{A} \models A[\eta_x^{\alpha}] \\
&\iff \text{für alle } \alpha \in |\mathcal{A}| \text{ gilt } \{\, i \in M \mid \mathcal{A}_i \models A[(\eta_i)_x^{\alpha(i)}] \,\} \in F \quad \text{nach IH} \\
&\iff \{\, i \in M \mid \text{für alle } a \in |\mathcal{A}_i| \text{ gilt } \mathcal{A}_i \models A[(\eta_i)_x^{a}] \,\} \in F \quad \text{Beweis siehe unten} \qquad (1.14) \\
&\iff \{\, i \in M \mid \mathcal{A}_i \models (\forall x A)[\eta_i] \,\} \in F.
\end{aligned}$$

Zu zeigen bleibt (1.14). Setze $X := \{\, i \in M \mid \text{für alle } a \in |\mathcal{A}_i| \text{ gilt } \mathcal{A}_i \models A[(\eta_i)_x^a] \,\}$ und $Y_\alpha := \{\, i \in M \mid \mathcal{A}_i \models A[(\eta_i)_x^{\alpha(i)}] \,\}$ für $\alpha \in |\mathcal{A}|$.

$\Longleftarrow$. Sei $\alpha \in |\mathcal{A}|$ und $X \in F$. Offenbar ist $X \subseteq Y_\alpha$, also auch $Y_\alpha \in F$. $\Longrightarrow$. Sei $Y_\alpha \in F$ für alle α. Annahme: $X \notin F$. Da F Ultrafilter ist, muß dann

$$M \setminus X = \{\, i \in M \mid \text{es gibt ein } a \in |\mathcal{A}_i| \text{ mit } \mathcal{A}_i \not\models A[(\eta_i)_x^a] \,\} \in F$$

sein. Wir wählen mit Hilfe des Auswahlaxioms ein $\alpha_0 \in |\mathcal{A}|$ mit

$$\alpha_0(i) = \begin{cases} \text{ein } a \in |\mathcal{A}_i| \text{ mit } \mathcal{A}_i \not\models A[(\eta_i)_x^a] & \text{falls } i \in M \setminus X, \\ \text{beliebig} \in |\mathcal{A}_i| & \text{sonst.} \end{cases}$$

Dann ist $Y_{\alpha_0} \cap (M \setminus X) = \emptyset$, im Widerspruch zu $Y_{\alpha_0}, M \setminus X \in F$. □

Wählt man $\mathcal{A}_i = \mathcal{B}$ konstant, so erfüllt $\mathcal{A} = \prod_{i \in M}^F \mathcal{B}$ dieselben Formeln wie $\mathcal{B}$ (solche Strukturen werden wir in Abschnitt 1.6 elementar äquivalent nennen; Schreibweise $\mathcal{A} \equiv \mathcal{B}$). $\prod_{i \in M}^F \mathcal{B}$ heißt *Ultrapotenz* von $\mathcal{B}$.

Korollar 1.5.9. *(Allgemeiner Kompaktheitssatz). Jede endlich erfüllbare Formelmenge Γ ist erfüllbar.*

Beweis. Sei $M := \{\, i \subseteq \Gamma \mid i \text{ endlich} \,\}$. Für $i \in M$ sei $\mathcal{A}_i$ ein Modell von i unter der Belegung η_i. Für $A \in \Gamma$ setzen wir $Z_A := \{\, i \in M \mid A \in i \,\} = \{\, i \subseteq \Gamma \mid i \text{ endlich und } A \in i \,\}$. Dann hat $F^- := \{\, Z_A \mid A \in \Gamma \,\}$ die endliche Durchschnittseigenschaft. Nach den Lemmata 1.5.6 und 1.5.7 existiert also ein Ultrafilter F über M mit $F^- \subseteq F$. Wir betrachten $\mathcal{A} := \prod_{i \in M}^F \mathcal{A}_i$ und die Produktbelegung η mit $\eta(x)(i) := \eta_i(x)$, und zeigen $\mathcal{A} \models \Gamma[\eta]$. Sei also $A \in \Gamma$. Nach dem Satz genügt es zu zeigen, daß $X_A := \{\, i \in M \mid \mathcal{A}_i \models A[\eta_i] \,\} \in F$. Dies folgt aber aus $Z_A \subseteq X_A$ und $Z_A \in F^- \subseteq F$. □

Hieraus ergibt sich unmittelbar, daß aus $\Gamma \models A$ die Existenz einer endlichen Teilmenge $\Gamma' \subseteq \Gamma$ mit $\Gamma' \models A$ folgt.

Für jede Formelmenge Γ sei $L(\Gamma)$ die Menge aller in Γ vorkommenden Funktions- und Relationssymbole. Ist $\mathcal{L}$ eine Teilsprache von $\mathcal{L}'$, $\mathcal{M}$ eine $\mathcal{L}$-Struktur und $\mathcal{M}'$ eine $\mathcal{L}'$-Struktur, so nennt man $\mathcal{M}'$ eine *Expansion* von $\mathcal{M}$ (und $\mathcal{M}$ ein *Redukt* von $\mathcal{M}'$), wenn

$$\begin{aligned} |\mathcal{M}| &= |\mathcal{M}'|, \\ f^{\mathcal{M}} &= f^{\mathcal{M}'} && \text{für alle } f \in \mathsf{Fun}_{\mathcal{L}}, \\ R^{\mathcal{M}} &= R^{\mathcal{M}'} && \text{für alle } R \in \mathsf{Rel}_{\mathcal{L}}. \end{aligned}$$

Das (eindeutig bestimmte) $\mathcal{L}$-Redukt von $\mathcal{M}'$ wird mit $\mathcal{M}' \restriction \mathcal{L}$ bezeichnet. Ist $\mathcal{M}'$ eine Expansion von $\mathcal{M}$ und η eine Belegung in $|\mathcal{M}|$, so gilt offenbar $t^{\mathcal{M}}[\eta] = t^{\mathcal{M}'}[\eta]$ für jeden $\mathcal{L}$-Term t und $\mathcal{M} \models A[\eta]$ genau dann, wenn $\mathcal{M}' \models A[\eta]$ für jede $\mathcal{L}$-Formel A. Die Gültigkeit von $\Gamma \models A$ hängt deshalb nicht von der zugrunde gelegten Sprache $\mathcal{L}$ ab, solange nur $L(\Gamma \cup \{A\}) \subseteq \mathcal{L}$ (genauer $\subseteq \mathsf{Fun}_{\mathcal{L}} \cup \mathsf{Rel}_{\mathcal{L}}$) ist.

Korollar 1.5.10. *(Allgemeiner Vollständigkeitssatz). Sei $\Gamma \cup \{A\}$ eine Menge von Formeln, wobei die zugrunde liegende Sprache überabzählbar sein kann. Dann gilt $\Gamma \vdash_c A \iff \Gamma \models A$.*

Beweis. Ein Richtung ist wieder der Korrektheitssatz. Für die umgekehrte Richtung können wir aufgrund der ersten vorangehenden Bemerkung annehmen, daß für ein endliches $\Gamma' \subseteq \Gamma$ gilt $\Gamma' \models A$. Dann gilt aber auch $\Gamma' \models A$ in einer abzählbaren Sprache (nach der zweiten vorangehenden Bemerkung). Nach dem Vollständigkeitssatz 1.5.2 für abzählbare Sprachen folgt $\Gamma' \vdash A$, also auch $\Gamma \vdash A$. $\square$

1.6 Anfänge der Modelltheorie

In diesem Abschnitt wollen wir wie in der Modelltheorie üblich auch überabzählbare Sprachen $\mathcal{L}$ zulassen. Wie wir eben gesehen haben, gelten der Vollständigkeitssatz und seine Folgerungen ebenfalls für überabzählbare Sprachen.

Wir wollen uns jetzt mit den *Gleichheitsaxiomen* befassen. Wir setzen deshalb in diesem Abschnitt stets voraus, daß die Sprache $\mathcal{L}$ ein zweistelliges Relationssymbol $=$ enthält. Die Menge $\mathsf{Eq}_{\mathcal{L}}$ der $\mathcal{L}$-Gleichheitsaxiome besteht dann aus den Allabschlüssen von

$$\begin{aligned} & x = x && \text{Reflexivität,} \\ & x = y \to y = x && \text{Symmetrie,} \\ & x = y \wedge y = z \to x = z && \text{Transitivität,} \\ & x_1 = y_1 \wedge \ldots \wedge x_n = y_n \to f x_1 \ldots x_n = f y_1 \ldots y_n \quad \text{für alle } f \in \mathsf{Fun}_{\mathcal{L}}^{(n)}, \\ & x_1 = y_1 \wedge \ldots \wedge x_n = y_n \wedge R x_1 \ldots x_n \to R y_1 \ldots y_n \quad \text{für alle } R \in \mathsf{Rel}_{\mathcal{L}}^{(n)}. \end{aligned}$$

Lemma 1.6.1. *(Gleichheitslemma).*

1. $\mathsf{Eq}_{\mathcal{L}} \vdash t = s \to r[x := t] = r[x := s]$.
2. $\mathsf{Eq}_{\mathcal{L}} \vdash t = s \to (A[x := t] \leftrightarrow A[x := s])$.

Beweis. 1. Induktion über r. 2. Induktion über A; wir beschränken uns auf den Fall $\forall y A$. Dann ist $(\forall y A)[x := r] = \forall y A[x := r]$, und nach IH haben wir $\mathsf{Eq}_{\mathcal{L}} \vdash t = s \wedge A[x := t] \to A[x := s]$. Daraus folgt die Behauptung. $\square$

Eine $\mathcal{L}$-Struktur $\mathcal{M}$ erfüllt die Gleichheitsaxiome genau dann, wenn $=^{\mathcal{M}}$ eine Kongruenzrelation ist (also eine Äquivalenzrelation, die mit den Funktionen und Relationen von $\mathcal{M}$ verträglich ist). In diesem Abschnitt setzen wir voraus, daß alle betrachteten $\mathcal{L}$-Strukturen $\mathcal{M}$ die Gleichheitsaxiome erfüllen. Das Koinzidenzlemma gilt dann auch mit $=^{\mathcal{M}}$ anstelle von $=$:

Lemma 1.6.2. *Seien η, ξ Belegungen in $|\mathcal{M}|$ mit $\mathsf{dom}(\eta) = \mathsf{dom}(\xi)$ und $\eta(x) =^{\mathcal{M}} \xi(x)$ für alle $x \in \mathsf{dom}(\eta)$. Dann gilt*

1. $t^{\mathcal{M}}[\eta] =^{\mathcal{M}} t^{\mathcal{M}}[\xi]$ *falls* $\mathsf{vars}(t) \subseteq \mathsf{dom}(\eta)$ *und*
2. $\mathcal{M} \models A[\eta] \iff \mathcal{M} \models A[\xi]$ *falls* $\mathsf{FV}(A) \subseteq \mathsf{dom}(\eta)$.

Beweis. Induktion über t bzw. A. □

Sei $\mathcal{M}/{=^{\mathcal{M}}}$ die *Quotientenstruktur*, deren Trägermenge aus den Kongruenzklassen besteht. Wir nennen eine Struktur $\mathcal{M}$ *unendlich* (*abzählbar*, *n-elementig*), wenn $\mathcal{M}/{=^{\mathcal{M}}}$ unendlich (abzählbar, n-elementig) ist.

Unter einem *Axiomensystem* Γ verstehen wir eine Menge geschlossener Formeln mit $\mathsf{Eq}_{L(\Gamma)} \subseteq \Gamma$. Ein *Modell* eines Axiomensystems Γ ist eine $\mathcal{L}$-Struktur $\mathcal{M}$ mit $L(\Gamma) \subseteq \mathcal{L}$ und $\mathcal{M} \models \Gamma$. Für Mengen Γ geschlossener Formeln schreiben wir

$$\mathsf{Mod}_{\mathcal{L}}(\Gamma) := \{\, \mathcal{M} \mid \mathcal{M} \text{ ist } \mathcal{L}\text{-Struktur und } \mathcal{M} \models \Gamma \cup \mathsf{Eq}_{\mathcal{L}} \,\}.$$

Offenbar ist Γ genau dann erfüllbar, wenn Γ ein Modell besitzt.

Satz 1.6.3. *Hat ein Axiomensystem beliebig große endliche Modelle, so hat es auch ein unendliches Modell.*

Beweis. Angenommen, Γ ist ein solches Axiomensystem. Seien $x_0, x_1, x_2, \dots$ verschiedene Variable und

$$\Gamma' := \Gamma \cup \{\, x_i \neq x_j \mid i, j \in \mathbb{N} \text{ mit } i < j \,\}.$$

Nach Annahme ist jede endliche Teilmenge von Γ' erfüllbar, also nach dem allgemeinen Kompaktheitssatz 1.5.9 auch Γ'. Wir haben also $\mathcal{M}$ und η mit $\mathcal{M} \models \Gamma'[\eta]$ und damit $\eta(x_i) \neq^{\mathcal{M}} \eta(x_j)$ für $i < j$. Also ist $\mathcal{M}$ unendlich. □

Mit $\overline{\mathcal{L}}$ bezeichnen wir die Menge aller geschlossenen $\mathcal{L}$-Formeln. Unter einer *Theorie* T verstehen wir ein gegen $\vdash$ abgeschlossenes Axiomensystem, also $\mathsf{Eq}_{L(T)} \subseteq T$ und

$$T = \{\, A \in \overline{L(T)} \mid T \vdash A \,\}.$$

Eine Theorie T heißt *vollständig*, wenn für jede Formel $A \in \overline{L(T)}$ gilt $T \vdash A$ oder $T \vdash \neg A$. Für jede $\mathcal{L}$-Struktur $\mathcal{M}$ (die die Gleichheitsaxiome erfüllt) bildet die Menge aller geschlossenen $\mathcal{L}$-Formeln A mit $\mathcal{M} \models A$ offenbar eine Theorie; sie heißt die *Theorie von* $\mathcal{M}$ und wird mit $\mathsf{Th}(\mathcal{M})$ bezeichnet. Zwei $\mathcal{L}$-Strukturen $\mathcal{M}$ und $\mathcal{M}'$ heißen *elementar äquivalent* (geschrieben $\mathcal{M} \equiv \mathcal{M}'$), wenn $\mathsf{Th}(\mathcal{M}) = \mathsf{Th}(\mathcal{M}')$ ist. Zwei $\mathcal{L}$-Strukturen $\mathcal{M}$ und $\mathcal{M}'$ heißen *isomorph* (geschrieben $\mathcal{M} \cong \mathcal{M}'$), wenn es eine Abbildung $\pi\colon |\mathcal{M}| \to |\mathcal{M}'|$ gibt, die eine Bijektion zwischen $|\mathcal{M}/{=^{\mathcal{M}}}|$ und $|\mathcal{M}'/{=^{\mathcal{M}'}}|$ induziert, also

$$\begin{aligned} &\forall a, b {\in} |\mathcal{M}|.\, a =^{\mathcal{M}} b \iff \pi(a) =^{\mathcal{M}'} \pi(b), \\ &\forall a' {\in} |\mathcal{M}'| \exists a {\in} |\mathcal{M}|\, \pi(a) =^{\mathcal{M}'} a', \end{aligned}$$

so daß für alle $a_1, \dots, a_n \in |\mathcal{M}|$ gilt

$$\begin{aligned} \pi(f^{\mathcal{M}}(a_1, \dots, a_n)) &=^{\mathcal{M}'} f^{\mathcal{M}'}(\pi(a_1), \dots, \pi(a_n)) && \text{für alle } f \in \mathsf{Fun}_{\mathcal{L}}^{(n)}, \\ R^{\mathcal{M}}(a_1, \dots, a_n) &\iff R^{\mathcal{M}'}(\pi(a_1), \dots, \pi(a_n)) && \text{für alle } R \in \mathsf{Rel}_{\mathcal{L}}^{(n)}. \end{aligned}$$

Wir stellen zunächst einige einfache Eigenschaften des Begriffs der Theorie einer Struktur $\mathcal{M}$ und der elementaren Äquivalenz zusammen.

Lemma 1.6.4. *1.* $\mathsf{Th}(\mathcal{M})$ *ist vollständig.*

2. *Ist Γ ein Axiomensystem mit $L(\Gamma) \subseteq \mathcal{L}$, so ist*

$$\{\, A \in \overline{\mathcal{L}} \mid \Gamma \vdash A \,\} = \bigcap \{\, \mathsf{Th}(\mathcal{M}) \mid \mathcal{M} \in \mathsf{Mod}_{\mathcal{L}}(\Gamma) \,\}.$$

3. $\mathcal{M} \equiv \mathcal{M}' \iff \mathcal{M} \models \mathsf{Th}(\mathcal{M}')$.
4. *Ist $\mathcal{L}$ abzählbar, so existiert zu jeder $\mathcal{L}$-Struktur $\mathcal{M}$ eine abzählbare $\mathcal{L}$-Struktur $\mathcal{M}'$ mit $\mathcal{M} \equiv \mathcal{M}'$.*

Beweis. 1. Sei $\mathcal{M}$ eine $\mathcal{L}$-Struktur und $A \in \overline{\mathcal{L}}$. Dann gilt $\mathcal{M} \models A$ oder $\mathcal{M} \models \neg A$, also $\mathsf{Th}(\mathcal{M}) \vdash A$ oder $\mathsf{Th}(\mathcal{M}) \vdash \neg A$.

2. Für alle $A \in \overline{\mathcal{L}}$ hat man

$$\begin{aligned} \Gamma \vdash A &\iff \Gamma \models A \\ &\iff \text{für alle } \mathcal{L}\text{-Strukturen } \mathcal{M} \text{ gilt } (\mathcal{M} \models \Gamma \implies \mathcal{M} \models A) \\ &\iff \text{für alle } \mathcal{L}\text{-Strukturen } \mathcal{M} \text{ gilt } (\mathcal{M} \in \mathsf{Mod}_{\mathcal{L}}(\Gamma) \implies A \in \mathsf{Th}(\mathcal{M})) \\ &\iff A \in \bigcap \{\, \mathsf{Th}(\mathcal{M}) \mid \mathcal{M} \in \mathsf{Mod}_{\mathcal{L}}(\Gamma) \,\}. \end{aligned}$$

3. $\Longrightarrow$. Gelte $\mathcal{M} \equiv \mathcal{M}'$ und $A \in \mathsf{Th}(\mathcal{M}')$. Dann folgt $\mathcal{M}' \models A$, also $\mathcal{M} \models A$.

$\Longleftarrow$. Gelte $\mathcal{M} \models \mathsf{Th}(\mathcal{M}')$. Offenbar ist dann $\mathsf{Th}(\mathcal{M}') \subseteq \mathsf{Th}(\mathcal{M})$. Zum Beweis der umgekehrten Inklusion sei $A \in \mathsf{Th}(\mathcal{M})$. Wäre $A \notin \mathsf{Th}(\mathcal{M}')$, so hätte man nach (1) auch $\neg A \in \mathsf{Th}(\mathcal{M}')$, also $\mathcal{M} \models \neg A$ im Widerspruch zu $A \in \mathsf{Th}(\mathcal{M})$.

4. Sei $\mathcal{L}$ abzählbar und $\mathcal{M}$ eine $\mathcal{L}$-Struktur. Dann ist $\mathsf{Th}(\mathcal{M})$ erfüllbar, besitzt also nach dem Satz 1.5.4 von Löwenheim und Skolem auch eine erfüllende $\mathcal{L}$-Struktur $\mathcal{M}'$ mit der abzählbaren Trägermenge $\mathsf{Ter}_{\mathcal{L}}$. Nach (3) ist $\mathcal{M} \equiv \mathcal{M}'$. □

Satz 1.6.5. *Es sei T eine Theorie und $\mathcal{L} = L(T)$. Dann sind folgende Aussagen äquivalent.*

1. *T ist vollständig.*
2. *Für jedes Modell $\mathcal{M} \in \mathsf{Mod}_{\mathcal{L}}(T)$ gilt $\mathsf{Th}(\mathcal{M}) = T$.*
3. *Je zwei Modelle $\mathcal{M}, \mathcal{M}' \in \mathsf{Mod}_{\mathcal{L}}(T)$ sind elementar äquivalent.*

Beweis. (1) $\Longrightarrow$ (2). Sei T vollständig und $\mathcal{M} \in \mathsf{Mod}_{\mathcal{L}}(T)$. Dann gilt $\mathcal{M} \models T$, also $T \subseteq \mathsf{Th}(\mathcal{M})$. Sei nun umgekehrt $A \in \mathsf{Th}(\mathcal{M})$. Dann folgt $\neg A \notin \mathsf{Th}(\mathcal{M})$, also $\neg A \notin T$ und damit $A \in T$.

(2) $\Longrightarrow$ (3) ist klar.

(3) $\Longrightarrow$ (1). Sei $A \in \overline{\mathcal{L}}$ und $T \nvdash A$. Dann existiert ein Modell $\mathcal{M}_0$ von $T \cup \{\neg A\}$. Sei nun $\mathcal{M} \in \mathsf{Mod}_{\mathcal{L}}(T)$ beliebig. Nach (3) ist $\mathcal{M} \equiv \mathcal{M}_0$, also $\mathcal{M} \models \neg A$. Damit haben wir $T \vdash \neg A$. □

Satz 1.6.6. *Ist π ein Isomorphismus von $\mathcal{M}$ auf $\mathcal{M}'$, so gilt für alle Terme t und Formeln A und für jede genügend große Belegung η in $|\mathcal{M}|$*

1. $\pi(t^{\mathcal{M}}[\eta]) =^{\mathcal{M}'} t^{\mathcal{M}'}[\pi \circ \eta]$ *und*
2. $\mathcal{M} \models A[\eta] \iff \mathcal{M}' \models A[\pi \circ \eta]$. *Insbesondere hat man also $\mathcal{M} \cong \mathcal{M}' \implies \mathcal{M} \equiv \mathcal{M}'$.*

Beweis. 1. Induktion über t. Zur Vereinfachung der Schreibweise behandeln wir nur den Fall einstelliger Funktionssymbole.

$$\begin{aligned} &\pi(x^{\mathcal{M}}[\eta]) = \pi(\eta(x)) = x^{\mathcal{M}'}[\pi \circ \eta] \\ &\pi(c^{\mathcal{M}}[\eta]) = \pi(c^{\mathcal{M}}) =^{\mathcal{M}'} c^{\mathcal{M}'} \\ &\pi(f(t)^{\mathcal{M}}[\eta]) = \pi(f^{\mathcal{M}}(t^{\mathcal{M}}[\eta])) =^{\mathcal{M}'} f^{\mathcal{M}'}(\pi(t^{\mathcal{M}}[\eta])) =^{\mathcal{M}'} f^{\mathcal{M}'}(t^{\mathcal{M}'}[\pi \circ \eta]) = f(t)^{\mathcal{M}'}[\pi \circ \eta]. \end{aligned}$$

2. Induktion über A. Wir behandeln nur den Fall einstelliger Relationssymbole und den Fall $\forall x\, A$.

$$\begin{aligned} \mathcal{M} \models R(t)[\eta] &\iff R^{\mathcal{M}}(t^{\mathcal{M}}[\eta]) \\ &\iff R^{\mathcal{M}'}(\pi(t^{\mathcal{M}}[\eta])) \\ &\iff R^{\mathcal{M}'}(t^{\mathcal{M}'}[\pi \circ \eta]) \\ &\iff \mathcal{M}' \models R(t)[\pi \circ \eta], \end{aligned}$$

$$\begin{aligned} \mathcal{M} \models \forall x\, A[\eta] &\iff \text{für alle } a \in |\mathcal{M}| \text{ gilt } \mathcal{M} \models A[\eta_x^a] \\ &\iff \text{für alle } a \in |\mathcal{M}| \text{ gilt } \mathcal{M}' \models A[\pi \circ \eta_x^a] \\ &\iff \text{für alle } a \in |\mathcal{M}| \text{ gilt } \mathcal{M}' \models A[(\pi \circ \eta)_x^{\pi(a)}] \\ &\iff \text{für alle } a' \in |\mathcal{M}'| \text{ gilt } \mathcal{M}' \models A[(\pi \circ \eta)_x^{a'}] \quad \text{nach 1.6.2} \\ &\iff \mathcal{M}' \models \forall x\, A[\pi \circ \eta] \end{aligned}$$

□

Satz 1.6.7. *Zu jeder unendlichen Struktur $\mathcal{M}$ gibt es eine elementar äquivalente Struktur $\mathcal{M}_0$, die nicht isomorph zu $\mathcal{M}$ ist.*

Beweis. OBdA sei $=^{\mathcal{M}}$ die Gleichheit auf $M := |\mathcal{M}|$. Mit $\mathcal{P}(M)$ bezeichnen wir die Potenzmenge von M. Zu jedem $\alpha \in \mathcal{P}(M)$ wählen wir eine neue Konstante c_α. In der Sprache $\mathcal{L}' := \mathcal{L} \cup \{ c_\alpha \mid \alpha \in \mathcal{P}(M) \}$ betrachten wir das Axiomensystem

$$\Gamma := \mathsf{Th}(\mathcal{M}) \cup \{ c_\alpha \neq c_\beta \mid \alpha, \beta \in \mathcal{P}(M) \text{ und } \alpha \neq \beta \} \cup \mathsf{Eq}_{\mathcal{L}'}.$$

Jede endliche Teilmenge von Γ ist erfüllbar durch eine geeignete Expansion von $\mathcal{M}$. Also ist nach dem allgemeinen Kompaktheitssatz 1.5.9 auch Γ erfüllbar, etwa durch $\mathcal{M}_0'$. Sei $\mathcal{M}_0 := M_0' {\restriction} \mathcal{L}$. OBdA sei $=^{\mathcal{M}_0}$ die Gleichheit auf $|\mathcal{M}_0|$. $\mathcal{M}_0$ ist nicht isomorph zu $\mathcal{M}$, denn andernfalls hätte man eine Injektion von $\mathcal{P}(M)$ in M und damit einen Widerspruch. □

Anmerkung. Nach Satz 1.6.7 ist es nicht möglich, eine unendliche Struktur durch ein Axiomensystem erster Stufe bis auf Isomorphie zu charakterisieren. Verläßt man jedoch die Logik erster Stufe und erlaubt auch die Quantifikation über Mengen X, so kann man etwa die folgenden *Peano-Axiome* anschreiben:

$$\begin{aligned}
&\forall n\, \mathsf{S}(n) \neq 0,\\
&\forall n \forall m. \mathsf{S}(n) = \mathsf{S}(m) \to n = m,\\
&\forall X. 0 \in X \wedge (\forall n. n \in X \to \mathsf{S}(n) \in X) \to \forall n\, n \in X.
\end{aligned}$$

Man kann zeigen, daß $(\mathbb{N}, 0, \mathsf{S})$ bis auf Isomorphie das einzige Modell der PEANO-Axiome ist. Eine Struktur, die elementar äquivalent, aber nicht isomorph zu $\mathcal{N} := (\mathbb{N}, 0, \mathsf{S})$ ist, heißt ein *Nichtstandardmodell* der natürlichen Zahlen. In Nichtstandardmodellen der natürlichen Zahlen gilt das Prinzip der vollständigen Induktion nicht für alle Mengen $X \subseteq \mathbb{N}$. Entsprechend heißt eine Struktur, die elementar äquivalent, aber nicht isomorph zu $(\mathbb{R}, 0, 1, +, \cdot, <)$ ist, ein Nichtstandardmodell der reellen Zahlen. In jedem Nichtstandardmodell der reellen Zahlen gilt das Vollständigkeitsaxiom

$$\forall X. \emptyset \neq X \text{ beschränkt} \to \exists y. y = \sup(X)$$

nicht für alle Mengen $X \subseteq \mathbb{R}$.

Satz 1.6.8. *Es gibt abzählbare Nichtstandardmodelle der natürlichen Zahlen.*

Beweis. Sei x eine Variable und

$$\Gamma := \mathsf{Th}(\mathcal{N}) \cup \{ x \neq \underline{n} \mid n \in \mathbb{N} \},$$

wobei $\underline{0} := 0$ und $\underline{n+1} := \mathsf{S}\underline{n}$. Offenbar ist jede endliche Teilmenge von Γ erfüllbar, also nach dem Kompaktheitssatz 1.5.3(2) auch Γ. Nach dem Satz 1.5.4 von LÖWENHEIM und SKOLEM haben wir dann ein höchstens abzählbares $\mathcal{M}$ und eine Belegung η mit $\mathcal{M} \models \Gamma[\eta]$. Wegen $\mathcal{M} \models \mathsf{Th}(\mathcal{N})$ ist $\mathcal{M} \equiv \mathcal{N}$ nach Satz 1.6.5; daraus folgt insbesondere, daß $\mathcal{M}$ abzählbar ist. Weiter ist $\eta(x) \neq^{\mathcal{M}} \underline{n}^{\mathcal{M}}$ für alle $n \in \mathbb{N}$, also $\mathcal{M} \not\cong \mathcal{N}$. □

Wir wollen jetzt einige Anwendungen auf aus der Mathematik bekannte Axiomensysteme diskutieren. Die Axiome der *Körpertheorie* sind (außer den Gleichheitsaxiomen) die Allabschlüsse von

$$\begin{aligned}
&x + (y + z) = (x + y) + z,\\
&0 + x = x,\\
&(-x) + x = 0,\\
&x + y = y + x,\\
&x \cdot (y \cdot z) = (x \cdot y) \cdot z,\\
&1 \cdot x = x,\\
&x \neq 0 \to x^{-1} \cdot x = 1,\\
&x \cdot y = y \cdot x,\\
&(x + y) \cdot z = (x \cdot z) + (y \cdot z),\\
&1 \neq 0.
\end{aligned}$$

Körper sind also genau die Modelle dieses Axiomensystems.

In der Theorie der *geordneten Körper* hat man zusätzlich ein zweistelliges Relationssymbol $<$ und als Axiome die Allabschlüsse von

$$
\begin{aligned}
& x \not< x, \\
& x < y \wedge y < z \to x < z, \\
& x < y \vee x = y \vee y < x, \\
& x < y \to x + z < y + z, \\
& 0 < x \wedge 0 < y \to 0 < x \cdot y.
\end{aligned}
$$

Geordnete Körper sind genau die Modelle dieses erweiterten Axiomensystems. Ein geordneter Körper heißt *archimedisch geordnet*, wenn es zu jedem Körperelement a eine natürliche Zahl n gibt, so daß a kleiner ist als das n-fache der Eins im Körper.

Satz 1.6.9. *Zu jedem archimedisch geordneten Körper gibt es einen elementar äquivalenten geordneten Körper, der nicht archimedisch geordnet ist.*

Beweis. Sei $\mathcal{K}$ ein archimedisch geordneter Körper, x eine Variable und

$$\Gamma := \mathsf{Th}(\mathcal{K}) \cup \{\, \underline{n} < x \mid n \in \mathbb{N} \,\}.$$

Offenbar ist jede endliche Teilmenge von Γ erfüllbar, also nach dem allgemeinen Kompaktheitssatz 1.5.9 auch Γ. Wir haben also $\mathcal{M}$ und η mit $\mathcal{M} \models \Gamma[\eta]$. Wegen $\mathcal{M} \models \mathsf{Th}(\mathcal{K})$ ist $\mathcal{M} \equiv \mathcal{K}$ und damit $\mathcal{M}$ ein geordneter Körper. Weiter ist $1^{\mathcal{M}} \cdot n <^{\mathcal{M}} \eta(x)$ für alle $n \in \mathbb{N}$, also $\mathcal{M}$ nicht archimedisch geordnet. □

Eine Klasse $\mathcal{S}$ von $\mathcal{L}$-Strukturen heißt (*endlich*) *axiomatisierbar*, wenn es ein (endliches) Axiomensystem Γ gibt mit $\mathcal{S} = \mathsf{Mod}_{\mathcal{L}}(\Gamma)$. Offenbar ist $\mathcal{S}$ endlich axiomatisierbar genau dann, wenn $\mathcal{S} = \mathsf{Mod}_{\mathcal{L}}(\{A\})$ für ein A. Gibt es zu jedem $\mathcal{M} \in \mathcal{S}$ ein elementar äquivalentes $\mathcal{M}' \notin \mathcal{S}$, so kann $\mathcal{S}$ offensichtlich nicht axiomatisierbar sein. Aus Satz 1.6.9 folgt also, daß die Klasse der archimedisch geordneten Körper nicht axiomatisierbar ist. Ebenso folgt, daß die Klasse der geordneten Körper, die nicht archimedisch geordnet sind, nicht axiomatisierbar ist.

Lemma 1.6.10. *Sei $\mathcal{S}$ eine Klasse von $\mathcal{L}$-Strukturen und Γ ein Axiomensystem.*

1. *$\mathcal{S}$ ist endlich axiomatisierbar genau dann, wenn $\mathcal{S}$ und das Komplement von $\mathcal{S}$ axiomatisierbar sind.*
2. *Ist $\mathsf{Mod}_{\mathcal{L}}(\Gamma)$ endlich axiomatisierbar, so gibt es ein endliches $\Gamma_0 \subseteq \Gamma$ mit $\mathsf{Mod}_{\mathcal{L}}(\Gamma_0) = \mathsf{Mod}_{\mathcal{L}}(\Gamma)$.*

Beweis. 1. Mit $1 - \mathcal{S}$ bezeichnen wir das Komplement von $\mathcal{S}$. $\Longrightarrow$. Sei $\mathcal{S} = \mathsf{Mod}_{\mathcal{L}}(\{A\})$. Dann gilt $\mathcal{M} \in 1 - \mathcal{S} \iff \mathcal{M} \models \neg A$, also $1 - \mathcal{S} = \mathsf{Mod}_{\mathcal{L}}(\{\neg A\})$. $\Longleftarrow$. Sei $\mathcal{S} = \mathsf{Mod}_{\mathcal{L}}(\Gamma_1)$ und $1 - \mathcal{S} = \mathsf{Mod}_{\mathcal{L}}(\Gamma_2)$. Dann ist $\Gamma_1 \cup \Gamma_2$ nicht erfüllbar, es gibt also ein endliches $\Gamma \subseteq \Gamma_1$ so daß $\Gamma \cup \Gamma_2$ nicht erfüllbar ist. Man erhält

$$\mathcal{M} \in \mathcal{S} \implies \mathcal{M} \models \Gamma \implies \mathcal{M} \not\models \Gamma_2 \implies \mathcal{M} \notin 1 - \mathcal{S} \implies \mathcal{M} \in \mathcal{S},$$

also $\mathcal{S} = \mathsf{Mod}_{\mathcal{L}}(\Gamma)$.

2. Sei $\mathsf{Mod}_{\mathcal{L}}(\Gamma) = \mathsf{Mod}_{\mathcal{L}}(\{A\})$. Dann gilt $\Gamma \models A$, also auch $\Gamma_0 \models A$ für ein endliches $\Gamma_0 \subseteq \Gamma$. Man erhält

$$\mathcal{M} \models \Gamma \implies \mathcal{M} \models \Gamma_0 \implies \mathcal{M} \models A \implies \mathcal{M} \models \Gamma,$$

also $\mathsf{Mod}_{\mathcal{L}}(\Gamma_0) = \mathsf{Mod}_{\mathcal{L}}(\Gamma)$. □

Zum Abschluß dieses Abschnitts behandeln wir als Beispiel einer vollständigen Theorie noch die Theorie DO der dichten linearen Ordnungen ohne Endpunkte. Die Axiome der Theorie DO sind (außer den Gleichheitsaxiomen) die Allabschlüsse von

$$
\begin{aligned}
& x \not< x, \\
& x < y \wedge y < z \to x < z, \\
& x < y \vee x = y \vee y < x, \\
& x < y \to \exists z. x < z \wedge z < y, \\
& \exists y\, x < y, \\
& \exists y\, y < x.
\end{aligned}
$$

Lemma 1.6.11. *Jedes abzählbare Modell von* DO *ist isomorph zur Struktur* $(\mathbb{Q}, <)$ *der rationalen Zahlen.*

Beweis. Sei $\mathcal{M} = (M, \prec)$ ein abzählbares Modell von DO; wir können annehmen, daß $=^{\mathcal{M}}$ die Gleichheit auf M ist. Sei $M = \{ b_n \mid n \in \mathbb{N} \}$ und $\mathbb{Q} = \{ a_n \mid n \in \mathbb{N} \}$, wobei wir $a_n \neq a_m$ und $b_n \neq b_m$ für $n < m$ annehmen können. Wir definieren rekursiv Funktionen $f_n \subseteq \mathbb{Q} \times M$ wie folgt. Sei $f_0 := \{(a_0, b_0)\}$. Angenommen, wir haben f_n schon konstruiert.

Fall $n + 1 = 2m$. Sei j minimal so daß $b_j \notin \mathsf{ran}(f_n)$. Wähle $a_i \notin \mathsf{dom}(f_n)$ so daß für alle $a \in \mathsf{dom}(f_n)$ gilt $a_i < a \leftrightarrow b_j < f_n(a)$; ein solches a_i existiert, da $\mathcal{M}$ und $(\mathbb{Q}, <)$ Modelle von DO sind. Setze $f_{n+1} := f_n \cup \{(a_i, b_j)\}$.

Den *Fall* $n+1 = 2m+1$ behandelt man ähnlich. Sei i minimal so daß $a_i \notin \mathsf{dom}(f_n)$. Wähle $b_j \notin \mathsf{ran}(f_n)$ so daß für alle $a \in \mathsf{dom}(f_n)$ gilt $a_i < a \leftrightarrow b_j < f_n(a)$; ein solches b_j existiert, da $\mathcal{M}$ und $(\mathbb{Q}, <)$ Modelle von DO sind. Setze $f_{n+1} := f_n \cup \{(a_i, b_j)\}$.

Nach Konstruktion ist $\{b_0, \dots, b_m\} \subseteq \mathsf{ran}(f_{2m})$ und $\{a_0, \dots, a_{m+1}\} \subseteq \mathsf{dom}(f_{2m+1})$ und schließlich $f := \bigcup_n f_n$ ein Isomorphismus von $(\mathbb{Q}, <)$ auf $\mathcal{M}$. □

Satz 1.6.12. *Die Theorie* DO *ist vollständig, und es gilt* $\mathsf{DO} = \mathsf{Th}(\mathbb{Q}, <)$.

Beweis. Da offenbar $(\mathbb{Q}, <)$ Modell von DO ist, genügt es nach Satz 1.6.5 zu zeigen, daß für jedes Modell $\mathcal{M}$ von DO gilt $\mathcal{M} \equiv (\mathbb{Q}, <)$. Sei also $\mathcal{M}$ Modell von DO. Nach Lemma 1.6.4(4) existiert ein abzählbares $\mathcal{M}'$ mit $\mathcal{M} \equiv \mathcal{M}'$. Nach dem vorangehenden Satz ist $\mathcal{M}' \cong (\mathbb{Q}, <)$, also $\mathcal{M} \equiv \mathcal{M}' \equiv (\mathbb{Q}, <)$. □

Ein weiteres Beispiel einer vollständigen Theorie ist die Theorie der algebraisch abgeschlossenen Körper. Hierfür und für viele weitere Gegenstände der Modelltheorie sei verwiesen auf die Literatur; ein gutes und umfassendes Buch über Modelltheorie ist das von Chang und Keisler [6].

1.7 Anmerkungen

Der Gödelsche Vollständigkeitssatz wurde zuerst von Gödel [12] für abzählbare Sprachen, und im allgemeinen Fall von Malzew [20] bewiesen. Der Satz von Löwenheim und Skolem wurde bereits vor dem Vollständigkeitssatz bewiesen von Löwenheim [19] und Skolem [35].

Beth-Strukturen für die intuitionistische Logik wurden von Beth in [1] eingeführt; die dort geführten (klassischen und intuitionistischen) Vollständigkeitsbeweise waren jedoch korrekturbedürftig. Beth hat seine Arbeit später in [2] revidiert. Eine seiner Absichten war es, *intuitionistische* Vollständigkeitsbeweise für die intuitionistische Logik zu geben. Dieser Aspekt ist besonders von Kreisel betont worden; er wird in der vorliegenden Darstellung ignoriert. Eine gute Übersicht über intuitionistische Vollständigkeitsbeweise findet man in Troelstras Buch [41]. Die Kripke-Semantik für die intuitionistische Logik wurde entwickelt in [17]; er gab dort einen (klassischen) Vollständigkeitsbeweis im Gödelschen Stil. Die Konstruktion einer Beth-Struktur aus einer Kripke-Struktur geht zurück auf Kripke [17]; die vorliegende Darstellung stützt sich auf das Buch [44] von Troelstra und van Dalen.

Der Beweis in Abschnitt 1.5 des Vollständigkeitssatzes für die klassische Logik aus dem Vollständigkeitssatz für die intuitionistische Logik stammt von Berger. Die Idee zum Begriff eines Ultraprodukts geht auf Skolem [36] zurück; er verwendete eine eingeschränkte Form zur Konstruktion eines Nichtstandardmodell der vollständigen Arithmetik. Der allgemeine Begriff eines Ultraprodukts wurde von Łoś [18] eingeführt; in dieser Arbeit hat er auch den Fundamentalsatz formuliert.

2. Beweistheorie

Das erste Ziel dieses Kapitels ist die Einführung einer kompakten und bequemen Bezeichnungsweise für Herleitungen: wir ersetzen die bisherige baumartige Bezeichnungsweise durch eine lineare. Um auch der Möglichkeit der Einführung und Beseitigung von Annahmen Rechnung zu tragen, verwenden wir freie (Annahme)variablen und einen Bindungsoperator für sie, der traditionsgemäß mit λ bezeichnet wird. Dies führt unmittelbar auf getypte λ-Kalküle.

Ferner zeigen wir in diesem Kapitel, daß sich jede Herleitung durch geeignete Konversionsschritte in eine Normalform bringen läßt. Eine normale Herleitung macht keine "Umwege", d.h. es kommt niemals vor, daß eine Beseitigung unmittelbar auf eine Einführung folgt. Herleitungen in Normalform haben viele angenehme Eigenschaften, und können zum Beweis vieler Resultate benutzt werden. Weiterhin zeigen wir, daß die Normalform eindeutig bestimmt ist.

Dies wird zunächst für das $\to$-Fragment der Aussagenlogik durchgeführt, da fast alle wesentlichen Aspekte hier schon zum Vorschein kommen. Die Erweiterung auf die volle Sprache wird anschließend diskutiert.

Als Anwendung beweisen wir die Teilformeleigenschaft und den HERBRANDschen Satz. Schließlich zeigen wir, daß die Forderung nach normalen Herleitungen manchmal unrealistisch sein kann. Wir geben Beispiele von Formeln C_k, die mit nicht normalen Herleitungen der Länge linear in k bewiesen werden können, für die aber jede normale Herleitung eine nicht mehr elementar rekursive Anzahl von Knoten erfordert.

2.1 $\boldsymbol{\lambda}$-Kalkül

Wir beginnen mit dem Fragment der Minimallogik, in dem nur Formeln zugelassen sind, die aus Aussagensymbolen mittels Implikation $\to$ aufgebaut sind; man nennt es das $\to$-Fragment der minimalen Aussagenlogik. Herleitungen sind dann mit den Einführungs- und Beseitigungsregeln für $\to$ gebildet. Sie haben eine besonders einfache Struktur, da man auf Variablenbedingungen wegen des Fehlens der Quantoren nicht zu achten braucht. Annahmenvariablen können jedoch nach wie vor abgebunden werden.

Es bietet sich nun an, eine kompakte Termschreibweise für Herleitungen zu verwenden. Die Herleitungsterme müssen dann aber – im Gegensatz zu den Termen der Logik erster Stufe – einen Bindungsoperator enthalten; er wird mit λ bezeichnet. Formeln erscheinen jetzt als (obere) Indizes solcher Herleitungsterme. Man nennt deshalb die Formeln des $\to$-Fragments der Aussagenlogik auch (einfache) Typen und die zugehörigen Herleitungsterme λ-Terme mit Typen.

Die zentrale Idee für unsere Termbezeichnung für Herleitungen ist also, *Formeln als Typen* von Herleitungstermen zu betrachten. Dieser Ansatz ist als CURRY-HOWARD *Korrespondenz* in der Literatur bekannt.

Formeln des $\to$-Fragments der minimalen Aussagenlogik heißen dann (einfache) *Typen*; als Mitteilungszeichen für Typen verwenden wir ρ, σ, τ. Spezielle Typen sind dann die Aussagensymbole, die wir hier *Grundtypen* nennen; als Mitteilungszeichen für Grundtypen verwenden wir μ. Also

– Jeder Grundtyp μ ist ein Typ.
– Sind ρ und σ Typen, so auch $\rho \to \sigma$.

Man beachte, daß sich jeder Typ ρ eindeutig schreiben läßt in der Form $\rho = \rho_1 \to \rho_2 \to \dots \to \rho_n \to \mu$; die ρ_i nennt man die*Argumenttypen* von ρ.

Definition 2.1.1. (Induktive Definition der λ-*Terme* mit Typen).

– Jede Variable x^ρ ist ein λ-Term.
– Ist M^σ ein λ-Term, so auch $(\lambda x^\rho M^\sigma)^{\rho\to\sigma}$ (*Abstraktion*).
– Sind $M^{\rho\to\sigma}$ und N^ρ λ-Terme, so auch $(M^{\rho\to\sigma}N^\rho)^\sigma$ (*Anwendung*).

λ-Terme nennen wir im folgenden einfach Terme; auch lassen wir die Typenindizes meistens weg. Mit Λ (Λ_ρ) bezeichnen wir die Menge aller Terme (des Typs ρ). Terme, die keine Abstraktionen sind, heißen *neutral*. Bei Anwendungstermen vereinbaren wir wie üblich Linksklammerung, also MNK meint $(MN)K$, nicht $M(NK)$. Ferner soll die Abstraktion stärker binden als die Anwendung, also λxMN meint $(\lambda xM)N$, nicht $\lambda x(MN)$. Wir verwenden wieder eine Punktnotation, wenn der Wirkungsbereich eines λ-Operators so groß wie durch die Klammerung möglich sein soll, also z.B. $(\lambda x.MN)K$ meint $(\lambda x(MN))K$. Die Menge $\mathsf{FV}(M)$ der freien Variablen eines λ-Terms definieren wir rekursiv durch

$$\begin{aligned}\mathsf{FV}(x) &:= \{x\}\\ \mathsf{FV}(\lambda xM) &:= \mathsf{FV}(M)\setminus\{x\}\\ \mathsf{FV}(MN) &:= \mathsf{FV}(M)\cup\mathsf{FV}(N)\end{aligned}$$

Hinzufügen des $\wedge$-Konstruktors für Typen und entsprechend von Paarbildung und Projektionen für Terme führt auf den getypten λ-Kalkül mit Paarbildung, der für Anwendungen meist benötigt wird. Diese Erweiterung ist jedoch recht harmlos. Wir schreiben anstelle von $\rho\wedge\sigma$ hier $\rho\times\sigma$. Die induktive Definition der λ-Terme mit Typen wird erweitert durch die Klauseln

– Sind M^ρ und N^σ λ-Terme, so auch $\langle M^\rho, N^\sigma\rangle^{\rho\times\sigma}$ (*Paarbildung*).
– Ist $M^{\rho\times\sigma}$ ein λ-Term, so sind es auch $\pi_0(M)^\rho$ und $\pi_1(M)^\sigma$ (*Projektionen*).

Manchmal ist es nützlich, $M0$ für $\pi_0(M)$ und $M1$ für $\pi_1(M)$ zu schreiben.

Wir können in ähnlicher Weise unsere Termdarstellung auf die volle $\wedge\!\to\!\forall$ Sprache erweitern. Dies bringt keine wesentlichen Schwierigleiten mit sich, nur ist einige Sorgfalt bei den Bezeichnungen erforderlich. Der Hauptunterschied in der Bezeichnungweise besteht darin, daß jetzt Formeln als Typensymbole verwendet werden, anstelle von $\rho, \sigma, \dots$ vorher.

Als Grundbausteine von Herleitungstermen oder kurz Termen verwenden wir mit Formeln versehene Annahmenvariable u^A; u ist wieder eines der Symbole $\Box_0, \Box_1, \dots$. Wir verlangen wie bisher, daß verschiedene Annahmeformeln stets verschiedene Marken u haben. Als Mitteilungszeichen für Herleitungsterme verwenden wir M, N, K, L und für Annahmevariablen u, v, w. Wir schreiben meist $u\colon A$ anstelle von u^A und $M\colon A$ anstelle von M^A.

Wir geben eine induktive Definition der Herleitungsterme in der Form der Tabelle 2.1, wobei wir zum besseren Verständnis links die entsprechenden Herleitungen aufschreiben.

Zur Formulierung der Variablenbedingungen benötigen wir die Menge $\mathsf{FA}(M)$ der im Herleitungsterm M *freien Annahmen*. Sie ist rekursiv definiert durch

$$\begin{aligned}\mathsf{FA}(u^A) &:= \{u^A\},\\ \mathsf{FA}(\langle M^A, N^B\rangle^{A\wedge B}) &:= \mathsf{FA}(M^A)\cup\mathsf{FA}(N^B),\\ \mathsf{FA}(\pi_0(M^{A\wedge B})^A) &:= \mathsf{FA}(M^{A\wedge B}),\\ \mathsf{FA}(\pi_1(M^{A\wedge B})^B) &:= \mathsf{FA}(M^{A\wedge B}),\\ \mathsf{FA}((\lambda u^A M^B)^{A\to B}) &:= \mathsf{FA}(M^B)\setminus\{u^A\},\\ \mathsf{FA}((M^{A\to B}N^A)^B) &:= \mathsf{FA}(M^{A\to B})\cup\mathsf{FA}(N^A),\\ \mathsf{FA}((\lambda x M^A)^{\forall xA}) &:= \mathsf{FA}(M^A),\\ \mathsf{FA}((M^{\forall xA}t)^{A[x:=t]}) &:= \mathsf{FA}(M^{\forall xA}).\end{aligned}$$

Die Variablenbedingung ist dann $x\notin\mathsf{FV}(A)$ für alle $u^A\in\mathsf{FA}(M)$.

Man beachte, daß in Herleitungstermen Formeln oft überflüssig sind. Wenn man $(\lambda uM)^{A\to B}$ schreibt, so ist klar, daß u die Formel A und M die Formel B haben müssen. Ebenso muß in $(M^{A\to B}N)$ der Gesamtterm die Formel B und N die Formel A haben. Wir werden im folgenden häufig solche redundanten Formeln weglassen. Auf diese Weise erhält man mit den Herleitungstermen eine sehr kompakte Schreibweise für Herleitungen.

Als Beispiele für Herleitungsterme schreiben wir die Herleitungen (1.1)-(1.4) als Herleitungsterme auf.

Herleitung	Term
$u\colon A$	u^A
$\dfrac{\begin{array}{c}\mid M\\ A\end{array} \qquad \begin{array}{c}\mid N\\ B\end{array}}{A \wedge B}\,\wedge^+$	$\langle M^A, N^B \rangle^{A\wedge B}$
$\dfrac{\begin{array}{c}\mid M\\ A \wedge B\end{array}}{A}\,\wedge_0^- \qquad \dfrac{\begin{array}{c}\mid M\\ A \wedge B\end{array}}{B}\,\wedge_1^-$	$\pi_0(M^{A\wedge B})^A \qquad \pi_1(M^{A\wedge B})^B$
$\dfrac{\begin{array}{c}[u\colon A]\\ \mid M\\ B\end{array}}{A \to B}\,\to^+ u$	$(\lambda u^A M^B)^{A\to B}$
$\dfrac{\begin{array}{c}\mid M\\ A \to B\end{array} \qquad \begin{array}{c}\mid N\\ A\end{array}}{B}\,\to^-$	$(M^{A\to B} N^A)^B$
$\dfrac{\begin{array}{c}\mid M\\ A\end{array}}{\forall x A}\,\forall^+$ (mit Var.Bed.)	$(\lambda x M^A)^{\forall x A}$ (mit Var.Bed.)
$\dfrac{\begin{array}{c}\mid M\\ \forall x A\end{array} \qquad t}{A[x := t]}\,\forall^-$	$(M^{\forall x A} t)^{A[x:=t]}$

Tab. 2.1. Curry-Howard Korrespondenz zwischen natürlichen Herleitungen und dem einfach getypten λ-Kalkül

(1.1) Mit $D := (A \wedge B \to C) \to (A \to (B \to C))$ haben wir

$$\left\{\lambda u^{A\wedge B\to C}\left(\lambda v^A\left[\lambda w^B\left(u^{A\wedge B\to C}\langle v^A, w^B\rangle^{A\wedge B}\right)^C\right]^{B\to C}\right)^{A\to B\to C}\right\}^D$$

Es ist klar, daß man hier viele redundante Formeln weglassen kann. Mit Annahmenvariablen u, v, w der Typen $u\colon A \wedge B \to C$, $v\colon A$ und $w\colon B$ schreiben wir den Herleitungsterm kurz als $\lambda u\lambda v\lambda w(u\langle vw\rangle)$.

(1.2) Wir verwenden Annahmenvariablen $u\colon A \to (B \to C)$ und $v\colon A \wedge B$. Der Herleitungsterm ist dann $\lambda u\lambda v[u(\pi_0 v)](\pi_1 v)$.

(1.3) Mit $u\colon \forall x(A \to B)$ und $v\colon A$ erhalten wir $\lambda u\lambda v\lambda x[(ux)v]$.

(1.4) Mit $u\colon A \to \forall xB$ und $v\colon A$ erhalten wir $\lambda u\lambda x\lambda v[(uv)x]$.

Neutral heißen jetzt Terme, die weder Abstraktionen (also von der Form λuM oder λxM) noch Paare sind. Es ist bequem, anstelle von $\pi_i(M)$ auch Mi mit $i = 0, 1$ zu schreiben: dann können wir jede Beseitigungsregel in der Form einer Anwendung mitteilen. Die Schreibweise $M\boldsymbol{N}$ bedeutet jetzt also einen (linksgeklammerten) verallgemeinerten Anwendungsterm, in dem einige N_i auch $0, 1$ (für die Projektionen π_0, π_1 bei $\wedge_0^-$, $\wedge_1^-$) oder Objektterme t (bei $\forall^-$) sein können.

Ein Herleitungsterm M^A heißt *geschlossen*, wenn $\mathsf{FA}(M^A) = \emptyset$. Wir schreiben

$$M^B[u_1^{A_1}, \dots, u_n^{A_n}]$$

um mitzuteilen, daß die freien Annahmen von M^B in der Liste $u_1^{A_1}, \dots, u_n^{A_n}$ enthalten sind.

Ferner benötigt man oft die Menge $\mathsf{FV}(M)$ der *freien* (Objekt-) *Variablen* in einem Herleitungsterm M; sie ist rekursiv definiert durch

$$\begin{aligned}
\mathsf{FV}(u^A) &:= \mathsf{FV}(A),\\
\mathsf{FV}(\langle M^A, N^B\rangle) &:= \mathsf{FV}(M^A) \cup \mathsf{FV}(N^B),\\
\mathsf{FV}(\pi_i(M^{A\wedge B})) &:= \mathsf{FV}(M^{A\wedge B}),\\
\mathsf{FV}(\lambda u^B M^A) &:= \mathsf{FV}(M^A),\\
\mathsf{FV}(M^{A\to B}N^A) &:= \mathsf{FV}(M^{A\to B}) \cup \mathsf{FV}(N^A),\\
\mathsf{FV}(\lambda x M^A) &:= \mathsf{FV}(M^A) \setminus \{x\},\\
\mathsf{FV}(M^{\forall xA}t) &:= \mathsf{FV}(M^{\forall xA}) \cup \mathsf{vars}(t).
\end{aligned}$$

Als Beispiele betrachten wir die folgende Herleitungsterme.

(1) $M_1 = \lambda v^P u^{Qx}$.

$$\dfrac{u\colon Qx}{P \to Qx}\,\to^+ v$$

Dann ist $\mathsf{FA}(M_1) = \{u^{Qx}\}$ und $\mathsf{FV}(M_1) = \{x\}$.

(2) $M_2 = \lambda u^{Qx}\lambda v^P u^{Qx}$.

$$\dfrac{\dfrac{u\colon Qx}{P \to Qx}\,\to^+ v}{Qx \to (P \to Qx)}\,\to^+ u$$

Dann ist $\mathsf{FA}(M_2) = \emptyset$ und $\mathsf{FV}(M_2) = \{x\}$.

(3) $M_3 = \lambda x\lambda u^{Qx}\lambda v^P u^{Qx}$.

$$\dfrac{\dfrac{\dfrac{u\colon Qx}{P \to Qx}\,\to^+ v}{Qx \to (P \to Qx)}\,\to^+ u}{\forall x.Qx \to (P \to Qx)}\,\forall^+$$

Dann ist $\mathsf{FA}(M_3) = \emptyset$ und $\mathsf{FV}(M_3) = \emptyset$.

Für Herleitungsterme haben wir zwei Arten von Substitutionen: Man kann einen Herleitungsterm M^A für eine Annahmenvariable u^A substituieren, und man kann auch einen Objektterm t für eine Objektvariable x substituieren.

Die Axiome für den starken Existenzquantor $\exists^*$ werden durch entsprechende Konstanten bezeichnet:

$$\exists^{*+}_{x,A} \colon \forall x.A \to \exists^* xA$$
$$\exists^{*-}_{x,A,B} \colon \exists^* xA \to (\forall x.A \to B) \to B$$

mit der üblichen Voraussetzung $x \notin \mathsf{FV}(B)$.

2.2 Normalisierung

Wir behandeln zunächst das $\to$-Fragment der minimalen Aussagenlogik; der Beweis wird sich dann leicht auf die volle Sprache übertragen lassen.

Zunächst definieren wir eine *Konversionsrelation* $\mapsto_\rho$ zwischen Termen vom Typ ρ durch

$$\begin{aligned} (\lambda x M)N &\mapsto M[x := N], && \beta\text{-Konversion} \\ \lambda x.Mx &\mapsto M \quad \text{falls } x \notin \mathsf{FV}(M) \text{ und } M \text{ neutral ist.} && \eta\text{-Konversion} \end{aligned}$$

Die *Einschritt-Reduktionsrelation* $\to^1$ kann man jetzt wie folgt definieren. $M \to^1 N$ gilt, wenn N aus M entsteht durch Ersetzen eines Teilterms M' in M durch N', wobei $M' \mapsto N'$. Die Reduktionsrelationen $\to^+$ and $\to^*$ sind der transitive und der reflexiv-transitive Abschluß von $\to^1$. Für $\boldsymbol{M} = M_1, \dots, M_n$ schreiben wir $\boldsymbol{M} \to^1 \boldsymbol{M'}$ falls $M_i \to^1 M_i'$ für ein $i \in \{1, \dots, n\}$ und $M_j = M_j'$ für alle $i \neq j \in \{1, \dots, n\}$.

Folgende Eigenschaft dieser Reduktionsrelation macht man sich leicht klar.

Lemma 2.2.1. *1. Wenn $x\boldsymbol{M} \to^1 N$, so ist $N = x\boldsymbol{M'}$ mit $\boldsymbol{M} \to^1 \boldsymbol{M'}$.*
2. Substitution und Reduktion sind miteinander verträglich, d.h. wenn $M \to^* M'$ und $N \to^* N'$, so gilt auch $M[x := N] \to^* M'[x := N']$. □

Definition 2.2.2. Ein Term M heißt *normal* (oder in *Normalform*), wenn es keinen Term N gibt mit $M \to^1 N$.

Definition 2.2.3. Die Menge sn der *stark normalisierenden* Terme wird induktiv definiert durch

$$(\forall N.M \to^1 N \implies N \in \mathsf{sn}) \implies M \in \mathsf{sn} \tag{2.1}$$

Man beachte, daß mit M offenbar auch jeder Teilterm von M stark normalisierend ist.

Zum Beweis, daß jeder Term M stark normalisierend ist, verwenden wir eine auf W.W. Tait [39] zurückgehende Methode, die auf der Einführung von sogenannten *starken Berechenbarkeitsprädikaten* sb^ρ beruht. Sie werden durch Induktion über den Typ ρ wie folgt definiert.

Definition 2.2.4.

$$M \in \mathsf{sb}^\mu \quad :\iff \quad \forall N.M \to^1 N \implies N \in \mathsf{sb} \tag{2.2}$$
$$M \in \mathsf{sb}^{\rho\to\sigma} \quad :\iff \quad (\forall N{\in}\mathsf{sb}^\rho)\, MN \in \mathsf{sb}^\sigma. \tag{2.3}$$

Lemma 2.2.5. *Ist $M \in \mathsf{sb}^\rho$ und $M \to^1 M'$, so ist auch $M' \in \mathsf{sb}$.*

Beweis. Induktion über ρ. *Fall* μ. Nach (2.2). *Fall* $\rho \to \sigma$. Gelte $M \in \mathsf{sb}^{\rho\to\sigma}$ und $M \to^1 M'$; zu zeigen ist $M' \in \mathsf{sb}$. Sei also $N \in \mathsf{sb}^\rho$; zu zeigen ist $M'N \in \mathsf{sb}^\sigma$. Dies folgt aber aus $MN \to^1 M'N$ und $MN \in \mathsf{sb}^\sigma$ nach Induktionshypothese (IH) über σ. □

Lemma 2.2.6. $(\forall \boldsymbol{M}{\in}\mathsf{sn}).\boldsymbol{M} \in \mathsf{sb} \implies (x\boldsymbol{M})^\mu \in \mathsf{sb}$.

Beweis. Induktion über $\boldsymbol{M} \in \mathsf{sn}$. Gelte $\boldsymbol{M} \in \mathsf{sn}$ und $\boldsymbol{M} \in \mathsf{sb}$; zu zeigen ist $(x\boldsymbol{M})^\mu \in \mathsf{sb}$. Sei also $x\boldsymbol{M} \to^1 N$; zu zeigen ist $N \in \mathsf{sb}$. Nach Lemma 2.2.1 muß N von der Gestalt $x\boldsymbol{M'}$ sein mit $\boldsymbol{M} \to^1 \boldsymbol{M'}$. Nach Lemma 2.2.5 gilt aber $\boldsymbol{M'} \in \mathsf{sb}$, also $x\boldsymbol{M'} \in \mathsf{sb}$ nach IH für $\boldsymbol{M'}$. □

Lemma 2.2.7.

$$\mathsf{sb}^\rho \subseteq \mathsf{sn}, \tag{2.4}$$

$$x \in \mathsf{sb}^\rho. \tag{2.5}$$

Beweis. Durch simultane Induktion über ρ. *Fall* μ. (2.4). Wir zeigen $M \in \mathsf{sb}^\mu \implies M \in \mathsf{sn}$ durch (Neben-) Induktion über $M \in \mathsf{sb}^\mu$. Gelte also $M \in \mathsf{sb}^\mu$; zu zeigen ist $M \in \mathsf{sn}$. Für jedes N mit $M \to^1 N$ haben wir $N \in \mathsf{sb}$ nach (2.2), also $N \in \mathsf{sn}$ nach der Nebeninduktionsvoraussetzung (NIH). (2.5). $x \in \mathsf{sb}^\mu$ gilt, da es kein N mit $x \to^1 N$ gibt.

Fall $\rho \to \sigma$. (2.4). Gelte $M \in \mathsf{sb}^{\rho\to\sigma}$; zu zeigen ist $M \in \mathsf{sn}$. Nach IH(2.5) für ρ haben wir $x \in \mathsf{sb}^\rho$, also $Mx \in \mathsf{sb}^\sigma$, also $Mx \in \mathsf{sn}$ nach IH(2.4) für σ. Nun impliziert $Mx \in \mathsf{sn}$ offenbar $M \in \mathsf{sn}$, wir man durch Induktion über $Mx \in \mathsf{sn}$ leicht sieht. (2.5). Sei $\boldsymbol{M} \in \mathsf{sb}^{\boldsymbol{\rho}}$ mit $\rho_1 = \rho$; zu zeigen ist $x\boldsymbol{M} \in \mathsf{sb}^\nu$. Dies folgt aber aus Lemma 2.2.6, unter Verwendung von IH(2.4) für $\boldsymbol{\rho}$. □

Lemma 2.2.8. $(\forall M, N, \boldsymbol{N} {\in} \mathsf{sn}).M[x := N]\boldsymbol{N} \in \mathsf{sb}^\mu \implies (\lambda x M)N\boldsymbol{N} \in \mathsf{sb}^\mu$.

Beweis. Durch Induktion über $M, N, \boldsymbol{N} \in \mathsf{sn}$. Seien $M, N, \boldsymbol{N} \in \mathsf{sn}$ und gelte $M[x := N]\boldsymbol{N} \in \mathsf{sb}$; zu zeigen ist $(\lambda x M)N\boldsymbol{N} \in \mathsf{sb}$. Gelte $(\lambda x M)N\boldsymbol{N} \to^1 K$; zu zeigen ist $K \in \mathsf{sb}$. *Fall* $K = (\lambda x M')N'\boldsymbol{N'}$ mit $M, N, \boldsymbol{N} \to^1 M', N', \boldsymbol{N'}$. Dann gilt $M[x := N]\boldsymbol{N} \to^* M'[x := N']\boldsymbol{N'}$ nach Lemma 2.2.1, also $M'[x := N']\boldsymbol{N'} \in \mathsf{sb}$ nach Lemma 2.2.5 bzw. der Annahme, also $(\lambda x M')N'\boldsymbol{N'} \in \mathsf{sb}$ nach IH. *Fall* $K = M[x := N]\boldsymbol{N}$. Dann ist $K \in \mathsf{sb}$ nach Annahme. □

Korollar 2.2.9. $(\forall M, N, \boldsymbol{N} {\in} \mathsf{sn}).M[x := N]\boldsymbol{N} \in \mathsf{sb}^\rho \implies (\lambda x M)N\boldsymbol{N} \in \mathsf{sb}^\rho$.

Beweis. Nach Induktion über ρ, mit Hilfe von (2.4). □

Definition 2.2.10. Ein Term M heißt *stark berechenbar unter Substitution*, wenn für alle $\boldsymbol{N} \in \mathsf{sb}$ der Term $M[\boldsymbol{x} := \boldsymbol{N}] \in \mathsf{sb}$ ist.

Satz 2.2.11. *Jeder Term ist stark berechenbar unter Substitution.*

Beweis. Induktion über den Term M. *Fall* x. Klar. *Fall* MN. Nach IH sind $M[\boldsymbol{x} := \boldsymbol{N}] \in \mathsf{sb}$ und $N[\boldsymbol{x} := \boldsymbol{N}] \in \mathsf{sb}$, also auch $(MN)[\boldsymbol{x} := \boldsymbol{N}] = M[\boldsymbol{x} := \boldsymbol{N}]N[\boldsymbol{x} := \boldsymbol{N}] \in \mathsf{sb}$. *Fall* $\lambda x M$. Sei $N \in \mathsf{sb}$; zu zeigen ist $(\lambda x M)[\boldsymbol{x} := \boldsymbol{N}] \in \mathsf{sb}$. Nach IH ist $M[\boldsymbol{x}, x := \boldsymbol{N}, N] \in \mathsf{sb}$, also $(\lambda x M[\boldsymbol{x} := \boldsymbol{N}])N \in \mathsf{sb}$ nach Korollar 2.2.9. □

Korollar 2.2.12. *Jeder Term ist stark normalisierbar.* □

Wir zeigen jetzt die *Eindeutigkeit der Normalform*. Ein Term M heißt *lokal konfluent*, wenn es zu jedem Paar N_1 und N_2 mit $M \to^1 N_1$ und $M \to^1 N_2$ einen Term N gibt mit $N_1 \to^* N$ und $N_2 \to^* N$.

Lemma 2.2.13. *Jeder Term M ist lokal konfluent.*

Beweis. Induktion über M. *Fall* x. Klar. *Fall* $\lambda x M$.

$$\begin{tikzcd} & \lambda x M \arrow[dl, "1"'] \arrow[dr, "1"] & \\ \lambda x M' \arrow[dr, "*"'] & & \lambda x M'' \arrow[dl, "*"] \\ & \lambda x M''' & \end{tikzcd} \qquad \begin{tikzcd} & \lambda x.Mx \arrow[dl, "1"'] \arrow[dr, "1"] & \\ \lambda x.M'x \arrow[dr, "1"'] & & M \arrow[dl, "1"] \\ & M' & \end{tikzcd} \qquad \text{falls } x \notin \mathsf{FV}(M)$$

Fall MN.

$$\begin{tikzcd} & MN \arrow[dl, "1"'] \arrow[dr, "1"] & \\ M'N \arrow[dr, "*"'] & & M''N \arrow[dl, "*"] \\ & M'''N & \end{tikzcd} \qquad \begin{tikzcd} & MN \arrow[dl, "1"'] \arrow[dr, "1"] & \\ MN' \arrow[dr, "*"'] & & MN'' \arrow[dl, "*"] \\ & MN''' & \end{tikzcd} \qquad \begin{tikzcd} & MN \arrow[dl, "1"'] \arrow[dr, "1"] & \\ M'N \arrow[dr, "*"'] & & MN' \arrow[dl, "*"] \\ & M'N' & \end{tikzcd}$$

sowie

$$
\begin{tikzcd}[column sep=tiny]
& (\lambda xM)N \arrow[dl, "1"'] \arrow[dr, "1"] & \\
M[x := N] \arrow[dr, "*"'] & & (\lambda xM')N \arrow[dl, "1"] \\
& M'[x := N] &
\end{tikzcd}
\qquad
\begin{tikzcd}[column sep=tiny]
& (\lambda xM)N \arrow[dl, "1"'] \arrow[dr, "1"] & \\
M[x := N] \arrow[dr, "*"'] & & (\lambda xM)N' \arrow[dl, "1"] \\
& M[x := N'] &
\end{tikzcd}
\qquad
\begin{tikzcd}[column sep=tiny]
& (\lambda x.Mx)N \arrow[dl, "1"'] \arrow[dr, "1"] & \\
MN & = & MN
\end{tikzcd}
$$

wobei wir zweimal links unten Lemma 2.2.1 und im letzten Diagramm einmal eine β- und einmal eine η-Konversion benutzt haben. Wegen der vorausgesetzten Neutralität von M im Fall der η-Konversion können keine weiteren Fälle auftreten. □

Satz 2.2.14. *Alle Terme haben eine eindeutige Normalform, d.h. zu jedem Term M gibt es genau einen normalen Term N mit $M \to^* N$.*

Beweis. Jeder Term M ist nach Korollar 2.2.12 stark normalisierbar, wir können also Induktion nach $M \in \mathsf{sn}$ verwenden. Ist M normal, so ist die Behauptung trivial. Gelte also $M \to^1 N_1$ und $M \to^1 N_2$. Nach IH haben N_1 und N_2 eindeutige Normalformen. Mit dem vorigen Lemma ergibt sich die Behauptung, wie man anhand der folgenden Skizze leicht verfiziert.

$$
\begin{tikzcd}[column sep=small]
& & M \arrow[dl] \arrow[dr] & & \\
& N_1 \arrow[dl, "*"'] \arrow[dr, "*"] & \text{lok. konf.} & N_2 \arrow[dl, "*"'] \arrow[dr, "*"] & \\
K \arrow[dr, "*"'] & \mathrm{IH}(N_1) & \exists P_1 \arrow[dl, "*"] & & L \arrow[ddll, "*"] \\
& \exists P_2 \arrow[dr, "*"'] & & \mathrm{IH}(N_2) & \\
& & \exists N & &
\end{tikzcd}
$$

□

Die bisherigen Begriffe und Resultate dieses Kapitels lassen sich leicht auf den Fall übertragen, daß neben der Implikation $\to$ auch noch die Konjunktion $\wedge$ und der Allquantor $\forall$ zugelassen sind. Wir nennen einen Term wieder *neutral*, wenn er weder ein Paar noch eine Abstraktion ist. Die Regeln der β- and η-Konversion sind jetzt

$$
\begin{aligned}
\langle M_0, M_1\rangle i &\mapsto M_i, \\
(\lambda uM)N &\mapsto M[u := N], \\
(\lambda xM)t &\mapsto M[x := t], \\
\langle M0, M1\rangle &\mapsto M \quad \text{falls } M \text{ neutral ist}, \\
\lambda u.Mu &\mapsto M \quad \text{falls } u \notin \mathsf{FA}(M) \text{ und } M \text{ neutral ist}, \\
\lambda x.Mx &\mapsto M \quad \text{falls } M \text{ is neutral.}
\end{aligned}
$$

Die Eigenschaften der Substitution, d.h. Lemma 2.2.1, bleiben gültig. Der Beweis der starken Normalisierung läßt sich leicht erweitern; auch die zusätzlichen Fälle im Beweis der lokalen Konfluenz sind sehr einfach.

2.3 Anwendungen

Wir wollen jetzt einige Folgerungen aus der Tatsache ziehen, daß sich jede Herleitung in Normalform bringen läßt. Zu diesem Zweck müssen wir die Form normaler Herleitungen genauer analysieren.

Mit $k, l \in \{0,1\}^*$ bezeichnen wir *Knoten* in einem Term M. $\mathsf{Node}(M)$ ist die Menge aller Knoten in M, und ε bezeichnet die leere Liste, d.h. den Wurzelknoten. $k0$ und $k1$ sind die Erweiterungen des Knotens k durch 0 oder 1. Der Teilterm von M am Knoten k wird mit M/k bezeichnet. Wir schreiben $k \preceq l$ falls k ein Anfangssegment von l ist. Weiter sei

$$\begin{aligned}
\mathsf{Leafnode}(M) &:= \{\, k \in \mathsf{Node}(M) \mid M/k \text{ Variable oder Konstante} \,\} \\
\mathsf{Elimnode}(M) &:= \{\, k \in \mathsf{Node}(M) \mid M/k \text{ gebildet durch eine Beseitigungsregel} \,\} \\
\mathsf{Intronode}(M) &:= \{\, k \in \mathsf{Node}(M) \mid M/k \text{ gebildet durch eine Einführungsregel} \,\}
\end{aligned}$$

Für jedes $k \in \mathsf{Elimnode}(M) \cup \mathsf{Leafnode}(M)$ definieren wir seinen zugeordneten minimalen Knoten durch

$$\begin{aligned}
\mathsf{minnode}_M(\varepsilon) &:= \varepsilon. \\
\mathsf{minnode}_M(ki) &:= \begin{cases} \mathsf{minnode}_M(k), & \text{falls } i = 0 \text{ und } k \in \mathsf{Elimnode}(M); \\ ki, & \text{falls } i = 1 \text{ und } k \in \mathsf{Elimnode}(M), \text{ oder} \\ & k \in \mathsf{Intronode}(M). \end{cases}
\end{aligned}$$

Offensichtlich ist $\mathsf{minnode}_M(k) \in \mathsf{Elimnode}(M) \cup \mathsf{Leafnode}(M)$. Sei

$$\mathsf{Minnode}(M) := \{\, k \in \mathsf{Elimnode}(M) \cup \mathsf{Leafnode}(M) \mid \mathsf{minnode}_M(k) = k \,\}.$$

Für jedes $k \in \mathsf{Node}(M)$ definieren wir seinen zugeordneten Endknoten durch

$$\begin{aligned}
\mathsf{endnode}_M(\varepsilon) &:= \varepsilon. \\
\mathsf{endnode}_M(ki) &:= \begin{cases} \mathsf{endnode}_M(k), & \text{falls } k \in \mathsf{Intronode}(M), \text{ oder} \\ & i = 0 \text{ und } k \in \mathsf{Elimnode}(M); \\ ki, & \text{falls } i = 1 \text{ und } k \in \mathsf{Elimnode}(M). \end{cases}
\end{aligned}$$

Sei

$$\mathsf{Endnode}(M) := \{\, k \in \mathsf{Node}(M) \mid \mathsf{endnode}_M(k) = k \,\}.$$

Für normale Terme M stehen die Blätter und die Minimalknoten in einer eineindeutigen Beziehung. Man beachte, daß zwei minimale Knoten denselben Endknoten haben können.

Sei M ein Term und $k \in \mathsf{Leafnode}(M)$. Dann heißt

$$\{\, l \in \mathsf{Node}(M) \mid k \succeq l \succeq \mathsf{endnode}_M(k) \,\}$$

der durch den Knoten k bestimmte *Ast* in M. Insbesondere ist $\mathsf{minnode}_M(k)$ ein Element des durch k bestimmten Astes. Im Fall $\mathsf{endnode}_M(k) = \varepsilon$ spricht man von einem *Hauptast*. In einem normalen Term hat jeder Ast $k_1 \succeq \cdots \succeq k_n$ eine besonders durchsichtige Form: alle Beseitigungsregeln kommen vor allen Einführungsregeln.

Im Fall eines normalen Herleitungsterms wollen wir einige Folgerungen über die an den Knoten angehefteten Formeln ziehen. Dazu brauchen wir die folgenden Begriffe. Die Formeln A, B heißen *unmittelbare Teilformeln* von $A \wedge B$ und $A \to B$, und für jeden Term t ist die Formel $A[x := t]$ eine unmittelbare Teilformel von $\forall x A$. Die Relation "A ist eine *Teilformel* von B" ist der reflexiv-transitive Abschluß der Relation "unmittelbare Teilformel".

Lemma 2.3.1. *Sei $M^B[u_1^{A_1}, \dots, u_n^{A_n}]$ ein normaler Herleitungsterm, der nicht mit einer Einführung endet. Dann ist B Teilformel einer Formel A_i.*

Beweis. Induktion über M. *Fall* $M^B = M_0^{D \to B} M_1^D$. Da M normal ist, kann M_0 nicht mit einer Einführungsregel enden. Nach IH sind deshalb $D \to B$ und also auch B Teilformeln eines A_i. Die anderen Fälle (d.h. $M^B = N^{B \wedge D}0$, $M^B = N^{D \wedge B}1$, $M^B = N^{\forall x D}t$ mit $B = D[x := t]$ und $M = u_i^{A_i}$ mit $B = A_i$) behandelt man ähnlich. $\square$

Satz 2.3.2. *(Teilformeleigenschaft). Sei $M^B[u_1^{A_1}, \dots, u_n^{A_n}]$ eine normale Herleitung und N^C eine Teilherleitung von M^B. Dann ist C Teilformel von B oder von einem A_i.*

Beweis. Induktion über M. Wir können annehmen, daß $N^C \neq M^B$. *Fall* $M^B = M_0^{D\to B} M_1^D$. Nach IH ist C Teilformel von $D \to B$ oder von einem A_i. Da M normal ist, kann M_0 nicht mit einer Einführungsregel enden. Nach dem vorangehenden Lemma ist also $D \to B$ eine Teilformel eines A_i. Also ist auch C Teilformel eines A_i. Die Fälle $M^B = N^{B\wedge D}0$, $M^B = N^{D\wedge B}1$ and $M^B = N^{\forall x D}t$ mit $B = D[x := t]$ behandelt man ähnlich.

Fall $M^B = \lambda u^{B_0} M_1^{B_1}$ mit $B = B_0 \to B_1$. Nach IH ist C Teilformel von B_1, B_0 oder einem A_i. Also ist C Teilformel von $B_0 \to B_1$ oder einem A_i. Die Fälle $M = \lambda x M_0$ und $M = \langle M_0, M_1\rangle$ behandelt man ähnlich. □

Lemma 2.3.3. *Sei $M[u_1^{\forall \boldsymbol{x}_1 A_1}, \dots, u_m^{\forall \boldsymbol{x}_m A_m}] \colon \forall \boldsymbol{y} \forall y B_0$ eine normale Herleitung mit quantorenfreien Formeln $A_1, \dots, A_m, B_0$, die nicht mit einer Einführungsregel endet. Dann ist M von der Form $u_i \boldsymbol{t}$.*

Beweis. Induktion über M. Sei $B := \forall \boldsymbol{y} \forall y B_0$. Im *Fall* $M^B = u_i$ ist die Behauptung trivial. *Fall* $M^B = M_0^{D\to B} M_1^D$. Da M normal ist, kann M_0 nicht mit einer Einführungsregel enden. Also ist nach Lemma 2.3.1 $D \to B$ eine Teilformel von einem $\forall \boldsymbol{x}_i A_i$. Dies ist aber unmöglich, da B mit $\forall$ beginnt; also kann dieser Fall nicht eintreten. Die Fälle $M^B = N^{B\wedge D}0$ and $M^B = N^{D\wedge B}1$ behandelt man ähnlich. *Fall* $M^B = N^{\forall z D}t$ mit $B = D[z := t]$. Dann haben wir $\forall z D = \forall z \forall \boldsymbol{y} \forall y D_0$ mit quantorenfreiem D_0. Nach IH folgt $N = u_i \boldsymbol{t}$, also $M = u_i \boldsymbol{t} t$. □

Satz 2.3.4. *(*Herbrand*). Sei $M[u_1^{\forall \boldsymbol{x}_1 A_1}, \dots, u_m^{\forall \boldsymbol{x}_m A_m}] \colon B$ eine normale Herleitung mit quantorenfreien Formeln $A_1, \dots, A_m, B$. Dann enthält M^B keine All-Einführungsregel, und jedes Vorkommen einer Annahmevariablen u_i in M ist in einem Kontext*

$$\frac{u_i \colon \forall \boldsymbol{x}_i A_i \qquad \boldsymbol{t}_{ij}}{A_i[\boldsymbol{x}_i := \boldsymbol{t}_{ij}]}.$$

Insbesondere findet man Terme $\boldsymbol{t}_{11}, \dots, \boldsymbol{t}_{1n_1}, \dots, \boldsymbol{t}_{m1}, \dots, \boldsymbol{t}_{mn_m}$ mit

$$\begin{aligned} &A_1[\boldsymbol{x}_1 := \boldsymbol{t}_{11}], \dots, A_1[\boldsymbol{x}_1 := \boldsymbol{t}_{1n_1}], \dots, \\ &A_m[\boldsymbol{x}_m := \boldsymbol{t}_{m1}], \dots, A_m[\boldsymbol{x}_m := \boldsymbol{t}_{mn_m}] \vdash B. \end{aligned}$$

Beweis. Teil 1. Im *Fall* $M = u_i$ ist nichts zu zeigen. *Fall* $M^B = M_0^{D\to B} M_1^D$. Da M normal ist, kann M_0 nicht mit einer Einführungsregel enden. Nach Lemma 2.3.1 ist $D \to B$ Teilformel eines $\forall \boldsymbol{x}_i A_i$. Also ist $D \to B$ quantorenfrei, und die IH läßt sich anwenden. Die Fälle $M^B = N^{B\wedge D}0$ und $M^B = N^{D\wedge B}1$ behandelt man ähnlich. *Fall* $M^B = N^{\forall z D}t$ mit $B = D[z := t]$, also D quantorenfrei. Da M^B normal ist, kann N nicht mit einer Einführungsregel enden. Nach Lemma 2.3.3 muß N dann die Form $u_i \boldsymbol{t}$ haben, also $M = u_i \boldsymbol{t} t$.

Fall $M^B = \lambda u^{B_0} M_1^{B_1}$ mit $B = B_0 \to B_1$. Dann sind B_0 und B_1 quantorenfrei, und die IH liefert die Behauptung. Den Fall $M = \langle M_0, M_1\rangle$ behandelt man ähnlich.

Teil 2. Ersetzt man alle Vorkommen von $u_i \boldsymbol{t}_{ij}$ durch neue Annahmevariablen $v_{ij} \colon A_i[\boldsymbol{x}_i := \boldsymbol{t}_{ij}]$, so erhält man eine korrekte Herleitung, da M nach Teil (1) keine All-Einführungsregel enthält, also auch keine der Variablen $\boldsymbol{t}_{ij}$ weiter unten in der Herleitung gebunden ist. □

2.4 Normale versus nicht-normale Herleitungen

Wir wollen jetzt zeigen, daß die Forderung, stets eine normale Herleitung für eine herleitbare Formel anzugeben, manchmal unrealistisch sein kann. Wir geben Beispiele von Formeln D_k, die leicht mit nicht normalen Herleitungen (deren Knotenzahl linear in k ist) bewiesen werden können, für die aber jede normale Herleitung eine (in k) superexponentielle Anzahl von Knoten erfordert.

Das Beispiel steht in enger Beziehung zu Gentzens Beweis (in [10]) der transfiniten Induktion bis ω_k in der Arithmetik. Dort spielt die Funktion $y \oplus \omega^x$ eine wesentliche Rolle, und auch die Zuordnung einer "lifting"-Formel A^+ zu jeder Formel A, nämlich

$$A^+ := \forall y.(\forall z{\prec}y\, A[x := z]) \to \forall z{\prec}y{\oplus}\omega^x\, A[x := z].$$

Hier betrachten wir stattdessen die numerische Funktion $y+2^x$, und axiomatisieren ihren Graphen durch Hornklauseln. Die Formel D_k sagt aus, daß aus diesen Axiomen die Existenz von 2_k folgt (wobei $2_0 := 1$ und $2_{k+1} := 2^{2_k}$). Einen kurzen, nicht normalen Beweis dieser Tatsache kann man durch eine Modifikation der Gentzenschen Idee erhalten. Man kann sich dann überzeugen, daß jede nicht normale Herleitung von D_k mindestens 2_k Knotenenthalten muß.

Zur Angabe der Herleitungen machen wir wesentlich Gebrauch von dem durch $\neg\forall\neg$ definierten Existenzquantor $\exists$ (vgl. Abschnitt 1.2).

Man beachte, daß die Stabilitätsannahme $\neg\neg B \to B$ nicht benötigt wird, falls B kein Atom $\neq \bot$ als strikt positive Teilformel[1] enthält. Dies wird für die unten anzugebenden Herleitungen der Fall sein; B ist dort immer eine existentielle Formel.

Wir fixieren zunächst unsere Sprache. Verwendet wird ein dreistelliges Relationssymbol R, das den Graphen der Funktion $y + 2^x$ beschreiben soll; $R(y, x, z)$ meint also $y + 2^x = z$. Wir axiomatisieren R mit Hilfe von Hornklauseln. Zur Vereinfachung verwenden wir ein einstelliges Funktionssymbol s (zu verstehen als die Nachfolgerfunktion) und eine Konstante 0; man könnte auch ohne Funktionssymbole auskommen, aber dies macht die Formeln weniger lesbar und die Beweise weniger durchsichtig.

$$\begin{aligned}&\mathsf{Hyp}_1 \colon \forall y R(y, 0, s(y))\\ &\mathsf{Hyp}_2 \colon \forall y, x, z, z_1.R(y, x, z) \to R(z, x, z_1) \to R(y, s(x), z_1)\end{aligned}$$

Die Zielformel ist

$$C_k := \exists z_k, \dots , z_0.R(0, 0, z_k) \wedge R(0, z_k, z_{k-1}) \wedge \dots \wedge R(0, z_1, z_0).$$

Um einen kurzen Beweis von $D_k := \mathsf{Hyp}_1 \to \mathsf{Hyp}_2 \to C_k$ zu erhalten, verwenden wir Formeln A_i mit einem freien Parameter x; zur besseren Lesbarkeit schreiben wir $A[r]$ anstelle von $A[x := r]$.

$$\begin{aligned}A_0 &:= \forall y \exists z\, R(y, x, z),\\ A_{i+1} &:= \forall y.A_i[y] \to \exists z.A_i[z] \wedge R(y, x, z).\end{aligned}$$

Lemma 2.4.1. $\vdash \mathsf{Hyp}_1 \to \mathsf{Hyp}_2 \to A_i[0]$.

Beweis. Wir geben ein informales Argument, das sich leicht in einen formalen Beweis umformen läßt. Man beachte, daß die Existenzbeseitigung nur mit Existenzformeln als Konklusionen verwendet wird. Es ist also nicht nötig, Stabilitätsannahmen zu machen und wir erhalten eine Herleitung in der Minimallogik.

Fall $i = 0$. Klar nach Hyp_1.

Fall $i = 1$. Sei x mit $A_0[x]$ gegeben. Es genügt zu zeigen $A_0[s(x)]$, also $\forall y \exists z_1 R(y, s(x), z_1)$. Sei also y gegeben. Wir wissen

$$A_0[x] = \forall y \exists z\, R(y, x, z). \tag{2.6}$$

Anwendung von (2.6) auf unser y ergibt z mit $R(y, x, z)$. Nochmalige Anwendung von (2.6) auf dieses z ergibt z_1 mit $R(z, x, z_1)$. Nach Hyp_2 erhalten wir $R(y, s(x), z_1)$.

Fall $i + 2$. Sei x mit $A_{i+1}[x]$ gegeben. Es genügt zu zeigen $A_{i+1}[s(x)]$, also $\forall y.A_i[y] \to \exists z.A_i[z] \wedge R(y, s(x), z)$. Sei also y mit $A_i[y]$ gegeben. Wir wissen

$$A_{i+1}[x] = \forall y.A_i[y] \to \exists z_1.A_i[z_1] \wedge R(y, x, z_1). \tag{2.7}$$

Anwendung von (2.7) auf unser y ergibt z mit $A_i[z]$ und $R(y, x, z)$. Nochmalige Anwendung von (2.7) auf dieses z ergibt z_1 mit $A_i[z_1]$ und $R(z, x, z_1)$. Nach Hyp_2 erhalten wir $R(y, s(x), z_1)$. □

Man beachte, daß diese Herleitungen eine feste Länge haben, die nicht von i abhängt.

[1] Die Formeln A, B sind *unmittelbare strikt positive Teilformeln* von $A \wedge B$, B ist eine unmittelbare strikt positive Teilformel von $A \to B$, und für jeden Term t ist die Formel $A[x := t]$ eine unmittelbare strikt positive Teilformel von $\forall xA$. Die Relation "A ist eine *strikt positive Teilformel* von B" ist der reflexiv-transitive Abschluß dieser Relation.

Lemma 2.4.2. $\vdash \mathsf{Hyp}_1 \to \mathsf{Hyp}_2 \to C_k$.

Beweis. Wir geben ein informales Argument, das sich leicht in einen formalen Beweis umformen läßt. Man beachte wieder, daß die Existenzbeseitigung nur mit Existenzformeln als Konklusionen verwendet wird, und wir deshalb eine Herleitung in der Minimallogik erhalten.

$A_k[0]$ angewandt auf 0 und $A_{k-1}[0]$ liefert z_k mit $A_{k-1}[z_k]$ und $R(0,0,z_k)$.
$A_{k-1}[z_k]$ angewandt auf 0 und $A_{k-2}[0]$ liefert z_{k-1} mit $A_{k-2}[z_{k-1}]$ und $R(0,z_k,z_{k-1})$.
$A_1[z_2]$ angewandt auf 0 und $A_0[0]$ liefert z_1 mit $A_0[z_1]$ und $R(0,z_2,z_1)$.
$A_0[z_1]$ angewandt auf 0 liefert z_0 mit $R(0,z_1,z_0)$. □

Man beachte, daß diese Herleitungen eine in k lineare Länge haben. Wir wollen jetzt die Länge einer beliebigen normalen Herleitung von D_k nach unten abschätzen.

Lemma 2.4.3. *Jede normale Herleitung von C_k aus Hyp_1 und Hyp_2 hat mindestens 2_k Knoten.*

Beweis. Sei M eine normale Herleitung von $\bot$ aus Hyp_1, Hyp_2 und der zusätzlichen Annahme

$$u\colon \forall z_k,\dots,z_0.R(0,0,z_k) \to R(0,z_k,z_{k-1}) \to \dots \to R(0,z_1,z_0) \to \bot.$$

Wir können annehmen, daß M keine freien Objektvariablen enthält (anderfalls ersetze man sie durch 0). Der Hauptast von M muß mit u beginnen, und seine Nebenprämissen sind alle von der Form $R(0,s^n(0),s^k(0))$.

Man beachte, daß jede normale Herleitung von $R(s^m(0),s^n(0),s^k(0))$ aus Hyp_1, Hyp_2 und u mindestens 2^n Vorkommen von Hyp_1 enthält, und daß $k = m + 2^n$ gelten muß. Dies sieht man leicht durch Induktion über n. Man beachte auch, daß eine solche Herleitung die Annahmenvariable u nicht enthalten kann.

Wendet man diese Beobachtung auf die obigen Herleitung der Nebenprämissen an, so sieht man, daß sie herleiten

$$R(0,0,s^{2^0}(0)),\quad R(0,s^{2^0}(0),s^{2^{2^0}}(0)),\quad \dots\quad R(0,s^{2_{k-1}}(0),s^{2_k}(0)).$$

Die letzte dieser Herleitungen verwendet mindestens $2^{2_{k-1}} = 2_k$-mal Hyp_1. □

2.5 Anmerkungen

Das System des natürlichen Schließens wurde von Gentzen [9] eingeführt. Den Beweis der starken Normalisierung haben wir mit einer auf W.W. Tait [39] zurückgehenden Methode geführt.

Das Beispiel von Herleitungen mit kurzen nicht normalen Beweisen und superexponentiell langen Normalformen geht zurück auf Statman [37] und Orevkov [22]. Orevkovs Ergebnis ist eine Adaption des Statmanschen für Sprachen mit Funktionssymbolen. Unsere Darstellung folgt [42].

3. Berechenbarkeit

3.1 Primitiv rekursive Funktionen

Wir beschränken uns auf die Diskussion der Berechenbarkeit von Funktionen $F\colon \mathbb{N}^n \to \mathbb{N}$. Eine solche Funktion soll *berechenbar* (im intuitiven Sinn) heißen, wenn es einen Algorithmus gibt, der für jedes Argumentetupel $a_1, \dots, a_n$ terminiert und den Funktionswert $f(a_1, \dots, a_n)$ liefert. In diesem Abschnitt behandeln wir besonders einfache berechenbare Funktionen, nämlich die von HILBERT eingeführten sogenannten primitiv rekursiven Funktionen, die aus einfachen Ausgangsfunktionen durch Komposition (oder Einsetzung) und primitive Rekursion erzeugt werden können. Obwohl diese Definitionsschemata recht speziell sind, wird sich zeigen, daß nahezu alle in der Mathematik verwendeten zahlentheoretischen Funktionen primitiv rekursiv sind. Insbesondere werden wir in Abschnitt 3.3 Terme, Formeln und Herleitungen durch natürliche Zahlen kodieren und zeigen, daß die charakteristischen Funktionen der entsprechenden Mengen von Kodenummern primitiv rekursiv sind.

Wir definieren induktiv für jedes $n \in \mathbb{N}$ eine Menge PR^n von n-stelligen Funktionssymbolen.

1. $0^n \in \mathsf{PR}^n$ $(n \geq 0)$, $\mathsf{S} \in \mathsf{PR}^1$, $I_i^n \in \mathsf{PR}^n$ $(1 \leq i \leq n)$.
2. Sind $g_1, \dots, g_m \in \mathsf{PR}^n$, $h \in \mathsf{PR}^m$ und $m \geq 1$, $n \geq 0$, so ist $(\circ h g_1 \dots g_m) \in \mathsf{PR}^n$.
3. Ist $g \in \mathsf{PR}^n$ und $h \in \mathsf{PR}^{n+2}$, so ist $(Rgh) \in \mathsf{PR}^{n+1}$.

Wir setzen $\mathsf{PR} := \bigcup_{n \in \mathbb{N}} \mathsf{PR}^n$ und $0 := 0^0$.

Zu der durch PR gegebenen Sprache definieren wir eine Standardstruktur $\mathcal{N}$ wie folgt. Es ist $|\mathcal{N}| := \mathbb{N}$ und

$$
\begin{aligned}
&(0^n)^{\mathcal{N}}(a_1, \dots, a_n) := 0, \\
&\mathsf{S}^{\mathcal{N}}(a) := a + 1, \\
&(I_i^n)^{\mathcal{N}}(a_1, \dots, a_n) := a_i, \\
&(\circ h g_1 \dots g_m)^{\mathcal{N}}(\boldsymbol{a}) := h^{\mathcal{N}}(g_1^{\mathcal{N}}(\boldsymbol{a}), \dots, g_m^{\mathcal{N}}(\boldsymbol{a})), \\
&(Rgh)^{\mathcal{N}}(0, \boldsymbol{b}) := g^{\mathcal{N}}(\boldsymbol{b}), \\
&(Rgh)^{\mathcal{N}}(a+1, \boldsymbol{b}) := h^{\mathcal{N}}(a, (Rgh)^{\mathcal{N}}(a, \boldsymbol{b}), \boldsymbol{b}).
\end{aligned}
$$

Eine Funktion $F\colon \mathbb{N}^n \to \mathbb{N}$ heißt *primitiv rekursiv*, wenn es ein $f \in \mathsf{PR}^n$ gibt mit $F = f^{\mathcal{N}}$. Eine Relation $R \subseteq \mathbb{N}^n$ heißt primitiv rekursiv, wenn ihre *charakteristische Funktion* 1_R primitiv rekursiv ist, wobei

$$
1_R(\boldsymbol{a}) := \begin{cases} 1, & \text{falls } \boldsymbol{a} \in R; \\ 0, & \text{sonst.} \end{cases}
$$

Wir wollen zunächst zeigen, daß die Menge der primitiv rekursiven Funktionen gegen *explizite Definitionen* abgeschlossen ist. Dazu definieren wir für jeden PR-Term t und verschiedene Variablen $x_1, \dots, x_n$ $(n \geq 1)$ mit $\mathsf{vars}(t) \subseteq \{x_1, \dots, x_n\}$ ein n-stelliges Funktionssymbol $\lambda x_1, \dots, x_n\, t$ wie folgt.

$$
\begin{aligned}
&\lambda x_1, \dots, x_n\, 0 := 0^n, \\
&\lambda x_1, \dots, x_n\, x_i := I_i^n, \\
&\lambda x_1, \dots, x_n . h t_1 \dots t_m := (\circ h g_1 \dots g_m), \quad \text{wobei } g_i := \lambda x_1, \dots, x_n\, t_i.
\end{aligned}
$$

Lemma 3.1.1. *Für jeden* PR-*Term* t *und verschiedene Variablen* $x_1, \ldots, x_n$ *(*$n \geq 1$*) mit* $\mathsf{vars}(t) \subseteq \{x_1, \ldots, x_n\}$ *gilt*

$$(\lambda x_1, \ldots, x_n\, t)^{\mathcal{N}}(a_1, \ldots, a_n) = t^{\mathcal{N}}[a_1, \ldots, a_n/x_1, \ldots, x_n].$$

Beweis. Induktion über t. Wir schreiben $t^{\mathcal{N}}[\boldsymbol{a}]$ für $t^{\mathcal{N}}[a_1, \ldots, a_n/x_1, \ldots, x_n]$.

$$(\lambda \boldsymbol{x}\, 0)^{\mathcal{N}}(\boldsymbol{a}) = (0^n)^{\mathcal{N}}(\boldsymbol{a}) = 0 = 0^{\mathcal{N}}[\boldsymbol{a}].$$
$$(\lambda \boldsymbol{x}\, x_i)^{\mathcal{N}}(\boldsymbol{a}) = (I_i^n)^{\mathcal{N}}(\boldsymbol{a}) = a_i = x_i^{\mathcal{N}}[\boldsymbol{a}].$$

Sei $t = ht_1 \ldots t_m$ und $g_i := \lambda x_1, \ldots, x_n\, t_i$. Dann ist

$$\begin{aligned}(\lambda \boldsymbol{x}\, t)^{\mathcal{N}}(\boldsymbol{a}) &= (\circ h g_1 \ldots g_m)^{\mathcal{N}}(\boldsymbol{a}) \\ &= h^{\mathcal{N}}(g_1^{\mathcal{N}}(\boldsymbol{a}), \ldots, g_m^{\mathcal{N}}(\boldsymbol{a})) \\ &= h^{\mathcal{N}}(t_1^{\mathcal{N}}[\boldsymbol{a}], \ldots, t_m^{\mathcal{N}}[\boldsymbol{a}]) \quad \text{nach IH} \\ &= t^{\mathcal{N}}[\boldsymbol{a}].\end{aligned}$$

□

Wir verwenden $a, b, c, i, j, k, \ell, m, n$ als Mitteilungszeichen für natürliche Zahlen. Ist $f \in \mathsf{PR}^n$, so bezeichnen wir die Funktion $f^{\mathcal{N}} : \mathbb{N}^n \to \mathbb{N}$ ebenfalls kurz mit f; in diesem Sinn ist PR^n die Menge der n-stelligen und PR die Menge aller primitiv rekursiven Funktionen. PR ist also die kleinste Funktionenmenge, die die *Ausgangsfunktionen* $0^n, I_i^n, \mathsf{S}$ enthält und abgeschlossen ist gegen *Komposition* (oder *Einsetzung*) $\circ$ und *primitive Rekursion* R.

Beispiele. Setzt man $\tilde{+} := RI_1^1(\lambda x, y, z.\mathsf{S}(y))$, so ist

$$\tilde{+}(0, b) = b,$$
$$\tilde{+}(a+1, b) = \mathsf{S}(\tilde{+}(a, b)),$$

und setzt man $\tilde{\cdot} := R0^1(\lambda x, y, z.\tilde{+}yz)$, so ist

$$\tilde{\cdot}(0, b) = 0,$$
$$\tilde{\cdot}(a+1, b) = \tilde{+}(\tilde{\cdot}(a, b), b).$$

Es gilt also $\tilde{+}(a, b) = a+b$ und $\tilde{\cdot}(a, b) = a \cdot b$; wir schreiben deshalb $+, \cdot$ für $\tilde{+}, \tilde{\cdot}$. Weiter sei $\mathsf{pd} := R0(\lambda x, y\, x)$ und $\dot{-} := RI_1^1(\lambda x, y, z.\mathsf{pd}(y))$. Dann gilt

$$\mathsf{pd}(0) = 0 \qquad\qquad b \dot{-} 0 = b,$$
$$\mathsf{pd}(a+1) = a \qquad\qquad b \dot{-} (a+1) = \mathsf{pd}(b \dot{-} a)$$

wobei wir für $\dot{-}$ die übliche Infixschreibweise verwendet haben. Es ist also

$$b \dot{-} a = \begin{cases} 0, & \text{falls } b \leq a; \\ b - a, & \text{sonst.} \end{cases}$$

Für $f \in \mathsf{PR}^{n+1}$ sei $\sum f := R0^n(\lambda x, y, \boldsymbol{z}\, (y + f(x, \boldsymbol{z})))$, also

$$\left(\sum f\right)(0, \boldsymbol{b}) = 0,$$
$$\left(\sum f\right)(a+1, \boldsymbol{b}) = \left(\sum f\right)(a, \boldsymbol{b}) + f(a, \boldsymbol{b}).$$

In Anlehnung an die übliche mathematische Terminologie schreiben wir $\sum_{i<a} f(i, \boldsymbol{b})$ für $\left(\sum f\right)(a, \boldsymbol{b})$. Ähnlich setzen wir $\prod f := R(\circ \mathsf{S}0^n)\lambda x, y, \boldsymbol{z}\, (y \cdot f(x, \boldsymbol{z}))$, also $(\prod f)(a, \boldsymbol{b}) = \prod_{x<a} f(x, \boldsymbol{b})$. Wir schreiben auch $\sum_{i \leq a} f(i, \boldsymbol{b})$ für $\sum_{i<\mathsf{S}(a)} f(i, \boldsymbol{b})$ und entsprechend bei Produkten.

Man beachte, daß eine Relation $R \subseteq \mathbb{N}^n$ genau dann primitiv rekursiv ist, wenn es ein $f \in \mathsf{PR}^n$ gibt mit $R = \{\, \boldsymbol{a} \in \mathbb{N}^n \mid f(\boldsymbol{a}) = 0 \,\}$; zum Beweis beachte man $1_R(\boldsymbol{a}) = 1 \dot{-} f(\boldsymbol{a})$ bzw. setze $f(\boldsymbol{a}) = 1 \dot{-} 1_R(\boldsymbol{a})$.

Um neben Termen auch Formeln betrachten zu können, nehmen wir im folgenden an, daß die Sprache PR noch das Gleichheitssymbol $=$ enthält, und daß $=^{\mathcal{N}}$ in der Standardstruktur $\mathcal{N}$ die Gleichheit auf $\mathbb{N}$ ist. Wir schreiben $s < t$ für $\mathsf{S}(s) \dot{-} t = 0$; offenbar gilt dann

$$\mathcal{N} \models (s < t)[\eta] \iff s^{\mathcal{N}}[\eta] < t^{\mathcal{N}}[\eta].$$

Eine Relation $R \subseteq \mathbb{N}^n$ heißt *definierbar* durch eine Formel $A[x_1, \ldots, x_n]$ der Sprache PR, wenn

$$R = \{\, (a_1, \ldots, a_n) \in \mathbb{N}^n \mid \mathcal{N} \models A[a_1, \ldots, a_n/x_1, \ldots, x_n] \,\}.$$

Offenbar sind genau die primitiv rekursiven Relationen durch atomare PR-Formeln definierbar. Wir wollen jetzt zeigen, daß dasselbe auch für eine reichere Formelmenge der Fall ist, nämlich für die sogenannten Δ_0-Formeln der Sprache PR.

Wir setzen im Fall $x \notin \mathsf{vars}(t)$

$$\begin{aligned}
\forall x{<}t\, A &:= \forall x.x < t \to A,\\
\exists x{<}t\, A &:= \neg\forall x.x < t \to \neg A,\\
\forall x{\le}t\, A &:= \forall x{<}\mathsf{S}(t)\, A,\\
\exists x{\le}t\, A &:= \exists x{<}\mathsf{S}(t)\, A.
\end{aligned}$$

Man beachte, daß $\exists x{<}t\, A$ logisch äquivalent ist zu $\exists x.x < t \wedge A$. Δ_0-*Formeln* werden induktiv definiert durch

- Jede PR-Primformel ist eine Δ_0-Formel; insbesondere ist also $\perp$ eine Δ_0-Formel.
- Sind A und B Δ_0-Formeln, so auch $A \to B$.
- Ist A eine Δ_0-Formel und t ein PR-Term mit $x \notin \mathsf{vars}(t)$, so ist auch $\forall x{<}t\, A$ eine Δ_0-Formel.

Es folgt, daß mit A im Fall $x \notin \mathsf{vars}(t)$ auch $\exists x{<}t\, A$ eine Δ_0-Formel ist.

Lemma 3.1.2. *Ist A eine Δ_0-Formel mit $\mathsf{FV}(A) \subseteq \{x_1, \ldots, x_n\}$, so ist die Relation*

$$\{\, (a_1, \ldots, a_n) \in \mathbb{N}^n \mid \mathcal{N} \models A[a_1, \ldots, a_n/x_1, \ldots, x_n] \,\}$$

primitiv rekursiv.

Beweis. Wir definieren für jede Δ_0-Formel A einen PR-Term r_A mit $\mathsf{vars}(r_A) = \mathsf{FV}(A)$ so daß für jede $\mathcal{N}$-Belegung η gilt $r_A^{\mathcal{N}}[\eta] = 0 \iff \mathcal{N} \models A[\eta]$.

$$\begin{aligned}
&r_\perp := \mathsf{S}(0),\\
&r_{s=t} := (s \dot{-} t) + (t \dot{-} s),\\
&r_{A\to B} := (1 \dot{-} r_A) \cdot r_B,\\
&r_{\forall x<t\, A} := \Big(\sum f\Big) t\boldsymbol{y} \quad \text{mit } f = \lambda x, \boldsymbol{y}\, r_A \text{ und } \mathsf{FV}(\forall x{<}t\, A) = \{\boldsymbol{y}\}.
\end{aligned}$$
□

Es folgt, daß die Menge aller primitiv rekursiven Relationen abgeschlossen ist gegen $\cap, \cup, \setminus$, beschränkte Quantifikation sowie gegen die Einsetzung von primitiv rekursiven Funktionen.

Lemma 3.1.3. *Sind $f_1, \ldots, f_{k+1} \in \mathsf{PR}^n$ und sind $R_1, \ldots, R_k \subseteq \mathbb{N}^n$ paarweise disjunkte primitiv rekursive Relationen, so ist auch die wie folgt definierte Funktion $f\colon \mathbb{N}^n \to \mathbb{N}$ primitiv rekursiv.*

$$f(\boldsymbol{a}) := \begin{cases} f_1(\boldsymbol{a}), & \textit{falls } R_1(\boldsymbol{a});\\ \vdots & \vdots\\ f_k(\boldsymbol{a}), & \textit{falls } R_k(\boldsymbol{a});\\ f_{k+1}(\boldsymbol{a}), & \textit{sonst.} \end{cases}$$

Beweis. Sei $R_{k+1} := \mathbb{N}^n \setminus (R_1 \cup \dots \cup R_k)$. Dann ist

$$f = \lambda \boldsymbol{x}. f_1(\boldsymbol{x}) \cdot 1_{R_1}(\boldsymbol{x}) + \dots + f_k(\boldsymbol{x}) \cdot 1_{R_k}(\boldsymbol{x}) + f_{k+1}(\boldsymbol{x}) \cdot 1_{R_{k+1}}(\boldsymbol{x}). \qquad \square$$

Der *beschränkte μ-Operator* macht aus jedem $g \colon \mathbb{N}^{n+1} \to \mathbb{N}$ ein $\overline{\mu}g \colon \mathbb{N}^n \to \mathbb{N}$, das wie folgt definiert ist.

$$(\overline{\mu}g)(a, \boldsymbol{b}) := \begin{cases} \min\{\, i \mid g(i, \boldsymbol{b}) = 0 \,\}, & \text{falls } \exists i{<}a\, (g(i, \boldsymbol{b}) = 0); \\ a, & \text{sonst.} \end{cases}$$

Wir schreiben auch $\mu i{<}a\, (g(i, \boldsymbol{b}) = 0)$ für $(\overline{\mu}g)(a, \boldsymbol{b})$ und $\mu i{\le}a\, (g(i, \boldsymbol{b}) = 0)$ für $\mu i{<}\mathsf{S}(a)\, (g(i, \boldsymbol{b}) = 0)$.

Lemma 3.1.4. *Mit $g \colon \mathbb{N}^{n+1} \to \mathbb{N}$ ist auch $\overline{\mu}g$ primitiv rekursiv.*

Beweis. Sei

$$h(c, \boldsymbol{b}) := \begin{cases} c, & \text{falls } g(c, \boldsymbol{b}) = 0 \text{ und } \forall i{<}c\, (g(i, \boldsymbol{b}) \neq 0); \\ 0, & \text{sonst.} \end{cases}$$

h ist nach den Lemmata 3.1.2 und 3.1.3 primitiv rekursiv, und wir haben

$$\mu i{<}a\, (g(i, \boldsymbol{b}) = 0) = \begin{cases} \sum_{i<a} h(i, \boldsymbol{b}), & \text{falls } \exists i{<}a\, (g(i, \boldsymbol{b}) = 0); \\ a, & \text{sonst.} \end{cases} \qquad \square$$

Beispiele.

$$\begin{aligned}
& b^a := \prod_{i<a} b, \\
& a! := \dots (\text{Übung}), \\
& a | b :\leftrightarrow \exists x{\le}b\, (b = a \cdot x), \\
& \lfloor a/b \rfloor := \mu x{\le}a\, (a < b \cdot (x+1)), \\
& r(a, b) := \dots (\text{Übung}), \\
& \mathrm{Pr}(a) :\leftrightarrow \dots (\text{Übung}), \\
& q(a) := \mu x{\le}a!{+}1 . x > a \wedge \mathrm{Pr}(x), \\
& p_i := \mu x{\le}2^{2^i} \big(\sum_{y \le x} 1_{\mathrm{Pr}}(y) = i + 1 \big).
\end{aligned}$$

Die Abschätzung $p_i \le 2^{2^i}$ für die i-te Primzahl p_i beweist man durch Induktion über i: Für $i = 0$ ist dies klar, und für $i \ge 1$ erhält man

$$p_i \le p_0 p_1 \cdots p_{i-1} + 1 = 2^{2^0} 2^{2^1} \cdots 2^{2^{i-1}} + 1 = 2^{2^i - 1} + 1 < 2^{2^i}.$$

Man kann die Eindeutigkeit der Darstellung natürlicher Zahlen als Primzahlpotenzprodukt verwenden, um eine endliche Folge $a_0, a_1, \dots, a_{n-1}$ natürlicher Zahlen durch

$$a := \prod_{i<n} p_i^{a_i}$$

zu kodieren; dies nennt man die *Primzahlpotenzkodierung*. Die i-te Komponente der durch a kodierten Folge kann man aus a und i ablesen durch

$$\exp(i, a) := \mu x{\le}a\, (p_i^{x+1} \nmid a).$$

Offenbar ist exp und für jedes feste n die Funktion $(a_0, \dots, a_{n-1}) \mapsto \prod_{i<n} p_i^{a_i}$ primitiv rekursiv.

Bei der Primzahlpotenzkodierung wird das Zahlenpaar a, b durch die Zahl $2^a 3^b$ kodiert. Dieses exponentielle Wachstum einer Paarkodierungsfunktion ist aber nicht notwendig, denn man kann $\mathbb{N} \times \mathbb{N}$ auch wie folgt abzählen.

$$
\begin{array}{lllll}
\vdots \\
10 \\
6 & \dots \\
3 & 7 & \dots \\
1 & 4 & 8 & \dots \\
0 & 2 & 5 & 9 & \dots
\end{array}
$$

An der Stelle $(0, b)$ steht offenbar die Summe der Längen der vorangehenden Diagonalen, und auf der nächsten Diagonalen bleibt $a + b$ konstant. Nennt man die an der Stelle (a, b) eingetragene Zahl $\pi(a, b)$, so ergibt sich

$$\pi(a, b) = \Big(\sum_{i \le a+b} i\Big) + a = \frac{1}{2}(a + b)(a + b + 1) + a.$$

Offenbar ist $\pi\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ bijektiv. Ferner gilt $a, b \le \pi(a, b)$ und im Fall $\pi(a, b) \ne 0$ auch $a < \pi(a, b)$. Setzt man

$$
\begin{aligned}
\pi_1(c) &:= \mu x{\le}c \exists y{\le}c\, (\pi(x, y) = c), \\
\pi_2(c) &:= \mu y{\le}c \exists x{\le}c\, (\pi(x, y) = c),
\end{aligned}
$$

so folgt unmittelbar $\pi_i(c) \le c$ für $i \in \{1, 2\}$ und

$$
\begin{aligned}
&\pi_1(\pi(a, b)) = a, \\
&\pi_2(\pi(a, b)) = b, \\
&\pi(\pi_1(c), \pi_2(c)) = c.
\end{aligned}
$$

Ferner sind π, π_1, π_2 aufgrund ihrer Definitionen primitiv rekursiv.

Mit Hilfe der *Paarfunktion* π kann man jetzt leicht (wie etwa in der Programmsprache Lisp oder Scheme) eine andere Kodierung von endlichen Folgen natürlicher Zahlen definieren. Wir setzen

$$
\begin{aligned}
&\langle\rangle := 0, \\
&\langle a_0, a_1, \dots, a_n\rangle := \pi(a_0, \langle a_1, \dots, a_n\rangle) + 1.
\end{aligned}
$$

Offenbar läßt sich dann jede Zahl a eindeutig in der Form $a = \langle a_0, a_1, \dots, a_{n-1}\rangle$ darstellen. Aus a lassen sich die Bestandteile dieser Darstellung wieder ablesen durch die folgenden Funktionen.

$$
\begin{aligned}
&\mathsf{hd}(a) := \pi_1(a \dot{-} 1) \quad \mathsf{hd} \text{ steht für "head"}, \\
&\mathsf{tl}(a) := \pi_2(a \dot{-} 1) \quad \mathsf{tl} \text{ steht für "tail"}, \\
&\tau(a, 0) := a, \\
&\tau(a, k + 1) := \mathsf{tl}(\tau(a, k)).
\end{aligned}
$$

Man beachte, daß für $\tau(a, k) \ne 0$ gilt $\tau(a, k + 1) < \tau(a, k)$.

$$
\begin{aligned}
&\mathsf{lh}(a) := \mu k{\le}a(\tau(a, k) = 0), \\
&(a)_i := \begin{cases} \mathsf{hd}(\tau(a, i)), & \text{falls } i < \mathsf{lh}(a); \\ 0, & \text{sonst.} \end{cases}
\end{aligned}
$$

Es folgt $\mathsf{lh}(a) \le a$ und $(a)_i \le a$; für $i < \mathsf{lh}(a)$ gilt sogar $(a)_i < a$, denn dann ist

$$(a)_i = \mathsf{hd}(\tau(a, i)) < \tau(a, i) \le a.$$

Offenbar sind die Funktionen $(a_0, \dots, a_n) \mapsto \langle a_0, \dots, a_n\rangle$ für jedes feste n, sowie $\mathsf{hd}, \mathsf{tl}, \tau, \mathsf{lh}$ und $(a, i) \mapsto (a)_i$ primitiv rekursiv. Wir schreiben $(a)_{i,j}$ für $((a)_i)_j$ und $(a)_{i,j,k}$ für $(((a)_i)_j)_k$.

Lemma 3.1.5.

$$\begin{aligned}
&\mathsf{hd}(\langle a_0, a_1, \dots, a_n\rangle) = a_0,\\
&\mathsf{tl}(\langle a_0, a_1, \dots, a_n\rangle) = \langle a_1, \dots, a_n\rangle,\\
&\tau(\langle a_0, \dots, a_k, \dots, a_n\rangle, k) = \langle a_k, \dots, a_n\rangle,\\
&\mathsf{lh}(\langle a_0, \dots, a_{n-1}\rangle) = n,\\
&(\langle a_0, \dots, a_{n-1}\rangle)_i = a_i \quad \textit{für } i < n.
\end{aligned}$$ □

Wir werden beide Kodierungen von endlichen Folgen natürlicher Zahlen nebeneinander verwenden. Ein Vorteil der Primzahlpotenzkodierung liegt darin, daß man die bekannte Produktschreibweise verwenden kann. Dies nutzen wir aus in dem folgenden Beweis, daß die Menge der primitiv rekursiven Funktionen abgeschlossen ist gegen *Wertverlaufsrekursion*. Für jedes $f\colon \mathbb{N}^{n+1} \to \mathbb{N}$ definieren wir seine *Wertverlaufsfunktion* $\overline{f}$ durch

$$\overline{f}(a, \boldsymbol{b}) := \prod_{i<a} p_i^{f(i,\boldsymbol{b})}.$$

Es folgt, daß mit f auch $\overline{f}$ primitiv rekursiv ist.

Lemma 3.1.6. *(Wertverlaufsrekursion). Mit g ist auch die durch*

$$f(a, \boldsymbol{b}) := g(a, \overline{f}(a, \boldsymbol{b}), \boldsymbol{b})$$

definierte Funktion f primitiv rekursiv.

Beweis. Es gilt

$$\begin{aligned}
&\overline{f}(0, \boldsymbol{b}) = 1,\\
&\overline{f}(a+1, \boldsymbol{b}) = \overline{f}(a, \boldsymbol{b}) \cdot p_a^{g(a,\overline{f}(a,\boldsymbol{b}),\boldsymbol{b})}.
\end{aligned}$$

Also ist mit g auch $\overline{f}$ primitiv rekursiv, und deshalb wegen $f(a, \boldsymbol{b}) = \exp(a, \overline{f}(a+1, \boldsymbol{b}))$ auch f. □

Wir zeigen jetzt, daß die Menge der primitiv rekursiven Funktionen auch gegen primitive Rekursion mit Einsetzungen in Parameterstellen abgeschlossen ist; dies wurde zuerst von R. Péter bewiesen (s. etwa [28]). Zur Vorbereitung beweisen wir ein Lemma über primitive Rekursion mit Abstieg bezüglich einer "Maßzahl" $h(a, \boldsymbol{b})$. Dieses Lemma werden wir später als Lemma 3.1.10 noch wesentlich verschärfen können.

Lemma 3.1.7. *Alle i mit $h(i, \boldsymbol{b}) < c$ seien abschätzbar durch $i < h'(c, \boldsymbol{b})$. Es gelte*

$$\begin{aligned}
&f(0, \boldsymbol{b}) = g_1(\boldsymbol{b}),\\
&f(a, \boldsymbol{b}) = g_2(a, f(t_1(a, \boldsymbol{b}), \boldsymbol{b}), \dots, f(t_k(a, \boldsymbol{b}), \boldsymbol{b}), \boldsymbol{b}) \quad \textit{falls } a > 0,\\
&h(t_j(a, \boldsymbol{b}), \boldsymbol{b}) < h(a, \boldsymbol{b}) \quad \textit{für } 1 \le j \le k, \textit{ falls } a > 0.
\end{aligned}$$

Dann ist mit $g_1, g_2, t_1, \dots, t_k, h$ und h' auch f primitiv rekursiv.

Beweis. Zur Vereinfachung lassen wir die Parameter $\boldsymbol{b}$ weg. Sei

$$\tilde{f}(c) := \prod_{h(i)<c} p_i^{f(i)}.$$

Dann gilt offenbar $f(a) = \exp(a, \tilde{f}(h(a)+1))$ und

$$\begin{aligned}
\tilde{f}(0) &= 1,\\
\tilde{f}(c+1) &= \tilde{f}(c) \cdot \prod_{h(i)=c} p_i^{f(i)}\\
&= \tilde{f}(c) \cdot \Bigl(\prod_{\substack{0<i<h'(c+1)\\ h(i)=c}} p_i^{g_2(i,\exp(t_1(i),\tilde{f}(c)),\dots,\exp(t_k(i),\tilde{f}(c)))}\Bigr) \cdot \begin{cases} p_0^{g_1}, & \text{falls } h(0) = c,\\ 1, & \text{sonst.}\end{cases}
\end{aligned}$$

Sind $g_1, g_2, t_1, \dots, t_k, h, h'$ primitiv rekursiv, so ist offenbar auch $\tilde{f}$ und damit auch f primitiv rekursiv. $\square$

Satz 3.1.8. *(Primitive Rekursion mit Einsetzungen in Parameterstellen). Sei*

$$\begin{aligned} &f(0, b, \boldsymbol{c}) = g_1(b, \boldsymbol{c}), \\ &f(a+1, b, \boldsymbol{c}) = g_2(a, f(a, h_1(a, b, \boldsymbol{c}), \boldsymbol{c}), \dots, f(a, h_k(a, b, \boldsymbol{c}), \boldsymbol{c}), b, \boldsymbol{c}). \end{aligned}$$

Dann ist mit $g_1, g_2, h_1, \dots, h_k$ auch f primitiv rekursiv.

Beweis. Zur Vereinfachung lassen wir die Parameter $\boldsymbol{c}$ weg. Sei

$$\hat{h}(a, b) := \big(\max_{\substack{i \le a, j \le b \\ 1 \le \ell \le k}} h_\ell(i, j)\big) + b + 1$$

und

$$\begin{aligned} &h^*(0, a, b) := b, \\ &h^*(c+1, a, b) := \hat{h}(a, h^*(c, a, b)). \end{aligned}$$

Dann sind mit $h_1, \dots, h_k$ auch $\hat{h}, h^*$ primitiv rekursiv. Man beachte hier, daß mit g auch $\max_{i<a} g(i, \boldsymbol{b})$ primitiv rekursiv ist. Ferner zeigt man leicht durch Induktion über c

$$\begin{aligned} &h^*(c+1, a, b) = h^*(c, a, \hat{h}(a, b)), \\ &h^*(c, a, b) \ge \max(c, b). \end{aligned}$$

Wir verwenden jetzt Lemma 3.1.7 mit $h(\pi(a, b)) := h^*(a, a, b)$ (und $h(0) := 0$). Sei also $h(x) < c$. Dann folgt

$$c > h^*(\pi_1(x), \pi_1(x), \pi_2(x)) \ge \max(\pi_1(x), \pi_2(x)) \ge \pi_i(x),$$

also

$$x = \pi(\pi_1(x), \pi_2(x)) < \pi(c, c) =: h'(c).$$

Sei jetzt

$$\begin{aligned} &\tilde{f}(\pi(0, b)) = g_1(b), \\ &\tilde{f}(\pi(a+1, b)) = g_2(a, \tilde{f}(\pi(a, h_1(a, b))), \dots, \tilde{f}(\pi(a, h_k(a, b))), b). \end{aligned}$$

Dann gilt offenbar $f(a, b) = \tilde{f}(\pi(a, b))$. Ferner hat man

$$h(\pi(a \dot{-} 1, h_\ell(a \dot{-} 1, b))) < h(\pi(a, b)) \quad \text{für } 1 \le \ell \le k \text{ und } a > 0,$$

denn es gilt

$$\begin{aligned} h(\pi(a \dot{-} 1, h_\ell(a \dot{-} 1, b))) &= h^*(a \dot{-} 1, a \dot{-} 1, h_\ell(a \dot{-} 1, b)) \\ &< h^*(a \dot{-} 1, a \dot{-} 1, \hat{h}(a \dot{-} 1, b)) \\ &= h^*(a, a \dot{-} 1, b) \\ &\le h^*(a, a, b) \\ &= h(\pi(a, b)). \end{aligned}$$

Für $a = 0$ und $b > 0$ kann man $t_j(\pi(0, b)) := 0$ setzen, denn dann ist $h(t_j(\pi(0, b))) = h(0) < h(\pi(0, b))$. Damit ergibt sich die Behauptung aus Lemma 3.1.7. $\square$

Wir zeigen schließlich noch, daß auch die Wertverlaufsrekursion mit Einsetzungen in Parameterstellen nicht aus den primitiv rekursiven Funktionen hinausführt.

Korollar 3.1.9. *Mit $g, h_1, \dots, h_k$ ist auch die durch*

$$f(a, b, \boldsymbol{c}) := g(a, \overline{f}(a, h_1(a, b, \boldsymbol{c}), \boldsymbol{c}), \dots, \overline{f}(a, h_k(a, b, \boldsymbol{c}), \boldsymbol{c}), b, \boldsymbol{c})$$

definierte Funktion f primitiv rekursiv.

Beweis. Es gilt

$$\begin{aligned}
&\overline{f}(0, b, \boldsymbol{c}) = 1,\\
&\overline{f}(a+1, b, \boldsymbol{c}) = \overline{f}(a, b, \boldsymbol{c}) \cdot p_a^{g(a, \overline{f}(a, h_1(a,b,\boldsymbol{c}), \boldsymbol{c}), \dots, \overline{f}(a, h_k(a,b,\boldsymbol{c}), \boldsymbol{c}), b, \boldsymbol{c})}.
\end{aligned}$$

Also ist mit g auch $\overline{f}$ primitiv rekursiv, also wegen $f(a, b, \boldsymbol{c}) = \exp(a, \overline{f}(a+1, b, \boldsymbol{c}))$ auch f. □

Jetzt können wir Lemma 3.1.7 wie folgt verschärfen.

Lemma 3.1.10. *(Primitive Rekursion nach einer Maßzahl). Es gelte*

$$\begin{aligned}
&f(0, \boldsymbol{b}) = g_1(\boldsymbol{b}),\\
&f(a, \boldsymbol{b}) = g_2(a, f(t_1(a, \boldsymbol{b}), \boldsymbol{b}), \dots, f(t_k(a, \boldsymbol{b}), \boldsymbol{b}), \boldsymbol{b}) \quad \textit{falls } a > 0,\\
&h(t_j(a, \boldsymbol{b}), \boldsymbol{b}) < h(a, \boldsymbol{b}) \quad \textit{für } 1 \le j \le k, \textit{ falls } a > 0.
\end{aligned}$$

Dann ist mit $g_1, g_2, t_1, \dots, t_k$ und h auch f primitiv rekursiv.

Beweis. Zur Vereinfachung lassen wir die Parameter $\boldsymbol{b}$ weg. Es gilt $f(a) = f^*(h(a), a)$ mit folgendem f^*:

$$f^*(c, a) := \begin{cases} g_2(a, f^*(h(t_1(a)), t_1(a)), \dots, f^*(h(t_k(a)), t_k(a))), & \text{falls } c = h(a) \text{ und } a > 0;\\ g_1, & \text{sonst.} \end{cases}$$

f^* ist primitiv rekursiv nach Korollar 3.1.9. □

3.2 Rekursive Funktionen, rekursiv aufzählbare Relationen

Aus den primitiv rekursiven Relationen erhält man durch Projektion die rekursiv aufzählbaren Relationen und daraus die rekursiven Funktionen als die Menge der Funktionen mit rekursiv aufzählbaren Graphen. Es wird sich zeigen, daß dieser Begriff einer rekursiven Funktion eine Präzisierung des intuitiven Begriffs der berechenbaren Funktion ist (Churchsche These).

Eine Relation $R \subseteq \mathbb{N}^n$ heißt *rekursiv aufzählbar*, wenn es eine primitiv rekursive Relation $Q \subseteq \mathbb{N}^{n+1}$ gibt mit

$$R = \{ (a_1, \dots, a_n) \mid \exists b\, (b, a_1, \dots, a_n) \in Q \}.$$

Insbesondere ist deshalb jede primitiv rekursive Relation rekursiv aufzählbar.

Lemma 3.2.1. *Eine nichtleere Menge $M \subseteq \mathbb{N}$ ist rekursiv aufzählbar genau dann, wenn es ein $f \in \mathsf{PR}^1$ gibt mit $M = \mathsf{ran}(f)$.*

Beweis. $\Longrightarrow$. Sei $M \subseteq \mathbb{N}$ rekursiv aufzählbar, also $M = \{ c \mid \exists b\, (b, c) \in Q \}$. Nach Annahme ist $M \neq \emptyset$, es gibt also ein $a_0 \in M$. Setze

$$f(a) := \begin{cases} \pi_2(a), & \text{falls } (\pi_1(a), \pi_2(a)) \in Q;\\ a_0, & \text{sonst.} \end{cases}$$

Dann ist $M = \mathsf{ran}(f)$, und nach Lemma 3.1.3 gilt $f \in \mathsf{PR}^1$.

$\Longleftarrow$. Sei $M = \mathsf{ran}(f)$ für ein $f \in \mathsf{PR}^1$. Dann ist $M = \{ c \mid \exists b\, f(b) = c \}$ und deshalb rekursiv aufzählbar. □

Unter denselben Annahmen, unter denen wir in Abschnitt 3.1 Δ_0-Formeln eingeführt hatten, definieren wir jetzt induktiv den Begriff einer Σ_1-*Formel* der Sprache PR.

- Jede PR-Primformel und jede negierte PR-Primformel ist eine Σ_1-Formel.
- Sind A und B Σ_1-Formeln, so auch $A \wedge B$ und $A \vee B$.
- Ist A eine Σ_1-Formel und t ein PR-Term mit $x \notin \mathsf{vars}(t)$, so ist $\forall x{<}t\, A$ eine Σ_1-Formel.
- Ist A eine Σ_1-Formel, so ist $\exists x A$ eine Σ_1-Formel.

Offenbar ist jede Δ_0-Formel logisch äquivalent zu einer Σ_1-Formel und jede rekursiv aufzählbare Relation definierbar durch eine Σ_1-Formel. Auch die Umkehrung ist richtig:

Lemma 3.2.2. *Ist A eine Σ_1-Formel mit $\mathsf{FV}(A) \subseteq \{x_1, \dots, x_n\}$, so ist die Relation*

$$\{\, (a_1, \dots, a_n) \in \mathbb{N}^n \mid \mathcal{N} \models A[a_1, \dots, a_n/x_1, \dots, x_n] \,\}$$

rekursiv aufzählbar.

Beweis. Eine Σ_1-Formel A heiße *strikt*, wenn sie von der Form $\exists x\, A$ mit einer Δ_0-Formel A ist. Wir konstruieren zu jeder Σ_1-Formel A eine strikte Σ_1-Formel A' mit $\mathsf{FV}(A) = \mathsf{FV}(A')$ und $\mathcal{N} \models \forall(A \leftrightarrow A')$, und zwar durch Rekursion über die Definition der Σ_1-Formeln. Daraus folgt offenbar die Behauptung. Ist A eine PR-Primformel oder negierte PR-Primformel, so setze man $A' := \exists z\, A$ mit $z \notin \mathsf{FV}(A)$. In den restlichen Fällen sei $A' = \exists x\, \tilde{A}$ und $B' = \exists y\, \tilde{B}$ mit Δ_0-Formeln $\tilde{A}$ und $\tilde{B}$. Wir setzen

$$\begin{aligned}
(A \vee B)' &:= \exists z. \tilde{A}[x := z] \vee \tilde{B}[y := z],\\
(A \wedge B)' &:= \exists z. \tilde{A}[x := \pi_1 z] \wedge \tilde{B}[y := \pi_2 z],\\
(\forall y{<}t\, A)' &:= \exists z \forall y{<}t\, \exists x{<}z\, \tilde{A},\\
(\exists y\, A)' &:= \exists z\, \tilde{A}[x, y := \pi_1 z, \pi_2 z]
\end{aligned}$$

wobei $z \notin \mathsf{FV}(\tilde{A}) \cup \mathsf{FV}(\tilde{B})$ bzw. $z \notin \mathsf{FV}(\tilde{A}) \cup \mathsf{vars}(t)$. □

Es folgt, daß die Menge aller rekursiv aufzählbaren Relationen abgeschlossen ist gegen $\cap, \cup$, beschränkte Allquantifikation, unbeschränkte Existenzquantifikation (oder Projektion) sowie gegen die Einsetzung von primitiv rekursiven Funktionen.

Eine Funktion $f \colon \mathbb{N}^n \to \mathbb{N}$ heißt *rekursiv*, wenn ihr Graph

$$G_f := \{\, (a_1, \dots, a_n, b) \mid f(a_1, \dots, a_n) = b \,\}$$

rekursiv aufzählbar ist. Eine Relation $R \subseteq \mathbb{N}^n$ heißt *rekursiv*, wenn ihre charakteristische Funktion 1_R rekursiv ist. Hieraus ergibt sich, daß jede primitiv rekursive Relation auch rekursiv ist.

Lemma 3.2.3. *Eine Relation $R \subseteq \mathbb{N}^n$ ist rekursiv genau dann, wenn R und ihr Komplement $\mathbb{N}^n \setminus R$ rekursiv aufzählbar sind.*

Beweis. $\Longrightarrow$. Sei R rekursiv, also 1_R rekursiv, also

$$G_{1_R} = \{\, (\boldsymbol{a}, 1) \mid \boldsymbol{a} \in R \,\} \cup \{\, (\boldsymbol{a}, 0) \mid \boldsymbol{a} \in \mathbb{N}^n \setminus R \,\}$$

rekursiv aufzählbar. Dann sind $R = \{\, \boldsymbol{a} \mid (\boldsymbol{a}, 1) \in G_{1_R} \,\}$ und $\mathbb{N}^n \setminus R = \{\, \boldsymbol{a} \mid (\boldsymbol{a}, 0) \in G_{1_R} \,\}$ rekursiv aufzählbar nach Lemma 3.2.2.

$\Longleftarrow$. Seien R und $\mathbb{N}^n \setminus R$ rekursiv aufzählbar. Dann gibt es Σ_1-Formeln A und A' mit

$$\begin{aligned}
R &= \{\, \boldsymbol{a} \mid \mathcal{N} \models A[\boldsymbol{a}] \,\},\\
\mathbb{N}^n \setminus R &= \{\, \boldsymbol{a} \mid \mathcal{N} \models A'[\boldsymbol{a}] \,\}.
\end{aligned}$$

Zu zeigen ist, daß die charakteristische Funktion 1_R rekursiv ist, also, daß ihr Graph G_{1_R} rekursiv aufzählbar ist. Es gilt

$$\begin{aligned}
G_{1_R} &= \{\, (\boldsymbol{a}, b) \mid 1_R(\boldsymbol{a}) = b \,\}\\
&= \{\, (\boldsymbol{a}, b) \mid (\boldsymbol{a} \in R \text{ und } b = 1) \text{ oder } (\boldsymbol{a} \in \mathbb{N}^n \setminus R \text{ und } b = 0) \,\}\\
&= \{\, (\boldsymbol{a}, b) \mid \mathcal{N} \models ((A \wedge y = 1) \vee (A' \wedge y = 0))[\boldsymbol{a}, b] \,\}.
\end{aligned}$$

Damit ist gezeigt, daß G_{1_R} durch eine Σ_1-Formel definierbar und deshalb rekursiv aufzählbar ist. □

Korollar 3.2.4. *Für jede rekursive Funktion f ist ihr Graph G_f rekursiv.*

Beweis. Sei $f\colon \mathbb{N}^n \to \mathbb{N}$ rekursiv. Es genügt zu zeigen, daß $\mathbb{N}^{n+1} \setminus G_f$ rekursiv aufzählbar ist. Dies folgt aber aus

$$\mathbb{N}^{n+1} \setminus G_f = \{\, (\boldsymbol{a}, b) \mid \exists c.c \neq b \wedge (\boldsymbol{a}, c) \in G_f \,\}. \qquad \square$$

Sei $\mathcal{F}$ die kleinste Menge von Funktionen mit den folgenden Eigenschaften. Wir schreiben $\mathcal{F}^n$ für die Teilmenge der n-stelligen Funktionen in $\mathcal{F}$.

1. $0^n \in \mathcal{F}^n$ $(n \geq 0)$, $\mathsf{S} \in \mathcal{F}^1$, $I_i^n \in \mathcal{F}^n$ $(1 \leq i \leq n)$.
2. Sind $g_1, \dots, g_m \in \mathcal{F}^n$, $h \in \mathcal{F}^m$ und $m \geq 1$, $n \geq 0$, so ist $(\circ h g_1 \dots g_m) \in \mathcal{F}^n$.
3. Ist $g \in \mathcal{F}^n$ und $h \in \mathcal{F}^{n+2}$, so ist $(Rgh) \in \mathcal{F}^{n+1}$.
4. Ist $g \in \mathcal{F}^{n+1}$ und gibt es zu jedem $\boldsymbol{a} \in \mathbb{N}^n$ ein $i \in \mathbb{N}$ mit $g(i, \boldsymbol{a}) = 0$, so ist die durch

   $$(\mu g)(\boldsymbol{a}) := \min\{\, i \mid g(i, \boldsymbol{a}) = 0 \,\}$$

   definierte Funktion μg in $\mathcal{F}^n$ (*unbeschränkter μ-Operator*). Wir schreiben meist $\mu i\, (g(i, \boldsymbol{a}) = 0)$ für $(\mu g)(\boldsymbol{a})$.

Die Funktionen in $\mathcal{F}$ nennen wir auch *μ-rekursiv*.

Satz 3.2.5. *$\mathcal{F}$ ist die Menge der rekursiven Funktionen.*

Beweis. Wir zeigen zunächst durch Induktion über die Definition von $\mathcal{F}$, daß jedes $f \in \mathcal{F}$ rekursiv ist. Dazu konstruieren wir zu jedem $f \in \mathcal{F}^n$ eine Σ_1-Formel $A_f[x_1, \dots, x_n, y]$ so daß $f(\boldsymbol{a}) = b$ genau dann, wenn $\mathcal{N} \models A_f[\boldsymbol{a}, b]$. Für die Ausgangsfunktionen ist dies trivial. *Fall* $f = (\circ h g_1 \dots g_m)$.

$$A_f := \exists z_1, \dots, z_m. \bigwedge_{i=1}^{m} A_{g_i}[\boldsymbol{x}, z_i] \wedge A_h[z_1, \dots, z_m, y].$$

Fall $f = (Rgh)$.

$$A_f[x, \boldsymbol{x}, y] := \exists z.\mathsf{lh}(z) = x + 1 \wedge A_g[\boldsymbol{x}, (z)_0] \wedge \forall i{<}x\, A_h[i, (z)_i, \boldsymbol{x}, (z)_{i+1}] \wedge y = (z)_x.$$

Fall $f = (\mu g)$.

$$A_f := A_g[y, \boldsymbol{x}, 0] \wedge \forall i{<}y \exists z\, A_g[i, \boldsymbol{x}, \mathsf{S}(z)].$$

Sei nun umgekehrt f rekursiv, also G_f rekursiv aufzählbar. Zu zeigen ist $f \in \mathcal{F}$. Nach dem Beweis von Lemma 3.2.2 ist G_f durch eine strikte Σ_1-Formel definierbar. Es gibt also eine Δ_0-Formel A so daß

$$G_f = \{\, (\boldsymbol{a}, b) \mid \mathcal{N} \models (\exists y\, A)[\boldsymbol{a}, b] \,\}.$$

Nach Lemma 3.1.2 ist

$$R := \{\, (i, \boldsymbol{a}) \mid \mathcal{N} \models A[\boldsymbol{a}, \pi_1(i), \pi_2(i)] \,\}$$

primitiv rekursiv. Es gibt also eine primitiv rekursive Funktion g mit $R = \{\, (i, \boldsymbol{a}) \mid g(i, \boldsymbol{a}) = 0 \,\}$. Nach Konstruktion gilt $\forall \boldsymbol{a} \exists i\, (g(i, \boldsymbol{a}) = 0)$ und $f(\boldsymbol{a}) = \pi_1(\mu i\, (g(i, \boldsymbol{a}) = 0))$. Also ist $f \in \mathcal{F}$, da $g, \pi_1 \in \mathsf{PR} \subseteq \mathcal{F}$. $\square$

Wir wollen jetzt noch den Zusammenhang zwischen den Begriffen einer rekursiven bzw. rekursiv aufzählbaren Relation und einer rekursiven Funktion einerseits und den intuitiven Begriffen der Entscheidbarkeit und der Aufzählbarkeit einer Relation und der Berechenbarkeit einer Funktion herstellen. Der Einfachheit halber beschränken wir uns wieder auf Funktionen und Relationen über den natürlichen Zahlen. Zunächst geben wir eine informale Definition der erwähnten intuitiven Begriffe, wobei wir auch *partielle Funktionen* zulassen, also Funktionen $f\colon \mathsf{dom}(f) \to \mathbb{N}$ mit $\mathsf{dom}(f) \subseteq \mathbb{N}^n$. Eine Funktion heißt *total*, wenn sie auf ganz $\mathbb{N}^n$ definiert ist.

Eine Relation $R \subseteq \mathbb{N}^n$ heißt *entscheidbar*, wenn es einen Algorithmus gibt, der für jedes Argumentetupel $\boldsymbol{a}$ terminiert und entscheidet, ob $\boldsymbol{a}$ zu R gehört oder nicht. Eine Relation $R \subseteq \mathbb{N}^n$ heißt *aufzählbar* (oder positiv berechenbar), wenn es einen Algorithmus gibt, der für jedes Argumentetupel $\boldsymbol{a}$ genau dann terminiert, wenn $\boldsymbol{a}$ zu R gehört. Eine Funktion $f\colon \mathsf{dom}(f) \to \mathbb{N}$ mit $\mathsf{dom}(f) \subseteq \mathbb{N}^n$ heißt *berechenbar*, wenn es einen Algorithmus gibt, der für jedes Argumentetupel $\boldsymbol{a}$ folgendes leistet.

1. Der Algorithmus terminiert auf $\boldsymbol{a}$ genau dann, wenn $\boldsymbol{a} \in \mathsf{dom}(f)$.
2. Für jedes $\boldsymbol{a} \in \mathsf{dom}(f)$ liefert der Algorithmus bei Eingabe von $\boldsymbol{a}$ den Wert $f(\boldsymbol{a})$.

Zunächst notieren wir einige einfache Eigenschaften entscheidbarer und aufzählbarer Relationen und berechenbarer Funktionen.

Lemma 3.2.6. *Eine Funktion f ist berechenbar genau dann, wenn ihr Graph G_f aufzählbar ist.*

Beweis. 1. Sei f berechenbar. Einen Algorithmus, der für jedes Argumentetupel $(\boldsymbol{a}, b)$ genau dann terminiert, wenn $(\boldsymbol{a}, b)$ zu G_f gehört, erhält man wie folgt. Man berechne $f(\boldsymbol{a})$ mittels des gegebenen Algorithmus für f. Man breche ab genau dann, wenn die Berechnung terminiert und den Wert b ergibt.

2. Sei G_f aufzählbar. Einen Algorithmus zur Berechnung von f erhält man wie folgt. Gegeben sei ein Argumentetupel $\boldsymbol{a}$. Man zähle alle Tupel $(\boldsymbol{c}, b)$ in G_f auf. Sobald dabei ein Tupel $(\boldsymbol{c}, b)$ mit $\boldsymbol{c} = \boldsymbol{a}$ auftaucht, breche man ab und gebe b aus. □

Lemma 3.2.7. *Eine Relation $R \subseteq \mathbb{N}^n$ ist entscheidbar genau dann, wenn sowohl sie selbst als auch ihr Komplement $\mathbb{N}^n \setminus R$ aufzählbar sind.*

Beweis. Übungsaufgabe. □

Das folgende Lemma gibt einige einfache Charakterisierungen des Begriffs der Aufzählbarkeit.

Lemma 3.2.8. *Für Relationen $R \subseteq \mathbb{N}^n$ sind folgende Bedingungen äquivalent.*

1. *R ist aufzählbar.*
2. *R ist Definitionsbereich einer berechenbaren Funktion.*
3. *R ist Projektion einer entscheidbaren Relation $Q \subseteq \mathbb{N}^{n+1}$.*

Beweis. Übungsaufgabe. □

Satz 3.2.9. *(CHURCHsche These).*

1. *Eine Relation $R \subseteq \mathbb{N}^n$ ist entscheidbar genau dann, wenn sie rekursiv ist.*
2. *Eine Relation $R \subseteq \mathbb{N}^n$ ist aufzählbar (oder positiv berechenbar) genau dann, wenn sie rekursiv aufzählbar ist.*
3. *Eine totale Funktion $f\colon \mathbb{N}^n \to \mathbb{N}$ ist berechenbar genau dann, wenn sie rekursiv ist.*

Die CHURCH*sche These* kann man aus prinzipiellen Gründen nicht mathematisch beweisen, da sie Begriffe verwendet, die nur intuitiv, aber nicht mathematisch präzise erklärt sind. Es besteht jedoch allgemeine Übereinstimmung darüber, daß die CHURCHsche These zutrifft. Diese Übereinstimmung beruht im wesentlichen auf Erfahrung: es ist keine im intuitiven Sinn berechenbare totale Funktion bekannt, die nicht rekursiv wäre.

3.3 Kodierung der Logik

In diesem Abschnitt wollen wir zur Vorbereitung unserer metamathematischen Studien die wesentlichen Begriffe der Syntax innerhalb der natürlichen Zahlen kodieren. Der Grund dafür ist, daß wir später im Rahmen einer formalen Theorie der Arithmetik über Beweise in einem formalen System sprechen wollen.

Wir hatten bisher der Einfachheit halber Formeln identifiziert, die sich nur durch gebundene Umbenennung unterscheiden. Diese Vorgehensweise ist für die Kodierung der Logik nicht mehr geeignet. Wir beginnen deshalb mit einer Rekapitulation des Begriffs der Substitution, wobei wir jetzt die konkreten Variablennamen berücksichtigen.

Eine *Substitution* ϑ ist eine endliche Menge der Form

$$\vartheta = \{t_1/x_1, \dots, t_n/x_n\},$$

so daß $t_i \neq x_i$ für $i = 1, \dots, n$ und $x_1, \dots, x_n$ paarweise verschieden sind. Ein Element t_i/x_i von ϑ wird als *Bindung* (von x_i an t_i) bezeichnet.

Ist r ein Term und ϑ die Substitution $\{t_1/x_1, \dots, t_n/x_n\}$, so verstehen wir unter $r\vartheta$ den Term, den wir erhalten, indem wir in r simultan jedes Vorkommen von x_i durch t_i ersetzen ($i = 1, \dots, n$). Wir sagen dann, daß ϑ auf r angewandt wurde und nennen $r\vartheta$ die von ϑ induzierte *Instanz* von r.

Ist ϑ eine Substitution $\{t_1/x_1, \dots, t_n/x_n\}$, so bezeichnen wir

$$\mathsf{dom}(\vartheta) := \{x_1, \dots, x_n\}$$

als den Bereich von ϑ und

$$\mathsf{ranv}(\vartheta) := \mathsf{vars}(t_1) \cup \ldots \cup \mathsf{vars}(t_n)$$

als die Wertevariablen von ϑ. (Wir schreiben nicht "ran" als Abkürzung, da dies zu Verwechslungen mit dem üblichen Wertebereich einer Funktion führen könnte). ϑ heißt *Grundsubstitution*, falls $\{t_1, \dots, t_n\}$ nur aus geschlossenen Termen besteht; ϑ heißt *Variablensubstitution*, falls $\{t_1, \dots, t_n\}$ nur aus Variablen besteht. Ist $\vartheta = \emptyset$, so sprechen wir von der leeren Substitution, die wir mit ε bezeichnen.

Zum Beispiel für $r = f(x, y, g(a))$ und $\vartheta = \{b/x, x/y\}$ mit Konstanten a, b ist $r\vartheta = f(b, x, g(a))$.

Sind r, t Terme und ist $\vartheta = \{t/x\}$ eine Substitution, so bezeichnen wir den Term $r\vartheta$ auch mit $r[x := t]$. Ferner setzen wir $r[x := x] := r$ (um später Fallunterscheidungen zu vermeiden).

Gegeben seien jetzt die Substitutionen

$$\vartheta = \{s_1/x_1, \dots, s_m/x_m\},$$
$$\sigma = \{t_1/y_1, \dots, t_n/y_n\}.$$

Dann ist die *Komposition* $\vartheta\sigma$ von ϑ und σ die Substitution, die wir erhalten, indem wir aus der Menge

$$\{s_1\sigma/x_1, \dots, s_m\sigma/x_m, t_1/y_1, \dots, t_n/y_n\}$$

alle Bindungen $s_i\sigma/x_i$ streichen, für die $s_i\sigma = x_i$ ist, sowie alle Bindungen t_j/y_j, für die $y_j \in \{x_1, \dots, x_m\}$.

Lemma 3.3.1. *Für Substitutionen ϑ, σ, τ gilt*

1. $\vartheta = \sigma \iff r\vartheta = r\sigma$ *für alle Terme* r.
2. $\vartheta = \sigma \iff x\vartheta = x\sigma$ *für alle Variablen* x.
3. $\vartheta\varepsilon = \varepsilon\vartheta = \vartheta$.
4. $(r\vartheta)\sigma = r(\vartheta\sigma)$ *für alle Terme* r.
5. $\vartheta(\sigma\tau) = (\vartheta\sigma)\tau$.

Beweis. (1)–(3) sind klar; (5) folgt unmittelbar aus (1) und (4). Wir zeigen nun (4). Sei also

$$\vartheta = \{s_1/x_1, \dots, s_m/x_m\} \quad \text{und} \quad \sigma = \{t_1/y_1, \dots, t_n/y_n\}.$$

Ist r eine Variable x, so unterscheiden wir drei Fälle:

Fall $x \notin \{x_1, \dots, x_m\} \cup \{y_1, \dots, y_n\}$. Dann ist $(x\vartheta)\sigma = x\sigma = x = x(\vartheta\sigma)$.

Fall $x \in \{x_1, \dots, x_m\}$, $x = x_i$. Dann ist $(x\vartheta)\sigma = s_i\sigma = x(\vartheta\sigma)$.

Fall $x \in \{y_1, \dots, y_n\} \setminus \{x_1, \dots, x_m\}$, $x = y_j$. Dann ist $(x\vartheta)\sigma = x\sigma = t_j = x(\vartheta\sigma)$.

Daraus folgt nun die Behauptung für beliebige Terme r durch eine triviale Induktion nach dem Termaufbau. □

Anmerkung. Aufgrund von (4) und (5) schreibt man statt $(r\vartheta)\sigma$ oder $r(\vartheta\sigma)$ kurz $r\vartheta\sigma$; ebenso schreibt man statt $\vartheta(\sigma\tau)$ oder $(\vartheta\sigma)\tau$ kurz $\vartheta\sigma\tau$.

Man kann leicht Beispiele einer Formel A und einer Substitution der Form $\vartheta = \{t_1/x_1, \dots, t_n/x_n\}$ finden, so daß bei der simultanen Ersetzung jedes freien Vorkommens von x_i durch t_i in A eine Variable aus t_i in den Bindungsbereich eines Quantors in A gerät. Um diese unerwünschte Erscheinung zu vermeiden, müssen wir gegebenenfalls einige gebundene Variablen in A umbenennen. Wir definieren deshalb zunächst die Menge $\mathsf{BV}(A)$ der in A *gebundenen Variablen*.

$$\begin{aligned} \mathsf{BV}(R(t_1, \dots, t_n)) &:= \emptyset. \\ \mathsf{BV}(A \wedge B) &:= \mathsf{BV}(A) \cup \mathsf{BV}(B). \\ \mathsf{BV}(A \to B) &:= \mathsf{BV}(A) \cup \mathsf{BV}(B). \\ \mathsf{BV}(\forall x A) &:= \mathsf{BV}(A) \cup \{x\}. \end{aligned}$$

Wir definieren jetzt $A\vartheta$ wie folgt durch Induktion über A.

$$R(t_1,\dots,t_n)\vartheta := R(t_1\vartheta,\dots,t_n\vartheta).$$
$$(A\wedge B)\vartheta := A\vartheta \wedge B\vartheta.$$
$$(A\to B)\vartheta := A\vartheta \to B\vartheta.$$

Im Fall $\forall x A$ bildet man zunächst $\sigma := \vartheta{\restriction}\mathsf{FV}(\forall xA)$ und setzt dann

$$(\forall xA)\vartheta := \begin{cases} \forall xA\sigma, & \text{falls } x\notin \mathsf{ranv}(\sigma) \\ \forall yA(\{y/x\}\sigma), & \text{falls } x\in \mathsf{ranv}(\sigma), \end{cases}$$

wobei y eine "neue" Variable ist, etwa die erste Variable echt oberhalb aller Variablen in $\mathsf{ranv}(\sigma)\cup\mathsf{FV}(A)\cup\mathsf{BV}(A)$. Man beachte, daß unter den angegebenen Voraussetzungen $\{y/x\}\sigma = \{y/x\}\cup\sigma$ ist. Diese Definition zeichnet sich dadurch aus, daß nur die unbedingt notwendigen Umbenennungen vorgenommen werden. – Ist $\vartheta = \{t/x\}$, so bezeichnen wir die Formel $A\vartheta$ auch mit $A[x:=t]$.

Wir beweisen jetzt einige einfache Eigenschaften von Substitutionen.

Lemma 3.3.2. *1. Gilt $x\vartheta = x\sigma$ für alle $x\in\mathsf{vars}(t)$, so ist $t\vartheta = t\sigma$.*
2. Gilt $x\vartheta = x\sigma$ für alle $x\in\mathsf{FV}(A)$, so ist $A\vartheta = A\sigma$.

Beweis. (1) ist klar. (2) beweisen wir durch Induktion über A. Wir behandeln nur den Fall $\forall xA$. Sei $\vartheta' := \vartheta{\restriction}\mathsf{FV}(\forall xA)$, $\sigma' := \sigma{\restriction}\mathsf{FV}(\forall xA)$. Nach Voraussetzung ist $\vartheta' = \sigma'$. Da $(\forall xA)\tau$ nur von x, A und $\tau{\restriction}\mathsf{FV}(\forall xA)$ abhängt, folgt $(\forall xA)\vartheta = (\forall xA)\sigma$. □

Lemma 3.3.3. *1. $\mathsf{vars}(t\vartheta)\subseteq[\mathsf{vars}(t)\setminus\mathsf{dom}(\vartheta)]\cup\mathsf{ranv}(\vartheta)$; ist $\mathsf{dom}(\vartheta)\subseteq\mathsf{vars}(t)$, so gilt $=$.*
2. $\mathsf{FV}(A\vartheta)\subseteq[\mathsf{FV}(A)\setminus\mathsf{dom}(\vartheta)]\cup\mathsf{ranv}(\vartheta)$; ist $\mathsf{dom}(\vartheta)\subseteq\mathsf{FV}(A)$, so gilt $=$.

Beweis. (1) ist klar. (2) beweisen wir durch Induktion über A. Wir behandeln nur den Fall $\forall xA$. Sei $\sigma := \vartheta{\restriction}\mathsf{FV}(\forall xA)$.

Unterfall $x\notin\mathsf{ranv}(\sigma)$. Dann hatten wir definiert $(\forall xA)\vartheta = \forall xA\sigma$. Man erhält

$$\begin{aligned}\mathsf{FV}(\forall xA)\vartheta &= \mathsf{FV}(\forall xA\sigma)\\ &= \mathsf{FV}(A\sigma)\setminus\{x\}\\ &= ([\mathsf{FV}(A)\setminus\mathsf{dom}(\sigma)]\cup\mathsf{ranv}(\sigma))\setminus\{x\} && \text{nach IH}\\ &= [(\mathsf{FV}(A)\setminus\{x\})\setminus\mathsf{dom}(\sigma)]\cup\mathsf{ranv}(\sigma) && \text{wegen } x\notin\mathsf{ranv}(\sigma)\\ &\subseteq [\mathsf{FV}(\forall xA)\setminus\mathsf{dom}(\vartheta)]\cup\mathsf{ranv}(\vartheta)\end{aligned}$$

Im Fall $\mathsf{dom}(\vartheta)\subseteq\mathsf{FV}(\forall xA)$ ist $\vartheta=\sigma$ und nach IH das $\subseteq$ ein $=$.

Unterfall $x\in\mathsf{ranv}(\sigma)$. Dann hatten wir definiert $(\forall xA)\vartheta = \forall xA(\{y/x\}\sigma)$ mit $y\notin\mathsf{ranv}(\sigma)\cup\mathsf{FV}(A)\cup\mathsf{BV}(A)$. Man erhält

$$\begin{aligned}\mathsf{FV}(\forall xA)\vartheta &= \mathsf{FV}(\forall yA(\{y/x\}\sigma))\\ &= \mathsf{FV}(A(\{y/x\}\sigma))\setminus\{y\}\\ &\subseteq \big([\mathsf{FV}(A)\setminus(\{x\}\cup\mathsf{dom}(\sigma))]\cup\{y\}\cup\mathsf{ranv}(\sigma)\big)\setminus\{y\} && \text{nach IH}\\ &= [(\mathsf{FV}(A)\setminus\{x\})\setminus\mathsf{dom}(\sigma)]\cup\mathsf{ranv}(\sigma) && \text{wegen } y\notin\mathsf{ranv}(\sigma)\cup\mathsf{FV}(A)\\ &\subseteq [\mathsf{FV}(\forall xA)\setminus\mathsf{dom}(\vartheta)]\cup\mathsf{ranv}(\vartheta).\end{aligned}$$

Im Fall $\mathsf{dom}(\vartheta)\subseteq\mathsf{FV}(\forall xA)$ ist wieder $\vartheta=\sigma$ und deshalb das zweite $\subseteq$ ein $=$. Auch das erste $\subseteq$ ist ein $=$, denn im Fall $x\in\mathsf{FV}(A)$ liefert die IH das $=$, und im Fall $x\notin\mathsf{FV}(A)$ hat man

$$\begin{aligned}\mathsf{FV}(A(\{y/x\}\sigma))\setminus\{y\} &= \mathsf{FV}(A\sigma)\setminus\{y\} && \text{nach Lemma 3.3.3(2)}\\ &= \big([\mathsf{FV}(A)\setminus\mathsf{dom}(\sigma)]\cup\mathsf{ranv}(\sigma)\big)\setminus\{y\} && \text{nach IH}\\ &= [\mathsf{FV}(A)\setminus\mathsf{dom}(\sigma)]\cup\mathsf{ranv}(\sigma) && \text{da } y\notin\mathsf{ranv}(\sigma)\cup\mathsf{FV}(A)\\ &= [\mathsf{FV}(\forall xA)\setminus\mathsf{dom}(\sigma)]\cup\mathsf{ranv}(\sigma).\end{aligned}$$

□

Kodierung

Gegeben sei eine abzählbare Sprache $\mathcal{L}$ erster Stufe. In injektiver Weise sei jedem $R \in \mathsf{Rel}_{\mathcal{L}}^{(n)}$ eine *Symbolnummer* $\mathsf{SN}(R)$ von der Form $\langle 1, n, i\rangle$ und jedem $f \in \mathsf{Fun}_{\mathcal{L}}^{(n)}$ eine Symbolnummer $\mathsf{SN}(f)$ von der Form $\langle 2, n, j\rangle$ zugeordnet. Wir nennen $\mathcal{L}$ *primitiv rekursiv präsentiert*, wenn

$$\mathsf{Symb}_{\mathcal{L}} := \bigcup_{n\in\mathbb{N}} \{\, \mathsf{SN}(R) \mid R \in \mathsf{Rel}_{\mathcal{L}}^{(n)} \,\} \cup \bigcup_{n\in\mathbb{N}} \{\, \mathsf{SN}(f) \mid f \in \mathsf{Fun}_{\mathcal{L}}^{(n)} \,\}$$

primitiv rekursiv ist. Insbesondere ist also jede Sprache mit endlich vielen Relations- und Funktionssymbolen primitiv rekursiv präsentiert. Für jede im folgenden betrachtete Sprache $\mathcal{L}$ nehmen wir an, daß sie primitiv rekursiv präsentiert ist.

Ferner sei $\mathsf{SN}(\wedge) := \langle 3, 1\rangle$, $\mathsf{SN}(\to) := \langle 3, 2\rangle$ und $\mathsf{SN}(\forall) := \langle 3, 3\rangle$. Schließlich ordnen wir der i-ten Variablen $*_i$ die Symbolnummer $\mathsf{SN}(*_i) := \langle 0, i\rangle$ zu.

Für jeden $\mathcal{L}$-Term t definieren wir rekursiv seine *Kodenummer* (oder GÖDEL*nummer*) $\ulcorner t \urcorner$ durch

$$\begin{aligned}
\ulcorner x \urcorner &:= \langle \mathsf{SN}(x)\rangle,\\
\ulcorner c \urcorner &:= \langle \mathsf{SN}(c)\rangle,\\
\ulcorner f t_1 \dots t_n \urcorner &:= \langle \mathsf{SN}(f), \ulcorner t_1 \urcorner, \dots, \ulcorner t_n \urcorner\rangle.
\end{aligned}$$

Ebenso definieren wir rekursiv für jede $\mathcal{L}$-Formel A ihre Kodenummer $\ulcorner A \urcorner$ durch

$$\begin{aligned}
\ulcorner R t_1 \dots t_n \urcorner &:= \langle \mathsf{SN}(R), \ulcorner t_1 \urcorner, \dots, \ulcorner t_n \urcorner\rangle,\\
\ulcorner A \wedge B \urcorner &:= \langle \mathsf{SN}(\wedge), \ulcorner A \urcorner, \ulcorner B \urcorner\rangle,\\
\ulcorner A \to B \urcorner &:= \langle \mathsf{SN}(\to), \ulcorner A \urcorner, \ulcorner B \urcorner\rangle,\\
\ulcorner \forall x\, A \urcorner &:= \langle \mathsf{SN}(\forall), \ulcorner x \urcorner, \ulcorner A \urcorner\rangle.
\end{aligned}$$

Sei $\mathsf{Var} := \ \mathit{set}\langle\langle 0, i\rangle\rangle i \in \mathbb{N}$. Var ist offenbar primitiv rekursiv, und es gilt $a \in \mathsf{Var}$ genau dann, wenn $a = \ulcorner x \urcorner$ für eine Variable x. Ferner definieren wir $\mathsf{Ter} \subseteq \mathbb{N}$ wie folgt durch Wertverlaufsrekursion.

$$a \in \mathsf{Ter} :\leftrightarrow a \in \mathsf{Var} \vee ((a)_0 \in \mathsf{Symb}_{\mathcal{L}} \wedge (a)_{0,0} = 2 \wedge \mathsf{lh}(a) = (a)_{0,1} + 1 \wedge \forall i_{0<i<\mathsf{lh}(a)}\, (a)_i \in \mathsf{Ter}).$$

Also ist Ter primitiv rekursiv, und es gilt $a \in \mathsf{Ter}$ genau dann, wenn $a = \ulcorner t \urcorner$ für einen Term t. Weiter sei $\mathsf{For} \subseteq \mathbb{N}$ definiert durch

$$\begin{aligned}
a \in \mathsf{For} :\leftrightarrow & ((a)_0 \in \mathsf{Symb}_{\mathcal{L}} \wedge (a)_{0,0} = 1 \wedge \mathsf{lh}(a) = (a)_{0,1} + 1 \wedge \forall i_{0<i<\mathsf{lh}(a)}\, (a)_i \in \mathsf{Ter})\\
& \vee (a = \langle \mathsf{SN}(\wedge), (a)_1, (a)_2\rangle \wedge (a)_1 \in \mathsf{For} \wedge (a)_2 \in \mathsf{For})\\
& \vee (a = \langle \mathsf{SN}(\to), (a)_1, (a)_2\rangle \wedge (a)_1 \in \mathsf{For} \wedge (a)_2 \in \mathsf{For})\\
& \vee (a = \langle \mathsf{SN}(\forall), (a)_1, (a)_2\rangle \wedge (a)_1 \in \mathsf{Var} \wedge (a)_2 \in \mathsf{For}).
\end{aligned}$$

Wieder ist For primitiv rekursiv, und es gilt $a \in \mathsf{For}$ genau dann, wenn $a = \ulcorner A \urcorner$ für eine Formel A. Für eine Menge S von Formeln setzen wir $\ulcorner S \urcorner := \{\, \ulcorner A \urcorner \mid A \in S \,\}$. Sei weiter $\mathsf{vars} \subseteq \mathbb{N} \times \mathbb{N}$ definiert durch

$$\mathsf{vars}(n, a) :\leftrightarrow (a \in \mathsf{Var} \wedge n = a) \vee ((a)_0 \in \mathsf{Symb}_{\mathcal{L}} \wedge \exists i_{0<i<\mathsf{lh}(a)}\, \mathsf{vars}(n, (a)_i)).$$

vars ist primitiv rekursiv, und es gilt $\mathsf{vars}(\ulcorner x \urcorner, \ulcorner t \urcorner)$ genau dann, wenn $x \in \mathsf{vars}(t)$. Entsprechend definieren wir

$$\begin{aligned}
\mathsf{FV}(n, a) :\leftrightarrow & ((a)_0 \in \mathsf{Symb}_{\mathcal{L}} \wedge \exists i_{0<i<\mathsf{lh}(a)}\, \mathsf{vars}(n, (a)_i))\\
& \vee ((a)_0 = \mathsf{SN}(\wedge) \wedge (\mathsf{FV}(n, (a)_1) \vee \mathsf{FV}(n, (a)_2)))\\
& \vee ((a)_0 = \mathsf{SN}(\to) \wedge (\mathsf{FV}(n, (a)_1) \vee \mathsf{FV}(n, (a)_2)))\\
& \vee ((a)_0 = \mathsf{SN}(\forall) \wedge n \neq (a)_1 \wedge \mathsf{FV}(n, (a)_2)).
\end{aligned}$$

Dann ist FV primitiv rekursiv, und es gilt $\mathsf{FV}(\ulcorner x \urcorner, \ulcorner A \urcorner)$ genau dann, wenn $x \in \mathsf{FV}(A)$.

Eine Substitution $\{t_1/x_1, \dots, t_n/x_n\}$ kann man kodieren durch die Zahl

$$\langle\langle\ulcorner x_1\urcorner,\ulcorner t_1\urcorner\rangle,\dots,\langle\ulcorner x_n\urcorner,\ulcorner t_n\urcorner\rangle\rangle$$

und natürlich auch durch viele andere Zahlen, da es auf die Reihenfolge der $\langle\ulcorner x_i\urcorner,\ulcorner t_i\urcorner\rangle$ nicht ankommt. Wir verwenden $\ulcorner\vartheta\urcorner$ zur Mitteilung einer solchen Kodenummer der Substitution ϑ. Mit $\mathsf{substval}(a,0) := a$ und im Fall $\ell > 0$

$$\mathsf{substval}(a,\ell) := \begin{cases} (\ell)_{0,1}, & \text{falls } a = (\ell)_{0,0}; \\ \mathsf{substval}(a,\mathsf{tl}(\ell)), & \text{sonst} \end{cases}$$

ist offenbar $\mathsf{substval}$ primitiv rekursiv und es gilt $\mathsf{substval}(\ulcorner x\urcorner,\ulcorner\vartheta\urcorner) = \ulcorner x\vartheta\urcorner$. Weiter sei $\mathsf{restrict}(0,a) := 0$ und im Fall $\ell > 0$

$$\mathsf{restrict}(\ell,a) := \begin{cases} \langle(\ell)_0\rangle * \mathsf{restrict}(\mathsf{tl}(\ell),a), & \text{falls } \mathsf{FV}((\ell)_{0,0},a); \\ \mathsf{restrict}(\mathsf{tl}(\ell),a), & \text{sonst.} \end{cases}$$

Dann ist $\mathsf{restrict}$ primitiv rekursiv und $\mathsf{restrict}(\ulcorner\vartheta\urcorner,\ulcorner A\urcorner)$ kodiert $\vartheta\restriction\mathsf{FV}(A)$. $\mathsf{Ranv}(n,\ell)$ definieren wir durch Rekursion über n; aus der Definition ergibt sich unmittelbar, daß Ranv primitiv rekursiv ist. $\mathsf{Ranv}(n,0)$ sei falsch und im Fall $\ell > 0$ setzen wir

$$\mathsf{Ranv}(n,\ell) :\leftrightarrow \mathsf{vars}(n,(\ell)_{0,1}) \lor \mathsf{Ranv}(n,\mathsf{tl}(\ell)).$$

Es gilt $\mathsf{Ranv}(\ulcorner x\urcorner,\ulcorner\vartheta\urcorner) \leftrightarrow x \in \mathsf{ranv}(\vartheta)$.

Zur Definition von $A\vartheta$ benötigen wir im $\forall$-Fall gegebenenfalls eine neue Variable. Dafür definieren wir $\mathsf{newvar}(0)$ beliebig und für $a > 0$

$$\mathsf{newvar}(a) := \begin{cases} \langle\langle 0,(a)_{0,1}+1\rangle\rangle, & \text{falls } a \in \mathsf{Var}; \\ \langle\langle 0,\max_{0<i<\mathsf{lh}(a)}(\mathsf{newvar}((a)_i))_{0,1}\rangle\rangle, & \text{falls } (a)_0 \in \mathsf{Symb}_{\mathcal{L}}; \\ \langle\langle 0,\max((\mathsf{newvar}((a)_1))_{0,1},(\mathsf{newvar}((a)_2))_{0,1})\rangle\rangle, & \text{falls } (a)_0 \in \{\mathsf{SN}(\land),\mathsf{SN}(\to)\}; \\ \langle\langle 0,\max((a)_{1,0,1}+1,(\mathsf{newvar}((a)_2))_{0,1})\rangle\rangle, & \text{falls } (a)_0 = \mathsf{SN}(\forall). \end{cases}$$

Dann ist newvar primitiv rekursiv, und $\mathsf{newvar}(\ulcorner t\urcorner)$ ist $\ulcorner y\urcorner$ für die kleinste Variable echt oberhalb aller Variablen in $\mathsf{vars}(t)$, und $\mathsf{newvar}(\ulcorner A\urcorner)$ ist $\ulcorner y\urcorner$ für die kleinste Variable echt oberhalb aller Variablen in $\mathsf{FV}(A)\cup\mathsf{BV}(A)$. Weiter sei $\mathsf{newranv}(0) := \langle\langle 0,0\rangle\rangle$ und für $\ell > 0$

$$\mathsf{newranv}(\ell) := \max(\mathsf{newvar}((\ell)_{0,1}),\mathsf{newranv}(\mathsf{tl}(\ell))).$$

Dann ist $\mathsf{newranv}$ primitiv rekursiv, und $\mathsf{newranv}(\ulcorner\vartheta\urcorner)$ ist die kleinste Variable echt oberhalb aller Wertevariablen von ϑ.

Wir definieren $\mathsf{sub}(a,\ell)$ wie folgt durch Rekursion über a. Im Fall $a \in \mathsf{Var}$ sei

$$\mathsf{sub}(a,\ell) := \mathsf{substval}(a,\ell).$$

Im Fall $(a)_0 = \mathsf{SN}(\forall)$ sei $k := \mathsf{restrict}(\ell,a)$ und

$$\mathsf{sub}(a,\ell) := \begin{cases} \langle\mathsf{SN}(\forall),(a)_1,\mathsf{sub}((a)_2,k)\rangle, & \text{falls } \neg\mathsf{Ranv}((a)_1,k); \\ \langle\mathsf{SN}(\forall),b,\mathsf{sub}((a)_2,\langle\langle(a)_1,b\rangle\rangle * k))\rangle, & \text{falls } \mathsf{Ranv}((a)_1,k), \end{cases}$$

wobei $b := \max(\mathsf{newranv}(k),\mathsf{newvar}((a)_2))$. Sonst sei

$$\mathsf{sub}(a,\ell) := \langle(a)_0,\mathsf{sub}((a)_1,\ell),\dots,\mathsf{sub}((a)_{\mathsf{lh}(a)-1},\ell)\rangle.$$

Lemma 3.3.4. *1.* $\mathsf{sub}(\ulcorner t\urcorner,\ulcorner\vartheta\urcorner) = \ulcorner t\vartheta\urcorner$.
2. $\mathsf{sub}(\ulcorner A\urcorner,\ulcorner\vartheta\urcorner) = \ulcorner A\vartheta\urcorner$.
3. Die Funktion sub *ist primitiv rekursiv.*

Beweis. 1. Induktion über t. 2. Induktion über A. Im Fall $\forall x\, A$ beachte man, daß in der obigen Definition von $(\forall x\, A)\vartheta$ die Komposition $\{y/x\}\sigma$ keine Streichung von Bindungen erfordert.

3. Es gilt $\mathsf{sub}(a,\ell) = g(a,\overline{\mathsf{sub}}(a,\ell),\overline{\mathsf{sub}}(a,\mathsf{restrict}(\ell,a)),\overline{\mathsf{sub}}(a,h_1(a,\ell)),\ell)$ mit

$$\begin{aligned}
b(a,\ell) &:= \max(\mathsf{newranv}(\mathsf{restrict}(\ell,a)), \mathsf{newvar}((a)_2)),\\
h_1(a,\ell) &:= \langle\langle (a)_1, b(a,\ell)\rangle\rangle * \mathsf{restrict}(\ell,a).
\end{aligned}$$

Hierbei ist

$$\begin{aligned}
&g(a,d_0,d_1,d_2,\ell)\\
&:= \begin{cases}
\mathsf{substval}(a,\ell), & \text{falls } a \in \mathsf{Var};\\
\langle \mathsf{SN}(\forall), (a)_1, \exp((a)_2, d_1)\rangle, & \text{falls } (a)_0 = \mathsf{SN}(\forall) \text{ und } \neg\mathsf{Ranv}((a)_1, \mathsf{restrict}(\ell,a));\\
\langle \mathsf{SN}(\forall), b(a,\ell), \exp((a)_2, d_2)\rangle, & \text{falls } (a)_0 = \mathsf{SN}(\forall) \text{ und } \mathsf{Ranv}((a)_1, \mathsf{restrict}(\ell,a));\\
h(a, d_0, \mathsf{lh}(a) \dot{-} 1), & \text{sonst.}
\end{cases}
\end{aligned}$$

Ferner ist h primitiv rekursiv definiert durch

$$\begin{aligned}
h(a,d,0) &:= \langle (a)_0\rangle,\\
h(a,d,n+1) &:= h(a,d,n) * \langle \exp((a)_{n+1}, d)\rangle.
\end{aligned}$$ □

In der bisherigen Darstellung von natürlichen Herleitungen findet man die an einem Knoten freien Annahmen, indem man den darüberliegenden Teil des Baums durchmustert. Mit mehr Schreibaufwand verbunden, aber hier und in anderen theoretischen Überlegungen manchmal nützlich ist eine alternative Darstellung, in der man an jedem Knoten noch die dort freien Annahmen hinzuschreibt. Dies hatten wir bereits in unserer “Sequenzenformulierung des natürlichen Schließens” in Abschnitt 1.2 getan. Für den gegenwärtigen Zweck der Kodierung von Herleitungen ist es bequem, diese Darstellung von Herleitungen noch einmal leicht abzuändern und die Annahmeformeln in der Form einer Multimenge mitzuführen.

Unter einer *Sequenz* $\Gamma \Rightarrow A$ verstehen wir ein Paar aus einer Multimenge $\Gamma = \{\{A_1,\dots,A_n\}\}$ von Formeln und einer Formel A. Wir definieren $\vdash_m \Gamma \Rightarrow A$ induktiv durch die folgenden Regeln. Eine Annahme kann man einführen durch

$$\Gamma \Rightarrow A, \quad \text{falls } A \text{ in } \Gamma.$$

Für die Konjuktion $\wedge$ haben wir eine Einführungsregel $\wedge$I und zwei Beseitigungsregeln $\wedge\mathrm{E_l}$ und $\wedge\mathrm{E_r}$.

$$\frac{\Gamma \Rightarrow A \quad \Delta \Rightarrow B}{\Gamma,\Delta \Rightarrow A\wedge B}\wedge\mathrm{I} \qquad \frac{\Gamma \Rightarrow A\wedge B}{\Gamma \Rightarrow A}\wedge\mathrm{E_r} \qquad \frac{\Gamma \Rightarrow A\wedge B}{\Gamma \Rightarrow B}\wedge\mathrm{E_l}$$

Γ,Δ bezeichnet hier die Multimengenvereinigung. Für die Implikation $\to$ gibt es eine Einführungsregel $\to$I (hier ohne Erwähnung einer Annahmevariablen u) und eine Beseitigungsregel $\to$E.

$$\frac{\Gamma \Rightarrow B}{\Delta \Rightarrow A \to B}\to\mathrm{I} \qquad \frac{\Gamma \Rightarrow A\to B \quad \Delta \Rightarrow A}{\Gamma,\Delta \Rightarrow B}\to\mathrm{E}$$

Dabei soll in $\to$I die Multimenge Δ aus Γ durch Streichen einiger Vorkommen von A entstehen. Für den Allquantor $\forall$ gibt es eine Einführungsregel $\forall$I und eine Beseitigungsregel $\forall$E, die wir hier ohne den zu substituierenden Term t als rechte Prämisse formulieren.

$$\frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow \forall x\, A}\forall\mathrm{I} \qquad \frac{\Gamma \Rightarrow \forall x\, A}{\Gamma \Rightarrow A[x:=t]}\forall\mathrm{E}$$

In $\forall$I muß wieder die Variablenbedingung erfüllt sein: für alle B in Γ muß gelten $x \notin \mathsf{FV}(B)$.

Anmerkung. In Kapitel 2 hatten wir die Schreibweise $M^B[u_1^{A_1},\dots,u_n^{A_n}]$ vereinbart um mitzuteilen, daß die freien Annahmen von M^B in der Liste $u_1^{A_1},\dots,u_n^{A_n}$ enthalten sind. Hierbei war selbstverständlich angenommen, daß im Fall $u_i = u_j$ auch die Formelindizes A_i und A_j gleich sind. Es ist für das folgende bequem, auch zuzulassen, daß in $u_1^{A_1},\dots,u_n^{A_n}$ ein $u_i^{A_i}$ mehrmals vorkommt.

Lemma 3.3.5. *1. Wenn* $\vdash_m \{\{A_1,\dots,A_n\}\} \Rightarrow A$, *so gibt es für alle (nicht notwendig verschiedenen)* $u_1,\dots,u_n$ *mit* $u_i = u_j \to A_i = A_j$ *einen Herleitungsterm* $M^A[u_1^{A_1},\dots,u_n^{A_n}]$.

2. *Zu jedem Herleitungsterm* $M^A[u_1^{A_1}, \dots, u_n^{A_n}]$ *kann man Vielfachheiten* $k_1, \dots, k_n \geq 0$ *angeben so daß* $\vdash_m \{\{A_1^{k_1}, \dots, A_n^{k_n}\}\} \Rightarrow A$*; hierbei bedeutet* A^k *ein k-faches Vorkommen von A.*

Beweis. 1. Gelte $\vdash_m \{\{A_1, \dots, A_n\}\} \Rightarrow A$. Wir führen den Beweis durch Induktion über $\vdash_m$. *Fall* Annahmeaxiom. Sei etwa $A = A_i$. Wähle $M = u_i$.

Fall →I. OBdA hat man

$$\frac{\{\{A_1, \dots, A_n, A, \dots, A\}\} \Rightarrow B}{\{\{A_1, \dots, A_n\}\} \Rightarrow A \to B} \to\mathrm{I}.$$

Seien $u_1, \dots, u_n$ gegeben mit $u_i = u_j \to A_i = A_j$. Man wähle ein neues u. Nach IH existiert ein M^B mit $\mathsf{FA}(M^B) \subseteq \{u_1^{A_1}, \dots, u_n^{A_n}, u^A\}$. Dann ist $(\lambda u^A\, M^B)^{A \to B}$ ein Herleitungsterm mit freien Annahmen unter $u_1^{A_1}, \dots, u_n^{A_n}$.

Fall →E. Gegeben sind Herleitungen von

$$\{\{A_1, \dots, A_n\}\} \Rightarrow A \to B \quad \text{und} \quad \{\{A_{n+1}, \dots, A_{n+m}\}\} \Rightarrow A.$$

Seien $u_1, \dots, u_{n+m}$ gegeben mit $u_i = u_j \to A_i = A_j$. Nach IH haben wir Herleitungsterme

$$M^{A \to B}[u_1^{A_1}, \dots, u_n^{A_n}] \quad \text{und} \quad N^A[u_{n+1}^{A_{n+1}}, \dots, u_{n+m}^{A_{n+m}}].$$

Dann ist aber auch

$$(MN)^B[u_1^{A_1}, \dots, u_{n+m}^{A_{n+m}}]$$

ein Herleitungsterm.

Die restlichen Fälle behandelt man ähnlich.

2. Gegeben sei ein Herleitungsterm $M^A[u_1^{A_1}, \dots, u_n^{A_n}]$. Wir führen den Beweis durch Induktion über M. *Fall* u^A. Dann ist $\vdash_m \{\{A\}\} \Rightarrow A$.

Fall →I, also $(\lambda u^A M^B)^{A \to B}$. Sei $\mathsf{FA}(M^B) \subseteq \{u_1^{A_1}, \dots, u_n^{A_n}, u^A\}$ mit $u_1, \dots, u_n, u$ verschieden. Nach IH hat man

$$\vdash_m \{\{A_1^{k_1}, \dots, A_n^{k_n}, A^k\}\} \Rightarrow B.$$

Mit der Regel →I folgt

$$\vdash_m \{\{A_1^{k_1}, \dots, A_n^{k_n}\}\} \Rightarrow A \to B.$$

Fall →E. Gegeben ist $(M^{A \to B} N^A)^B[u_1^{A_1}, \dots, u_n^{A_n}]$. Nach IH hat man

$$\vdash_m \{\{A_1^{k_1}, \dots, A_n^{k_n}\}\} \Rightarrow A \to B \quad \text{und} \quad \vdash_m \{\{A_1^{\ell_1}, \dots, A_n^{\ell_n}\}\} \Rightarrow A.$$

Mit der Regel →E folgt

$$\{\{A_1^{k_1+\ell_1}, \dots, A_n^{k_n+\ell_n}\}\} \Rightarrow B.$$

Die restlichen Fälle behandelt man ähnlich. □

Die Multimenge $\{\{A_1, \dots, A_n\}\}$ kann man durch $\langle \ulcorner A_1 \urcorner, \dots, \ulcorner A_n \urcorner \rangle$ kodieren. Wie bei Substitutionen hat auch hier dieselbe Multimenge i.a. mehrere Kodenummern. Wir verwenden $\ulcorner \Gamma \urcorner$ zur Mitteilung einer Kodenummer der Multimenge Γ. Die Verkettungsfunktion $*$ entspricht offenbar der Multimengenvereinigung. Zur Behandlung der →I-Regel benötigen wir noch die wie folgt definierten Funktionen msm (für "multisetminus") und msrm (für "multisetremove"). Es sei $\mathsf{msrm}(0, b) := 0$, und für $a > 0$ setzen wir

$$\mathsf{msrm}(a, b) := \begin{cases} \mathsf{tl}(a), & \text{falls } \mathsf{hd}(a) = b; \\ \langle \mathsf{hd}(a) \rangle * \mathsf{msrm}(\mathsf{tl}(a), b), & \text{sonst.} \end{cases}$$

Dann kodiert offenbar $\mathsf{msrm}(\ulcorner \Gamma \urcorner, \ulcorner A \urcorner)$ diejenige Multimenge, die aus Γ durch Entfernen eines Vorkommens von A entsteht, falls A in Γ vorkommt, und Γ sonst.

Weiter sei $\mathsf{msm}(a, 0) := a$, und für $b > 0$ sei $\mathsf{msm}(a, b) := \mathsf{msrm}(\mathsf{msm}(a, \mathsf{tl}(b)), \mathsf{hd}(b))$. $\mathsf{msm}(\ulcorner \Gamma \urcorner, \ulcorner \Delta \urcorner)$ kodiert dann die Multimenge, die aus Γ durch Entfernen aller Elemente aus Δ entsteht. Ferner sei die Relation $\mathsf{mseq}(a, b)$ definiert durch $\mathsf{msm}(a, b) = 0 \land \mathsf{msm}(b, a) = 0$. Offenbar gilt dann $\mathsf{mseq}(\ulcorner \Gamma \urcorner, \ulcorner \Delta \urcorner)$ genau dann, wenn Γ und Δ als Multimengen gleich sind. Wir schreiben kurz $a =_m b$ für $\mathsf{mseq}(a, b)$.

Eine Sequenz $\Gamma \Rightarrow A$ kodieren wir durch $\langle\ulcorner\Gamma\urcorner,\ulcorner A\urcorner\rangle$. Solche Kodifikate bezeichnen wir mit $\ulcorner\Gamma \Rightarrow A\urcorner$. Eine Herleitung für die induktive Definition von $\vdash_m$ kodieren wir durch $\langle\ulcorner\Gamma_0 \Rightarrow A_0\urcorner,\dots,\ulcorner\Gamma_n \Rightarrow A_n\urcorner\rangle$.

Wir definieren jetzt

$$
\begin{aligned}
&\mathsf{Herl}(d) :\leftrightarrow \forall i{<}\mathsf{lh}(d).\\
&\quad(\forall m{<}\mathsf{lh}((d)_{i,0})\,\mathsf{For}((d)_{i,0,m}) \wedge \exists n{<}\mathsf{lh}((d)_{i,0})\,((d)_{i,1} = (d)_{i,0,n})) && (\mathrm{Ann})\\
&\quad\vee (\exists j,k{<}i.(d)_{i,1} = \langle\mathsf{SN}(\wedge),(d)_{j,1},(d)_{k,1}\rangle \wedge (d)_{i,0} =_m (d)_{j,0} * (d)_{k,0}) && (\wedge\mathrm{I})\\
&\quad\vee (\exists j{<}i.(d)_{j,1} = \langle\mathsf{SN}(\wedge),(d)_{i,1},(d)_{j,1,2}\rangle \wedge (d)_{i,0} =_m (d)_{j,0}) && (\wedge\mathrm{E_r})\\
&\quad\vee (\exists j{<}i.(d)_{j,1} = \langle\mathsf{SN}(\wedge),(d)_{j,1,1},(d)_{i,1}\rangle \wedge (d)_{i,0} =_m (d)_{j,0}) && (\wedge\mathrm{E_l})\\
&\quad\vee (\exists j{<}i.(d)_{i,1} = \langle\mathsf{SN}(\to),(d)_{i,1,1},(d)_{j,1}\rangle \wedge \mathsf{For}((d)_{i,1,1}) && (\to\mathrm{I})\\
&\qquad\wedge \mathsf{msm}((d)_{i,0},(d)_{j,0}) = 0\\
&\qquad\wedge \forall n{<}\mathsf{lh}(\mathsf{msm}((d)_{j,0},(d)_{i,0}))\,((\mathsf{msm}((d)_{j,0},(d)_{i,0}))_n = (d)_{i,1,1}))\\
&\quad\vee (\exists j,k{<}i.(d)_{j,1} = \langle\mathsf{SN}(\to),(d)_{k,1},(d)_{i,1}\rangle \wedge (d)_{i,0} =_m (d)_{j,0} * (d)_{k,0}) && (\to\mathrm{E})\\
&\quad\vee (\exists j{<}i.(d)_{i,1} = \langle\mathsf{SN}(\forall),(d)_{i,1,1},(d)_{j,1}\rangle \wedge \mathsf{Var}((d)_{i,1,1}) \wedge (d)_{i,0} =_m (d)_{j,0} && (\forall\mathrm{I})\\
&\qquad\wedge \forall n{<}\mathsf{lh}((d)_{i,0})\,\neg\mathsf{FV}((d)_{i,1,1},(d)_{i,0,n}))\\
&\quad\vee (\exists j{<}i.(d)_{j,1,0} = \mathsf{SN}(\forall) \wedge (d)_{i,0} =_m (d)_{j,0} && (\forall\mathrm{E})\\
&\qquad\wedge ((d)_{i,1} = (d)_{j,1,2} \vee \exists n{<}(d)_{i,1}.\mathsf{Ter}(n) \wedge (d)_{i,1} = \mathsf{sub}((d)_{j,1,2},\langle\langle(d)_{j,1,1},n\rangle\rangle)))).
\end{aligned}
$$

Lemma 3.3.6. *1.* $\mathsf{Herl}(d)$ *gilt genau dann, wenn* d *eine Herleitung kodiert.*
2. Herl *ist primitiv rekursiv.*

Beweis. 1. $\Longrightarrow$. Gelte $\mathsf{Herl}(d)$. Man sieht dann leicht durch Induktion über i, daß für jedes $i < \mathsf{lh}(d)$ gilt $(d)_{i,0} = \ulcorner\Gamma\urcorner$ und $(d)_{i,1} = \ulcorner A\urcorner$ mit $\vdash_m \Gamma \Rightarrow A$.
$\Longleftarrow$. d kodiere eine Herleitung, etwa $d = \langle\ulcorner\Gamma_0 \Rightarrow A_0\urcorner,\dots,\ulcorner\Gamma_n \Rightarrow A_n\urcorner\rangle$. Durch Induktion über n und Fallunterscheidung nach der zuletzt angewandten Regel zeigt man leicht, daß dann $\mathsf{Herl}(d)$ gilt.

2. Die primitive Rekursivität von Herl folgt unmittelbar aus der Definition. □

Eine Menge S von Formeln heißt *rekursiv* (*primitiv rekursiv*, *rekursiv aufzählbar*), wenn $\ulcorner S\urcorner := \{\,\ulcorner A\urcorner \mid A \in S\,\}$ rekursiv (primitiv rekursiv, rekursiv aufzählbar) ist. Um auch die klassische Logik behandeln zu können, zeigen wir jetzt, daß die Menge $\mathsf{Stabax}_{\mathcal{L}}$ primitiv rekursiv ist, also die Menge der Formeln

$$\forall x_1,\dots,x_n.\neg\neg Rx_1\dots x_n \to Rx_1\dots x_n.$$

mit $R \in \mathsf{Rel}^{(n)}_{\mathcal{L}}$. Offenbar können wir hier annehmen, daß $x_1,\dots,x_n$ die ersten n Variablen $*_0,\dots,*_{n-1}$ sind. Wir definieren deshalb zunächst

$$
\begin{aligned}
\mathsf{app}(a,0) &:= \langle a\rangle,\\
\mathsf{app}(a,n+1) &:= \mathsf{app}(a,n) * \langle\langle\langle 0,n\rangle\rangle\rangle.
\end{aligned}
$$

Dann ist app primitiv rekursiv und wir haben $\mathsf{app}(\mathsf{SN}(R),n) = \ulcorner R*_0\dots*_{n-1}\urcorner$ und auch $\mathsf{app}(\mathsf{SN}(f),n) = \ulcorner f*_0\dots*_{n-1}\urcorner$. Weiter definieren wir

$$
\begin{aligned}
\mathsf{gen}(0,a) &:= a,\\
\mathsf{gen}(n+1,a) &:= \mathsf{gen}(n,\langle\mathsf{SN}(\forall),\langle\langle 0,n\rangle\rangle,a\rangle).
\end{aligned}
$$

Dann ist gen primitiv rekursiv und es gilt $\mathsf{gen}(n,\ulcorner A\urcorner) = \ulcorner\forall *_0\dots\forall *_{n-1}\ A\urcorner$. Setzt man noch $a \dot\to b := \langle\mathsf{SN}(\to),a,b\rangle$ und $\dot\neg a := a \dot\to \ulcorner\bot\urcorner$, so gilt

$$\ulcorner\mathsf{Stabax}_{\mathcal{L}}\urcorner(a) \leftrightarrow \exists k{<}a.\mathsf{Symb}_{\mathcal{L}}(k) \wedge (k)_0 = 1 \wedge a = \mathsf{gen}((k)_1,\dot\neg\dot\neg\mathsf{app}(k,(k)_1) \dot\to \mathsf{app}(k,(k)_1)).$$

Offenbar ist $\ulcorner\mathsf{Stabax}_{\mathcal{L}}\urcorner$ primitiv rekursiv.

Weiter zeigen wir, daß auch die Menge $\mathsf{Eq}_{\mathcal{L}}$ der $\mathcal{L}$-Gleichheitsaxiome primitiv rekursiv ist; hierbei setzen wir natürlich voraus, daß unsere primitiv rekursiv präsentierte Sprache $\mathcal{L}$ das Gleichheitssymbol $=$ enthält. Zum Beweis konstruiert man primitiv rekursive Hilfsfunktionen $\mathsf{app}_g, \mathsf{app}_u$ und h mit

$$\begin{aligned}
&\mathsf{app}_g(\mathsf{SN}(R),n) = \ulcorner R{*_0}{*_2}{*_4}\dots{*_{2n-2}}\urcorner, \qquad \mathsf{app}_g(\mathsf{SN}(f),n) = \ulcorner f{*_0}{*_2}{*_4}\dots{*_{2n-2}}\urcorner,\\
&\mathsf{app}_u(\mathsf{SN}(R),n) = \ulcorner R{*_1}{*_3}{*_5}\dots{*_{2n-1}}\urcorner, \qquad \mathsf{app}_u(\mathsf{SN}(f),n) = \ulcorner f{*_1}{*_3}{*_5}\dots{*_{2n-1}}\urcorner,\\
&h(n) = \ulcorner *_0 = *_1 \wedge *_2 = *_3 \wedge \dots \wedge *_{2n-2} = *_{2n-1}\urcorner.
\end{aligned}$$

Setzt man noch $a \doteq b := \langle \mathsf{SN}(=), a, b\rangle$ und $a \mathbin{\dot\wedge} b := \langle \mathsf{SN}(\wedge), a, b\rangle$, so gilt

$$\begin{aligned}
a \in \ulcorner \mathsf{Eq}_{\mathcal{L}}\urcorner \leftrightarrow a &= \ulcorner \forall *_0\,(*_0 = *_0)\urcorner\\
&\vee\ a = \ulcorner \forall *_0, *_1 . *_0 = *_1 \to *_1 = *_0\urcorner\\
&\vee\ a = \ulcorner \forall *_0, *_1, *_2 . *_0 = *_1 \wedge *_1 = *_2 \to *_0 = *_2\urcorner\\
&\vee\ \exists k, n{<}a.\mathsf{Symb}_{\mathcal{L}}(k) \wedge (k)_0 = 1 \wedge (k)_1 = n > 0\\
&\qquad \wedge\ a = \mathsf{gen}(2n, h(n) \mathbin{\dot\wedge} \mathsf{app}_g(k,n) \mathbin{\dot\to} \mathsf{app}_u(k,n))\\
&\vee\ \exists k, n{<}a.\mathsf{Symb}_{\mathcal{L}}(k) \wedge (k)_0 = 2 \wedge (k)_1 = n > 0\\
&\qquad \wedge\ a = \mathsf{gen}(2n, h(n) \mathbin{\dot\to} \mathsf{app}_g(k,n) \doteq \mathsf{app}_u(k,n)).
\end{aligned}$$

Offenbar ist $\ulcorner \mathsf{Eq}_{\mathcal{L}}\urcorner$ primitiv rekursiv.

Sei jetzt $\mathcal{L}$ eine primitiv rekursiv präsentierte Sprache mit $=$ in $\mathcal{L}$. Eine Theorie T mit $L(T) \subseteq \mathcal{L}$ heißt *(primitiv) rekursiv axiomatisierbar*, wenn es eine (primitiv) rekursive Menge S geschlossener $\mathcal{L}$-Formeln gibt so daß $T = \{\, A \in \overline{\mathcal{L}} \mid S \cup \mathsf{Eq}_{\mathcal{L}} \vdash_c A \,\}$.

Satz 3.3.7. *Für Theorien T mit $L(T) \subseteq \mathcal{L}$ sind die folgenden Aussagen äquivalent.*

1. *T ist rekursiv axiomatisierbar.*
2. *T ist primitiv rekursiv axiomatisierbar.*
3. *T ist rekursiv aufzählbar.*

Beweis. (3) $\Longrightarrow$ (2). Sei $\ulcorner T\urcorner$ rekursiv aufzählbar. Nach Lemma 3.2.1 existiert dann ein $f \in \mathsf{PR}^1$ mit $\ulcorner T\urcorner = \mathsf{ran}(f)$. Sei $f(n) = \ulcorner A_n\urcorner$. Wir definieren ein primitiv rekursives g mit $g(n) = \ulcorner A_0 \wedge \dots \wedge A_n\urcorner$ durch

$$\begin{aligned}
&g(0) := f(0),\\
&g(n+1) := g(n) \mathbin{\dot\wedge} f(n+1).
\end{aligned}$$

Für $S := \{\, A_0 \wedge \dots \wedge A_n \mid n \in \mathbb{N} \,\}$ ist $\ulcorner S\urcorner = \mathsf{ran}(g)$, und diese Menge ist primitiv rekursiv wegen $a \in \mathsf{ran}(g) \leftrightarrow \exists n{<}a\,(a = g(n))$. T ist also primitiv rekursiv axiomatisierbar, denn offenbar gilt $T = \{\, A \in \overline{\mathcal{L}} \mid S \cup \mathsf{Eq}_{\mathcal{L}} \vdash_c A \,\}$.

(2) $\Longrightarrow$ (1) ist klar.

(1) $\Longrightarrow$ (3). Sei T axiomatisiert durch S mit $\ulcorner S\urcorner$ rekursiv. Dann gilt

$$a \in \ulcorner T\urcorner \leftrightarrow \exists d \exists c{<}d.\mathsf{Herl}(d) \wedge (d)_{\mathsf{lh}(d)\dot-1} = \langle c, a\rangle \wedge \forall i{<}\mathsf{lh}(c)\,((c)_i \in \ulcorner \mathsf{Stabax}\urcorner \cup \ulcorner \mathsf{Eq}\urcorner \cup \ulcorner S\urcorner).$$

Also ist $\ulcorner T\urcorner$ rekursiv aufzählbar. □

Eine Theorie T in unserer primitiv rekursiv präsentierten Sprache $\mathcal{L}$ heißt *axiomatisiert*, wenn sie durch ein rekursiv aufzählbares Axiomensystem Ax_T gegeben ist. Nach dem eben bewiesenenen Satz können wir dann sogar annehmen, daß Ax_T primitiv rekursiv ist. Für solche axiomatisierten Theorien definieren wir $\mathsf{Abl}_T \subseteq \mathbb{N} \times \mathbb{N}$ durch

$$\mathsf{Abl}_T(d,a) :\leftrightarrow \mathsf{Herl}(d) \wedge \exists c{<}d.(d)_{\mathsf{lh}(d)\dot-1} = \langle c, a\rangle \wedge \forall i{<}\mathsf{lh}(c)\,((c)_i \in \ulcorner \mathsf{Stabax}\urcorner \cup \ulcorner \mathsf{Eq}\urcorner \cup \ulcorner \mathsf{Ax}_T\urcorner).$$

Offenbar ist Abl_T primitiv rekursiv und es gilt $\mathsf{Abl}_T(d,a)$ genau dann, wenn d eine Herleitung einer Sequenz $\Gamma \Rightarrow A$ kodiert mit $a = \ulcorner A\urcorner$ und Γ zusammengesetzt aus Stabilitätsaxiomen, Gleichheitsaxiomen und Formeln aus Ax_T.

Eine Theorie T heißt *konsistent*, wenn es eine geschlossene Formel A gibt mit $A \notin T$; andernfalls heißt T *inkonsistent*.

Korollar 3.3.8. *Jede axiomatisierte vollständige Theorie T ist rekursiv.*

Beweis. Ist T inkonsistent, so ist $\ulcorner T\urcorner$ rekursiv. Andernfalls folgt aus der Vollständigkeit von T

$$a \in \mathbb{N} \setminus \ulcorner T\urcorner \leftrightarrow a \notin \mathsf{For} \vee \exists b{<}a\,\mathsf{FV}(b,a) \vee \dot\neg a \in \ulcorner T\urcorner.$$

Also ist mit $\ulcorner T\urcorner$ auch $\mathbb{N} \setminus \ulcorner T\urcorner$ rekursiv aufzählbar und damit nach Lemma 3.2.3 $\ulcorner T\urcorner$ rekursiv. □

3.4 Herbrand-Gödel-Kleene-rekursive Funktionen

Wir geben in diesem Abschnitt eine weitere, auf Herbrand, Gödel und Kleene zurückgehende Charakterisierung der rekursiven Funktionen. Diese Charakterisierung verwendet logische Herleitungen aus Gleichungssystemen und stellt deshalb eine Verbindung zwischen der Logik und der Berechenbarkeitstheorie dar. Als Anwendung beweisen wir die Existenz einer rekursiv aufzählbaren, aber nicht rekursiven Relation.

Unter einem *Gleichungssystem* verstehen wir eine endliche Menge von generalisierten Gleichungen, also von Formeln der Gestalt $\forall x_1, \dots, x_n(t[x_1, \dots, x_n] = s[x_1, \dots, x_n])$. Wir setzen in diesem Abschnitt stets voraus, daß $0, \mathsf{S}$ in der Sprache vorhanden sind.

Eine Funktion $f\colon \mathbb{N}^n \to \mathbb{N}$ heißt Herbrand-Gödel-Kleene-*rekursiv* (oder kurz: *HGK-rekursiv*), wenn es ein Gleichungssystem E_f mit einem ausgezeichneten Funktionssymbol $\hat{f}$ gibt so daß für alle $a_1, \dots, a_n, b \in \mathbb{N}$ gilt

$$f(a_1, \dots, a_n) = b \quad \text{genau dann, wenn} \quad E_f \cup \mathsf{Eq} \vdash \hat{f}(\underline{a_1}, \dots, \underline{a_n}) = \underline{b}.$$

Satz 3.4.1. *Jede μ-rekursive Funktion ist HGK-rekursiv.*

Beweis. Durch Rekursion über die induktive Definition der μ-rekursiven Funktionen f konstruieren wir ein Gleichungssystem E_f mit einem ausgezeichneten Funktionssymbol $\hat{f}$ so daß

1. E_f nur die Funktionssymbole $0, \mathsf{S}, \hat{\cdot}, \hat{q}$ sowie $\hat{g}, \widehat{\prod g}$ für jede in der Definition von f vorkommende Funktion g enthält,
2. E_f in $\mathbb{N}$ gültig ist, wenn man 0 durch die Zahl 0, S durch die Nachfolgerfunktion, $\hat{\cdot}$ durch die Multiplikation, $\hat{q}$ durch $q(\mathsf{S}(a), 0, b) := b$ und beliebig sonst, jedes $\hat{g}$ durch g und jedes $\widehat{\prod g}$ durch $b, \boldsymbol{a} \mapsto \prod_{c<b} g(c, \boldsymbol{a})$ interpretiert, und
3. für alle $a_1, \dots, a_n, b \in \mathbb{N}$ gilt
$$f(a_1, \dots, a_n) = b \quad \text{genau dann, wenn} \quad E_f \cup \mathsf{Eq} \vdash \hat{f}(\underline{a_1}, \dots, \underline{a_n}) = \underline{b}.$$

Die Gültigkeit von (1)-(3) ergibt sich jeweils unmittelbar aus der im folgenden angegebenen Konstruktion, wobei man zum Beweis von (3)$\Leftarrow$ verwendet, daß (2) zutrifft. Bei Gleichungen lassen wir der Kürze halber die Generalisierungen immer weg.

Fall 0^n. Wähle $\widehat{0^n}(x_1, \dots, x_n) = 0$.

Fall S. Wähle $\widehat{\mathsf{S}}(x) = \mathsf{S}(x)$.

Fall I_i^n. Wähle $\widehat{I_i^n}(x_1, \dots, x_n) = x_i$.

Fall $(\circ h g_1 \dots g_m)$. Nach IH haben wir $E_h, E_{g_1}, \dots, E_{g_m}$ mit ausgezeichneten Funktionssymbolen $\hat{h}, \widehat{g_1}, \dots, \widehat{g_m}$. Wähle $E_h \cup E_{g_1} \cup \dots \cup E_{g_m}$ zusammen mit der Generalisierung von

$$\hat{f}(x_1, \dots, x_n) = \hat{h}(\widehat{g_1}(x_1, \dots, x_n), \dots, \widehat{g_m}(x_1, \dots, x_n)).$$

Fall (Rgh). Nach IH haben wir E_g, E_h mit ausgezeichneten Funktionssymbolen $\hat{g}, \hat{h}$. Wähle $E_g \cup E_h$ zusammen mit den Generalisierungen von

$$\begin{aligned} &\hat{f}(0, y_1, \dots, y_n) = \hat{g}(y_1, \dots, y_n), \\ &\hat{f}(\mathsf{S}(x), y_1, \dots, y_n) = \hat{h}(x, \hat{f}(x, y_1, \dots, y_n), y_1, \dots, y_n). \end{aligned}$$

Fall (μg). Nach IH haben wir E_g mit ausgezeichnetem Funktionssymbol $\hat{g}$. Wähle E_g zusammen mit den Generalisierungen von

$$\begin{aligned} &\widehat{(\prod g)}(0, x_1, \dots, x_n) = 1, \\ &\widehat{(\prod g)}(\mathsf{S}(y), x_1, \dots, x_n) = \hat{\cdot}(\widehat{(\prod g)}(y, x_1, \dots, x_n), \hat{g}(y, x_1, \dots, x_n)), \\ &\hat{q}(\mathsf{S}(x), 0, y) = y, \\ &\hat{f}(x_1, \dots, x_n) = \hat{q}(\widehat{(\prod g)}(y, x_1, \dots, x_n), \widehat{(\prod g)}(\mathsf{S}(y), x_1, \dots, x_n), y). \end{aligned}$$ □

Zum Beweis der Umkehrung definieren wir für jedes n eine primitiv rekursive Relation $T_n \subseteq \mathbb{N}^{n+2}$ so daß für jedes Gleichungssystem E mit ausgezeichnetem Funktionssymbol $\hat{f}$ mit Symbolnummer $\mathsf{SN}(\hat{f}) = \langle 2, n, 0\rangle$ gilt $T_n(\ulcorner E \urcorner, a_1, \ldots, a_n, d)$ genau dann, wenn d eine Herleitung einer Sequenz $\Gamma \Rightarrow \hat{f}(\underline{a_1}, \ldots, \underline{a_n}) = \underline{b}$ kodiert, wobei Γ aus Stabilitätsaxiomen, Gleichheitsaxiomen und Elementen aus E besteht. Ferner definieren wir eine primitiv rekursive Funktion $U\colon \mathbb{N} \to \mathbb{N}$, die aus einem solchen d die Zahl b abliest. Wir setzen

$$
\begin{aligned}
T_n(e, a_1, \ldots, a_n, d) :\leftrightarrow \mathsf{Herl}(d) \wedge \exists c, a, b{<}d((d)_{\mathsf{lh}(d)\dot{-}1} &= \langle c, a\rangle \\
&\wedge a = \langle \mathsf{SN}(=), \langle\langle 2, n, 0\rangle, \ulcorner \underline{a_1} \urcorner, \ldots, \ulcorner \underline{a_n} \urcorner\rangle, \ulcorner \underline{b} \urcorner\rangle \\
&\wedge \forall i{<}\mathsf{lh}(c).(c)_i \in \ulcorner \mathsf{Stabax} \urcorner \cup \ulcorner \mathsf{Eq} \urcorner \vee \exists j{<}\mathsf{lh}(e).(c)_i = (e)_j)
\end{aligned}
$$

und $U(d) := \mathsf{dec}((d)_{\mathsf{lh}(d)\dot{-}1,1,2})$, wobei dec eine primitiv rekursive Funktion ist mit $\mathsf{dec}(\ulcorner \underline{b} \urcorner) = b$; dec definieren wir durch $\mathsf{dec}(0) := 0$ und für $a > 0$

$$
\mathsf{dec}(a) := \begin{cases} \mathsf{dec}((a)_1) + 1, & \text{falls } (a)_0 = \mathsf{SN}(\mathsf{S}); \\ 0, & \text{sonst.} \end{cases}
$$

Nach Konstruktion sind T_n und U primitiv rekursiv und haben die verlangten Eigenschaften. Deshalb gilt offenbar folgender Satz.

Satz 3.4.2. *(*Kleene*sches Normalformentheorem). Zu jeder HGK-rekursiven n-stelligen Funktion f gibt es ein $e \in \mathbb{N}$ so daß folgendes gilt.*

1. $\forall a_1, \ldots, a_n \exists d\, T_n(e, a_1, \ldots, a_n, d)$
2. $f(a_1, \ldots, a_n) = U(\mu d\, T_n(e, a_1, \ldots, a_n, d))$ □

Insbesondere ergibt sich hieraus, daß jede μ-rekursive Funktion mit einem geeigneten e in der angegebenen Form definiert werden kann, wobei der μ-Operator nur einmal angewandt wird. Dies erklärt den Namen "Normalformentheorem".

Korollar 3.4.3. *Jede HGK-rekursive Funktion ist μ-rekursiv.* □

Wir kommen nun noch einmal auf den Fall (μg) mit g μ-rekursiv zurück, aber jetzt ohne vorauszusetzen, daß $\forall \boldsymbol{a} \exists i\, g(i, \boldsymbol{a}) = 0$. Sei E das oben im Fall (μg) angegebene Gleichungssystem. Dann gilt: (1) Wenn $\exists i\, g(i, \boldsymbol{a}) = 0$, so ist $E \cup \mathsf{Eq} \vdash \hat{f}(\underline{a_1}, \ldots, \underline{a_n}) = \underline{i}$ für das kleinste i mit $g(i, \boldsymbol{a}) = 0$. (2) Gilt $E \cup \mathsf{Eq} \vdash \hat{f}(\underline{a_1}, \ldots, \underline{a_n}) = \underline{i}$, so ist $g(i, \boldsymbol{a}) = 0$ und für $j < i$ ist $g(j, \boldsymbol{a}) \neq 0$. Daraus ergibt sich aufgrund der Definition von T_n:

Satz 3.4.4. *(*Kleene*sches Aufzählungstheorem). Zu jeder rekursiven Relation $R \subseteq \mathbb{N}^{n+1}$ findet man eine Zahl e so daß für alle $a_1, \ldots, a_n \in \mathbb{N}$ gilt*

$$
\exists b\, R(b, a_1, \ldots, a_n) \leftrightarrow \exists d\, T_n(e, a_1, \ldots, a_n, d). \qquad \square
$$

Die Relationen $\{ (a_1, \ldots, a_n) \mid \exists d\, T_n(e, a_1, \ldots, a_n, d) \}$ durchlaufen also für $e = 0, 1, \ldots$ genau die n-stelligen rekursiv aufzählbaren Relationen. Sei jetzt $n = 1$. Mit dem Cantorschen Diagonalargument folgt dann, daß $\{ a \mid \neg \exists d\, T_1(a, a, d) \}$ *nicht* rekursiv aufzählbar sein kann. Andernfalls gäbe es nämlich ein e so daß für alle a gilt

$$
\neg \exists d\, T_1(a, a, d) \leftrightarrow \exists d\, T_1(e, a, d),
$$

und durch Einsetzung von e für a erhielte man einen Widerspruch. Da nach Lemma 3.2.3 eine rekursiv aufzählbare Relation genau dann rekursiv ist, wenn ihr Komplement rekursiv aufzählbar ist, erhalten wir:

Satz 3.4.5. $\{ a \mid \exists d\, T_1(a, a, d) \}$ *ist rekursiv aufzählbar, aber nicht rekursiv.* □

Setzt man die Churchsche These voraus, so haben wir damit ein erstes Beispiel eines *unentscheidbaren Problems* gefunden: Es kann keinen Algorithmus geben, der für ein beliebiges $a \in \mathbb{N}$ stets terminiert und entscheidet, ob es ein $d \in \mathbb{N}$ gibt so daß die primitiv rekursive Relation $T_1(a, a, d)$ besteht.

3.5 Anmerkungen

Primitiv rekursive Funktionen wurden von Hilbert eingeführt. Ackermann zeigte 1925 durch Angabe eines Beispiels, daß es eine berechenbare, aber nicht primitiv rekursive Funktion gibt. Eine ausführliche Diskussion der primitiv rekursiven Funktionen findet man in den Büchern von Pèter [25] und Rose [27]. Der Begriff der rekursiven Funktion wurde nach Vorarbeiten von Herbrand, Gödel und Kleene von Church geprägt, der mit seiner Churchschen These die Äuivalenz zum intuitiven Begriff der Berechenbarkeit postulierte. Die hier gegebene Darstellung der Theorie der Berechenbarkeit stützt sich teilweise auf die Bücher von Shoenfield [34] und Kleene [16].

4. Metamathematik

4.1 Undefinierbarkeit des Wahrheitsbegriffs

Sei $\mathcal{M}$ eine $\mathcal{L}$-Struktur. Eine Relation $R \subseteq |\mathcal{M}|^n$ heißt in $\mathcal{M}$ *definierbar*, wenn es eine $\mathcal{L}$-Formel $A[x_1, \ldots, x_n]$ gibt so daß

$$R = \{\, (a_1, \ldots, a_n) \in |\mathcal{M}|^n \mid \mathcal{M} \models A[a_1, \ldots, a_n/x_1, \ldots, x_n] \,\}.$$

Wir wollen in diesem Abschnitt annehmen, daß $|\mathcal{M}| = \mathbb{N}$ und $0 \in \mathsf{Fun}_{\mathcal{L}}^{(0)}$, $\mathsf{S} \in \mathsf{Fun}_{\mathcal{L}}^{(1)}$ mit $0^{\mathcal{M}} = 0$ und $\mathsf{S}^{\mathcal{M}}(a) = a + 1$. Dann kann man für jedes $a \in \mathbb{N}$ das *Numeral* $\underline{a} \in \mathsf{Ter}_{\mathcal{L}}$ definieren durch $\underline{0} := 0$ und $\underline{n+1} := \mathsf{S}(\underline{n})$. Man beachte auch, daß in diesem Fall die Definierbarkeit von $R \subseteq \mathbb{N}^n$ durch $A[x_1, \ldots, x_n]$ äquivalent ist zu

$$R = \{\, (a_1, \ldots, a_n) \in \mathbb{N}^n \mid \mathcal{M} \models A[x_1, \ldots, x_n := \underline{a_1}, \ldots, \underline{a_n}] \,\}.$$

Sei weiter $\mathcal{L}$ eine primitiv rekursiv präsentierte Sprache. Wir werden in diesem Kapitel stets voraussetzen, jede primitiv rekursive Relation in $\mathcal{M}$ definierbar ist. Ein Beispiel für eine solche Situation ist die in Abschnitt 3.1 eingeführte Standardstruktur $\mathcal{N}$ zur Sprache PR. Eine Menge S von Formeln heißt *definierbar* in $\mathcal{M}$, wenn $\ulcorner S \urcorner := \{\, \ulcorner A \urcorner \mid A \in S \,\}$ in $\mathcal{M}$ definierbar ist.

Wir zeigen, daß bereits aus diesen recht schwachen Voraussetzungen folgt, daß der *Wahrheitsbegriff* von $\mathcal{M}$, genauer die Menge $\mathsf{Th}(\mathcal{M})$ aller in $\mathcal{M}$ gültigen geschlossenen Formeln, nicht in $\mathcal{M}$ definierbar ist. Da aus der vorausgesetzten Definierbarkeit aller primitiv rekursiven Relationen in $\mathcal{M}$ sofort folgt, daß auch alle rekursiven aufzählbaren Relationen in $\mathcal{M}$ definierbar sind, erhält man hieraus mit der Churchschen These, daß der Wahrheitsbegriff von $\mathcal{M}$ nicht aufzählbar sein kann und folglich insbesondere unentscheidbar ist.

Zum Beweis benötigen wir das folgende Fixpunktlemma, das wir im nächsten Abschnitt (als Lemma 4.2.2) noch verallgemeinern werden. Wir schreiben hier und im folgenden oft $A[\boldsymbol{t}]$ für $A[\boldsymbol{x} := \boldsymbol{t}]$, wenn die zu substituierenden Variablen aus dem Zusammenhang klar sind.

Lemma 4.1.1. *(Semantisches Fixpunktlemma). Ist jede primitiv rekursive Relation in $\mathcal{M}$ definierbar, so findet man zu jeder $\mathcal{L}$-Formel $B[z]$ eine geschlossene $\mathcal{L}$-Formel A mit*

$$\mathcal{M} \models A \quad \textit{genau dann, wenn} \quad \mathcal{M} \models B[\ulcorner A \urcorner].$$

Beweis. Wir definieren eine primitiv rekursive Funktion s durch

$$s(b, k) := \mathsf{sub}(b, \langle\langle \ulcorner z \urcorner, \ulcorner \underline{k} \urcorner \rangle\rangle).$$

z ist hierbei die spezielle, durch $B[z]$ vorgegebene Variable, etwa $*_0$. Dann gilt für jede Formel $C[z]$

$$s(\ulcorner C \urcorner, k) = \mathsf{sub}(\ulcorner C \urcorner, \langle\langle \ulcorner z \urcorner, \ulcorner \underline{k} \urcorner \rangle\rangle) = \ulcorner C[\underline{k}] \urcorner,$$

also insbesondere

$$s(\ulcorner C \urcorner, \ulcorner C \urcorner) = \ulcorner C[\underline{\ulcorner C \urcorner}] \urcorner.$$

Nach Annahme ist der Graph G_s von s definierbar in $\mathcal{M}$, etwa durch $A_s[x_1, x_2, x_3]$. Setze

$$\begin{aligned} C &:= \exists x. B[x] \wedge A_s[z, z, x], \\ A &:= C[\underline{\ulcorner C \urcorner}], \end{aligned}$$

also

$$A = \exists x.B[x] \wedge A_s[\underline{\ulcorner C \urcorner}, \underline{\ulcorner C \urcorner}, x].$$

Damit gilt $\mathcal{M} \models A$ genau dann, wenn $\exists a{\in}\mathbb{N}.\mathcal{M} \models B[a]$ und $a = \ulcorner C[\underline{\ulcorner C \urcorner}]\urcorner$, also genau dann, wenn $\mathcal{M} \models B[\ulcorner A \urcorner]$. □

Satz 4.1.2. *(*TARSKI*s Undefinierbarkeitssatz). Ist jede primitiv rekursive Relation in $\mathcal{M}$ definierbar, so ist $\mathsf{Th}(\mathcal{M})$ in $\mathcal{M}$ undefinierbar, insbesondere also nicht rekursiv aufzählbar.*

Beweis. Nehmen wir an, $\ulcorner \mathsf{Th}(\mathcal{M}) \urcorner$ wäre definierbar durch $B_W[z]$. Dann gilt für alle geschlossenen Formeln A

$$\mathcal{M} \models A \quad \text{genau dann, wenn} \quad \mathcal{M} \models B_W[\ulcorner A \urcorner].$$

Wir betrachten nun die Formel $\neg B_W[z]$ und wählen uns nach dem Fixpunktlemma 4.1.1 eine geschlossene $\mathcal{L}$-Formel A mit

$$\mathcal{M} \models A \quad \text{genau dann, wenn} \quad \mathcal{M} \models \neg B_W[\ulcorner A \urcorner].$$

Dies widerspricht der obigen Äquivalenz.

Bereits oben hatten wir bemerkt, daß alle rekursiv aufzählbaren Relationen in $\mathcal{M}$ definierbar sind; deshalb folgt, daß $\ulcorner \mathsf{Th}(\mathcal{M}) \urcorner$ nicht rekursiv aufzählbar ist. □

4.2 Der Wahrheitsbegriff in formalen Theorien

Wir wollen die Überlegungen des vorgehenden Abschnitts verallgemeinern. Dort hatten wir mit dem Wahrheitsbegriff in einer Struktur $\mathcal{M}$ gearbeitet, also der Relation $\mathcal{M} \models A$. Die Menge der Sätze A mit $\mathcal{M} \models A$ hatten wir als Theorie vom $\mathcal{M}$ bezeichnet, geschrieben $\mathsf{Th}(\mathcal{M})$.

Jetzt gehen wir statt von $\mathsf{Th}(\mathcal{M})$ allgemeiner von einer beliebigen Theorie T aus, und stellen uns die Frage, ob in T ein "Wahrheitsbegriff" (in der Gestalt einer "Wahrheitsformel" $B[z]$) existiert, so daß $B[z]$ "bedeutet", daß z "wahr" ist.

Was soll das heißen? Wir müssen die verwendeten Begriffe ohne semantische Konzepte erkären.

– z läuft also über geschlossene Formeln oder Sätze A bzw. genauer $\ulcorner A \urcorner$.
– A "wahr" ist zu ersetzen durch $T \vdash A$.
– C ist "gleichbedeutend" mit D ist zu ersetzen durch $T \vdash C \leftrightarrow D$.

Wir wollen also untersuchen, ob es eine Wahrheitsformel $B[z]$ geben kann, so daß für alle Sätze A gilt $T \vdash A \leftrightarrow B[\underline{\ulcorner A \urcorner}]$. Es wird sich zeigen, daß dies schon unter recht schwachen Voraussetzungen an die Theorie T unmöglich ist.

Technisch wird es darum gehen, an die Stelle der Definierbarkeit in $\mathcal{M}$ die "Repräsentierbarkeit" innerhalb einer formalen Theorie setzen.

Sei im folgenden $\mathcal{L}$ wieder eine primitiv rekursiv präsentierte Sprache mit $0, \mathsf{S}, =$ in $\mathcal{L}$ und T eine Theorie mit $\mathsf{Eq}_{\mathcal{L}} \subseteq T$. Wir nennen eine Relation $R \subseteq \mathbb{N}^n$ *repräsentierbar* in T, wenn es eine Formel $A[x_1, \ldots, x_n]$ gibt mit

$$\begin{aligned} &T \vdash A[\underline{a_1}, \ldots, \underline{a_n}], && \text{falls } (a_1, \ldots, a_n) \in R, \\ &T \vdash \neg A[\underline{a_1}, \ldots, \underline{a_n}], && \text{falls } (a_1, \ldots, a_n) \notin R. \end{aligned}$$

Eine Funktion $f\colon \mathbb{N}^n \to \mathbb{N}$ heißt *repräsentierbar* in T, wenn es eine Formel $A[x_1, \ldots, x_n, y]$ gibt, die den Graphen $G_f \subseteq \mathbb{N}^{n+1}$ repräsentiert, also die

$$T \vdash A[\underline{a_1}, \ldots, \underline{a_n}, \underline{f(a_1, \ldots, a_n)}], \tag{4.1}$$

$$T \vdash \neg A[\underline{a_1}, \ldots, \underline{a_n}, \underline{c}], \qquad \text{falls } c \neq f(a_1, \ldots, a_n) \tag{4.2}$$

erfüllt, und für die zusätzlich gilt

$$T \vdash A[\underline{a_1}, \ldots, \underline{a_n}, y] \wedge A[\underline{a_1}, \ldots, \underline{a_n}, z] \to y = z \quad \text{für alle } a_1, \ldots, a_n \in \mathbb{N}. \tag{4.3}$$

Man beachte, daß für den Fall $T \vdash \underline{b} \neq \underline{c}$ für $b < c$ die Bedingung (4.2) aus (4.1) und (4.3) folgt.

Lemma 4.2.1. *Ist die charakteristische Funktion 1_R einer Relation $R \subseteq \mathbb{N}^n$ in T repräsentierbar, so auch die Relation R selbst.*

Beweis. Übung OBdA sei $n = 1$. Sei $A[x, y]$ eine 1_R repräsentierende Formel. Wir zeigen, daß dann $A[x, \underline{1}]$ die Relation R repräsentiert. Sei also zunächst $a \in R$. Dann ist $1_R(a) = 1$, also $(a, 1) \in G_{1_R}$, also $T \vdash A[\underline{a}, \underline{1}]$. Sei nun $a \notin R$. Dann ist $1_R(a) = 0$, also $(a, 1) \notin G_{1_R}$, also $T \vdash \neg A[\underline{a}, \underline{1}]$. □

Lemma 4.2.2. *(Fixpunktlemma). Sind alle primitiv rekursiven Funktionen in T repräsentierbar, so findet man zu jeder Formel $B[z]$ eine geschlossene Formel A mit*

$$T \vdash A \leftrightarrow B[\underline{\ulcorner A \urcorner}].$$

Beweis. Wir gehen zunächst wie im Beweis des semantischen Fixpunktlemmas 4.1.1 vor. Sei also $A_s[x_1, x_2, x_3]$ eine die primitiv rekursive Funktion $s(b, k) := \mathsf{sub}(b, \langle\langle \ulcorner z \urcorner, \ulcorner \underline{k} \urcorner \rangle\rangle)$ repräsentierende Formel. Setze

$$C := \exists x . B[x] \wedge A_s[z, z, x],$$
$$A := C[\underline{\ulcorner C \urcorner}],$$

d.h.

$$A = \exists x . B[x] \wedge A_s[\underline{\ulcorner C \urcorner}, \underline{\ulcorner C \urcorner}, x].$$

Wegen $s(\ulcorner C \urcorner, \ulcorner C \urcorner) = \ulcorner C[\underline{\ulcorner C \urcorner}] \urcorner = \ulcorner A \urcorner$ ist in T herleitbar

$$A_s[\underline{\ulcorner C \urcorner}, \underline{\ulcorner C \urcorner}, x] \leftrightarrow x = \underline{\ulcorner A \urcorner},$$

also nach Definition von A auch

$$A \leftrightarrow \exists x . B[x] \wedge x = \underline{\ulcorner A \urcorner}$$

und damit

$$A \leftrightarrow B[\underline{\ulcorner A \urcorner}].$$ □

Mit $T = \mathsf{Th}(\mathcal{M})$ ergibt sich das obige (semantische) Fixpunktlemma 4.1.1 als Spezialfall.

Satz 4.2.3. *(Undefinierbarkeit der Wahrheitsbegriffs). Sei T eine konsistente Theorie, in der alle primitiv rekursiven Funktionen repräsentierbar sind. Dann kann es keine Formel $B[z]$ geben, so daß für alle geschlossenen Formeln A gilt*

$$T \vdash A \leftrightarrow B[\underline{\ulcorner A \urcorner}].$$

Beweis. Angenommen, wir hätten so ein $B[z]$. Wir betrachten nun die Formel $\neg B[z]$ und wählen nach dem Fixpunktlemma 4.2.2 eine geschlossene Formel A mit

$$T \vdash A \leftrightarrow \neg B[\underline{\ulcorner A \urcorner}].$$

Für dieses A hätte man also $T \vdash A \leftrightarrow \neg A$ im Widerspruch zur Konsistenz von T. □

Mit $T = \mathsf{Th}(\mathcal{M})$ ergibt sich der TARSKIsche Undefinierbarkeitssatz 4.1.2 wieder als Spezialfall.

4.3 Unentscheidbarkeit und Unvollständigkeit

In diesem Abschnitt betrachten wir eine konsistente formale Theorie, in der alle rekursiven Funktionen repräsentierbar sind. Dies ist eine sehr schwache Voraussetzung, wie wir im nächsten Abschnitt zeigen werden. Sie ist erfüllt, sobald die Theorie ein gewisses Minimum an Arithmetik zu beweisen gestattet.

Wir zeigen, daß eine solche Theorie notwendigerweise unentscheidbar ist. Weiter zeigen wir den ersten GÖDELschen Unvollständigkeitssatz, der aussagt, daß eine axiomatisierte derartige Theorie immer unvollständig sein muß. Diesen Satz beweisen wir dann noch in einer von ROSSER verschärften Form, in der eine geschlossene Formel A angegeben wird, so daß weder A noch $\neg A$ in der Theorie beweisbar ist.

In diesem Abschnitt sei $\mathcal{L}$ wieder eine primitiv rekursiv präsentierte Sprache mit $0, \mathsf{S}, =$ in $\mathcal{L}$ und T eine Theorie mit $\mathsf{Eq}_{\mathcal{L}} \subseteq T$.

Satz 4.3.1. *(Unentscheidbarkeit). Ist T eine konsistente Theorie, in der alle rekursiven Funktionen repräsentierbar sind, so ist T nicht rekursiv.*

Beweis. Nehmen wir an, T ist rekursiv. Nach Voraussetzung ist dann $\ulcorner T \urcorner$ durch eine Formel $B[z]$ in T repräsentierbar. Wir wählen nach dem Fixpunktlemma 4.2.2 eine geschlossene Formel A mit

$$T \vdash A \leftrightarrow \neg B[\underline{\ulcorner A \urcorner}]$$

und zeigen (1) $T \nvdash A$ und (2) $T \vdash A$; dies ist der gewünschte Widerspruch.

Zu (1). Gelte $T \vdash A$. Dann ist $A \in T$, also $\ulcorner A \urcorner \in \ulcorner T \urcorner$ und damit $T \vdash B[\underline{\ulcorner A \urcorner}]$ (da $B[z]$ die Menge $\ulcorner T \urcorner$ in T repräsentiert). Es folgt $T \vdash \neg A$ nach Wahl von A und damit ein Widerspruch zur Konsistenz von T.

Zu (2). Nach (1) wissen wir $T \nvdash A$. Daher ist $A \notin T$, also $\ulcorner A \urcorner \notin \ulcorner T \urcorner$ und damit $T \vdash \neg B[\underline{\ulcorner A \urcorner}]$. Es folgt $T \vdash A$ nach Wahl von A. □

Satz 4.3.2. *(Erster* GÖDEL*scher Unvollständigkeitssatz). Jede axiomatisierte konsistente Theorie, in der alle rekursiven Funktionen repräsentierbar sind, ist unvollständig.*

Beweis. Die Behauptung ergibt sich aus Korollar 3.3.8 und Satz 4.3.1. □

Wie oben schon erwähnt, wollen wir jetzt unter Verwendung einer Idee von ROSSER den Unvollständigkeitssatz in dem Sinn verschärfen, daß wir eine geschlossene Formel A konkret angegeben, für die weder A noch $\neg A$ in der Theorie beweisbar ist.

Satz 4.3.3. *(*GÖDEL-ROSSER*). Sei T eine axiomatisierte konsistente $\mathcal{L}$-Theorie mit $0, \mathsf{S}, =$ in $\mathcal{L}$ und $\mathsf{Eq}_{\mathcal{L}} \subseteq T$. Ferner gebe es eine Formel $K[x,y]$ – geschrieben $x < y$ – so daß*

$$T \vdash \forall x. x < \underline{n} \to x = \underline{0} \vee \cdots \vee x = \underline{n-1}, \tag{4.4}$$
$$T \vdash \forall x. x = \underline{0} \vee \cdots \vee x = \underline{n} \vee \underline{n} < x. \tag{4.5}$$

Weiter sei jede primitiv rekursive Funktion in T repräsentierbar. Dann findet man eine geschlossene Formel A, für die weder A noch $\neg A$ in T beweisbar ist.

Beweis. Wir definieren zunächst $\mathsf{Wdl}_T \subseteq \mathbb{N} \times \mathbb{N}$ durch

$$\mathsf{Wdl}_T(d,a) :\leftrightarrow \mathsf{Abl}_T(d, \dot\neg a).$$

Dann ist Wdl_T primitiv rekursiv, und es gilt $\mathsf{Wdl}_T(d,a)$ genau dann, wenn d eine Widerlegung von a in T ist, d.h. wenn d eine Herleitung einer Sequenz $\Gamma \Rightarrow \neg A$ kodiert mit $a = \ulcorner \neg A \urcorner$ und Γ zusammengesetzt aus Stabilitätsaxiomen und Formeln aus Ax_T. Seien $B_{\mathsf{Abl}_T}[x_1,x_2]$ und $B_{\mathsf{Wdl}_T}[x_1,x_2]$ repräsentierende Formeln zu Abl_T und Wdl_T. Wir wählen nun nach dem Fixpunktlemma 4.2.2 eine geschlossene Formel A mit

$$T \vdash A \leftrightarrow \forall x. B_{\mathsf{Abl}_T}[x, \underline{\ulcorner A \urcorner}] \to \exists y. y < x \wedge B_{\mathsf{Wdl}_T}[y, \underline{\ulcorner A \urcorner}].$$

A drückt also seine eigene Unbeweisbarkeit aus, und zwar in der (auf ROSSER zurückgehenden) Form "Zu jedem Beweis von mir gibt es einen kürzeren Beweis meiner Negation".

Wir zeigen (1) $T \nvdash A$ und (2) $T \nvdash \neg A$. Zu (1). Gelte $T \vdash A$. Man wähle ein n mit

$$\mathsf{Abl}_T(n, \ulcorner A \urcorner).$$

Dann gilt auch

$$\text{nicht } \mathsf{Wdl}_T(m, \ulcorner A \urcorner) \qquad \text{für alle } m,$$

da T konsistent ist. Folglich haben wir

$$T \vdash B_{\mathsf{Abl}_T}[\underline{n}, \underline{\ulcorner A \urcorner}],$$
$$T \vdash \neg B_{\mathsf{Wdl}_T}[\underline{m}, \underline{\ulcorner A \urcorner}] \qquad \text{für alle } m.$$

Daraus folgt wegen (4.4)

$$T \vdash B_{\mathsf{Abl}_T}[\underline{n}, \underline{\ulcorner A \urcorner}] \land \forall y.y < \underline{n} \to \neg B_{\mathsf{Wdl}_T}[y, \underline{\ulcorner A \urcorner}].$$

Also gilt

$$T \vdash \exists x.B_{\mathsf{Abl}_T}[x, \underline{\ulcorner A \urcorner}] \land \forall y.y < x \to \neg B_{\mathsf{Wdl}_T}[y, \underline{\ulcorner A \urcorner}],$$
$$T \vdash \neg A.$$

Dies ist ein Widerspruch zur Konsistenz von T.

Zu (2). Gelte $T \vdash \neg A$. Man wähle ein n mit

$$\mathsf{Wdl}_T(n, \ulcorner A \urcorner).$$

Dann gilt auch

$$\text{nicht } \mathsf{Abl}_T(m, \ulcorner A \urcorner) \qquad \text{für alle } m,$$

da T konsistent ist. Folglich haben wir

$$T \vdash B_{\mathsf{Wdl}_T}[\underline{n}, \underline{\ulcorner A \urcorner}],$$
$$T \vdash \neg B_{\mathsf{Abl}_T}[\underline{m}, \underline{\ulcorner A \urcorner}] \qquad \text{für alle } m.$$

Daraus erhält man aber

$$T \vdash \forall x.B_{\mathsf{Abl}_T}[x, \underline{\ulcorner A \urcorner}] \to \exists y.y < x \land B_{\mathsf{Wdl}_T}[y, \underline{\ulcorner A \urcorner}],$$

wie man durch Fallunterscheidung nach x mit Hilfe von (4.5) leicht beweist. Also gilt $T \vdash A$. Dies ist wieder ein Widerspruch zur Konsistenz von T. □

Schließlich wollen wir noch eine Variante dieses Satzes formulieren, in der wir nicht mehr davon ausgehen, daß die Theorie T nur etwas über Zahlen aussagt.

Satz 4.3.4. *(*Gödel-Rosser*). Sei T eine axiomatisierte konsistente $\mathcal{L}$-Theorie mit $0, \mathsf{S}, =$ in $\mathcal{L}$ und $\mathsf{Eq}_{\mathcal{L}} \subseteq T$. Ferner gebe es Formeln $N[x]$ und $K[x,y]$ – geschrieben Nx bzw. $x < y$ – so daß $T \vdash N0$, $T \vdash \forall x{\in}N\, N[\mathsf{S}(x)]$ und*

$$T \vdash \forall x{\in}N.x < \underline{n} \to x = \underline{0} \lor \cdots \lor x = \underline{n-1},$$
$$T \vdash \forall x{\in}N.x = \underline{0} \lor \cdots \lor x = \underline{n} \lor \underline{n} < x.$$

Hierbei steht $\forall x{\in}N\, A$ für $\forall x.Nx \to A$. Ferner sei jede primitiv rekursive Funktion in T repräsentierbar. Dann findet man eine geschlossene Formel A, für die weder A noch $\neg A$ in T beweisbar ist.

Beweis. Wie für den Satz 4.3.3; man muß lediglich die auftretenden Quantoren auf N relativieren. □

4.4 Repräsentierbarkeit

Wir zeigen in diesem Abschnitt, daß schon recht einfache formale Theorien die Eigenschaft besitzen, daß in ihnen jede rekursive Funktion repräsentierbar ist. Als Hilfsmittel für den Beweis zeigen wir zunächst, daß die Menge der rekursiven Funktionen auch ohne das Schema der primitiven Rekursion erzeugt werden kann, also alleine durch Komposition und den unbeschränkten μ-Operator aus einfachen Ausgangsfunktionen.

Wir definieren die Menge $\mathcal{G}$ der *μ'-rekursiven* Funktionen als die kleinste Menge von Funktionen mit den folgenden Eigenschaften. Wir schreiben $\mathcal{G}^n$ für die Teilmenge der n-stelligen Funktionen in $\mathcal{G}$.

1. $0^n \in \mathcal{G}^n$ $(n \geq 0)$, $\mathsf{S} \in \mathcal{G}^1$, $I_i^n \in \mathcal{G}^n$ $(1 \leq i \leq n)$, $+, \cdot, 1_< \in \mathcal{G}^2$.
2. Sind $g_1, \ldots, g_m \in \mathcal{G}^n$, $h \in \mathcal{G}^m$ und $m \geq 1$, $n \geq 0$, so ist $(\circ h g_1 \ldots g_m) \in \mathcal{G}^n$.
3. Ist $g \in \mathcal{G}^{n+1}$ und gilt $\forall \boldsymbol{a} \exists i\, (g(i, \boldsymbol{a}) = 0)$, so ist μg in $\mathcal{G}^n$.

Lemma 4.4.1. *Die Funktionen* $\dot{-}$ *sowie* π, π_1, π_2 *sind* μ'*-rekursiv.*

Beweis. $a \dot{-} b = \mu i\, (a < i + b + 1)$. Es war $\pi(a, b) = \frac{1}{2}(a+b)(a+b+1) + a$. Zum Beweis von $\pi \in \mathcal{G}$ genügt es offenbar, ein $h \in \mathcal{G}$ zu finden mit $h(2c) = c$. Eine solche Funktion ist $h(d) := \mu i\, (d < 2i + 1)$.

Wir zeigen jetzt $\pi_1 \in \mathcal{G}^1$. Dazu beachte man, daß

$$2\pi(a, b) = (a + b)(a + b + 1) + 2a. \tag{4.6}$$

Nun gilt

$$(a + b)(a + b + 1) \leq (a + b)(a + b + 1) + 2a < (a + b + 1)(a + b + 2).$$

Setzt man also

$$g(c) := \mu i\, (2c < (i + 1)(i + 2)),$$

so ist $g(\pi(a, b)) = a + b$. Wegen (4.6) können wir also aus $c := \pi(a, b)$ die Zahl $2a$ gewinnen als $2c - g(c)(g(c) + 1)$. Damit gilt

$$\pi_1(c) = h(2c \dot{-} g(c)(g(c) + 1))$$

und wir haben $\pi_1 \in \mathcal{G}$ gezeigt.

$\pi_2 \in \mathcal{G}$ folgt jetzt aus $\pi_2(c) = g(c) \dot{-} \pi_1(c)$. □

Lemma 4.4.2. (Gödel). *Es gibt eine* μ'*-rekursive Funktion* β *mit folgender Eigenschaft. Zu beliebigen* $a_0, \ldots, a_{n-1}$ *findet man ein* c *mit* $\beta(c, i) = a_i$ *für alle* $i < n$.

Beweis. Sei

$$a := \max_{i<n} \pi(a_i, i).$$

Man beachte zunächst, daß alle $1 + ka!$ für $k \leq a$ relativ prim sind. Hätten nämlich $1 + ia!$ und $1 + ja!$ mit $i < j \leq a$ einen gemeinsamen Primteiler p, so wäre $p | (j - i)a!$, also auch $p | a!$ und damit $p | 1$, was nicht sein kann.

Wir setzen jetzt

$$b := a!,$$
$$d := \prod_{i<n} \bigl(1 + \pi(a_i, i)a!\bigr).$$

Aus b und d kann man zu gegebenem $i < n$ die Zahl a_i wie folgt wiedergewinnen. Man betrachte das kleinste x mit

$$1 + \pi(x, i)b | d.$$

Offenbar ist a_i ein solches x. Würde nun ein $x < a_i$ die Bedingung auch erfüllen, so wäre wegen $\pi(x, i) < a$ und der Tatsache, daß alle $1 + ka!$ für $k \leq a$ relativ prim sind, $\pi(x, i) = \pi(a_j, j)$ für ein $j < n$ nach Konstruktion von d, also $x = a_j$ und $i = j$ im Widerspruch zu $x < a_i$.

Wir können also definieren

$$\beta(c, i) := \pi_1\bigl(\mu y\, [(1 + \pi(\pi_1(y), i) \cdot \pi_1(c)) \cdot \pi_2(y) \geq \pi_2(c)]\bigr).$$

Diese Funktion heißt Gödel*sche* β*-Funktion.* β ist μ'-rekursiv, denn für jedes c, i existiert offenbar ein solches y. Für $c := \pi(b, d)$ ist dann $\pi(a_i, \lceil d / 1 + \pi(a_i, i)b \rceil)$ das kleinste solche y und wir haben $\beta(c, i) = a_i$. □

Satz 4.4.3. *Die Menge der* μ'*-rekursiven Funktionen ist abgeschlossen unter primitiver Rekursion, stimmt also mit der Menge der* μ*-rekursiven Funktionen überein.*

Beweis. Sei $g \in \mathcal{G}^n$, $h \in \mathcal{G}^{n+2}$ und $f = (Rgh)$. Nach dem vorangehenden Lemma findet man zu gegebenen $a, \boldsymbol{b}$ ein c mit $\beta(c,i) = f(i,\boldsymbol{b})$ für alle $i \le a$, also

$$\begin{aligned} \beta(c,0) &= g(\boldsymbol{b}), \\ \beta(c,i+1) &= h(i,\beta(c,i),\boldsymbol{b}) \quad \text{für alle } i < a. \end{aligned}$$

Setze

$$\begin{aligned} |a-b| &:= (a \mathbin{\dot-} b) + (b \mathbin{\dot-} a), \\ h_1(a,\boldsymbol{b},c) &:= \mu i \left((a \mathbin{\dot-} i)(1 \mathbin{\dot-} |h(i,\beta(c,i),\boldsymbol{b}) - \beta(c,i+1)|) = 0\right), \\ h_2(a,\boldsymbol{b},c) &:= |g(\boldsymbol{b}) - \beta(c,0)| + (a \mathbin{\dot-} h_1(a,\boldsymbol{b},c)). \end{aligned}$$

Dann ist offenbar $h_2 \in \mathcal{G}$ und $h_2(a,\boldsymbol{b},c) = 0$ ist äquivalent zu der Gültigkeit der obigen Gleichungen für β. Insbesondere gibt es zu jedem $a, \boldsymbol{b}$ ein c mit $h_2(a,\boldsymbol{b},c) = 0$. Daher gilt

$$f(a,\boldsymbol{b}) = \beta(\mu c\,(h_2(a,\boldsymbol{b},c) = 0), a)$$

und f ist deshalb μ'-rekursiv. □

Wir haben also gesehen, daß die μ-rekursiven und die μ'-rekursiven Funktionen übereinstimmen; im folgenden sprechen wir deshalb nur noch von μ-rekursiven Funktionen.

Satz 4.4.4. *Sei T eine $\mathcal{L}$-Theorie mit $0, \mathsf{S}, =$ in $\mathcal{L}$ und $\mathsf{Eq}_{\mathcal{L}} \subseteq T$. Ferner gebe es Formeln $N[x]$ und $K[x,y]$ – geschrieben Nx bzw. $x < y$ – so daß $T \vdash N0$, $T \vdash \forall x{\in}N\, N[\mathsf{S}(x)]$ und folgendes gilt.*

$$\begin{aligned} &T \vdash \mathsf{S}(\underline{a}) \ne 0 && \text{für alle } a \in \mathbb{N}, && (4.7) \\ &T \vdash \mathsf{S}(\underline{a}) = \mathsf{S}(\underline{b}) \to \underline{a} = \underline{b} && \text{für alle } a,b \in \mathbb{N}, && (4.8) \\ &\text{die Funktionen } + \text{ und } \cdot \text{ sind in } T \text{ repräsentierbar,} && && (4.9) \\ &T \vdash \forall x{\in}N\,(x \not< 0) && && (4.10) \\ &T \vdash \forall x{\in}N.x < \mathsf{S}(\underline{b}) \to x < \underline{b} \lor x = \underline{b} && \text{für alle } b \in \mathbb{N}, && (4.11) \\ &T \vdash \forall x{\in}N.x < \underline{b} \lor x = \underline{b} \lor \underline{b} < x && \text{für alle } b \in \mathbb{N}. && (4.12) \end{aligned}$$

Hierbei steht wieder $\forall x{\in}N\,A$ für $\forall x.Nx \to A$. Dann erfüllt T die Voraussetzungen von Satz 4.3.4, also des (auf N relativierten) Satzes von GÖDEL-ROSSER. *Genauer gilt für alle $a \in \mathbb{N}$*

$$T \vdash \forall x{\in}N.x < \underline{a} \to x = \underline{0} \lor \dots \lor x = \underline{a-1}, \tag{4.13}$$

$$T \vdash \forall x{\in}N.x = \underline{0} \lor \dots \lor x = \underline{a} \lor \underline{a} < x, \tag{4.14}$$

und jede rekursive Funktion ist in T repräsentierbar.

Beweis. Wir zeigen zunächst, daß die Formel $x = y$ die Gleichheit und die Formel $x < y$ die Kleinerbeziehung in T repräsentiert. Aus (4.7) and (4.8) folgt sofort $T \vdash \underline{a} \ne \underline{b}$ für $a \ne b$. Gelte nun $a \not< b$. Wir zeigen $T \vdash \underline{a} \not< \underline{b}$ durch Induktion nach b. $T \vdash \underline{a} \not< 0$ folgt aus (4.10). Im Schritt haben wir $a \not< b+1$, also $a \not< b$ und $a \ne b$, also nach IH und der obigen Bemerkung $T \vdash \underline{a} \not< \underline{b}$ und $T \vdash \underline{a} \ne \underline{b}$, also nach (4.11) $T \vdash \underline{a} \not< \mathsf{S}(\underline{b})$. Sei jetzt $a < b$. Dann gilt $T \vdash \underline{a} \ne \underline{b}$ und $T \vdash \underline{b} \not< \underline{a}$, also nach (4.12) $T \vdash \underline{a} < \underline{b}$.

(4.13) ergibt sich jetzt leicht durch Induktion über a. Die Basis folgt aus (4.10), und der Schritt aus der IH und (4.11). (4.14) ergibt sich aus (4.12) unmittelbar mit (4.13).

Wir zeigen jetzt durch Induktion über die Definition der μ'-rekursiven Funktionen, daß jede μ'-rekursive (und damit nach Satz 4.4.3 auch jede rekursive) Funktion in T repräsentierbar ist. Man beachte hierbei, daß wegen $T \vdash \underline{a} \ne \underline{b}$ für $a \ne b$ die Bedingung (4.2) in der Definition der Repräsentierbarkeit einer Funktion nicht mehr überprüft werden muß.

Die *Ausgangsfunktionen* 0^n, S und I_i^n werden offenbar durch die Formeln $0 = y$, $\mathsf{S}(x) = y$ und $x_i = y$ repräsentiert. $+$ und $\cdot$ sind nach Voraussetzung (4.9) in T repräsentierbar. Wir zeigen noch, daß die Funktion $1_<$ in T durch

$$A := (x_1 < x_2 \land y = 1) \lor (x_1 \not< x_2 \land y = 0)$$

repräsentiert wird. Gelte also $a_1 < a_2$. Dann ist $T \vdash \underline{a_1} < \underline{a_2}$, also $T \vdash A[\underline{a_1}, \underline{a_2}, 1]$. Sei jetzt $a_1 \not< a_2$. Dann ist $T \vdash \underline{a_1} \not< \underline{a_2}$, also $T \vdash A[\underline{a_1}, \underline{a_2}, 0]$. Ferner ist $A[x_1, x_2, y] \wedge A[x_1, x_2, z] \to y = z$ schon logisch aus den Gleichheitsaxiomen herleitbar (Fallunterscheidung nach $x_1 < x_2$).

Fall $(\circ h g_1 \dots g_m)$. OBdA seien $g_1, \dots, g_m$ einstellig. Nach IH haben wir repräsentierende Formeln $A_{g_i}[x, y_i]$ und $A_h[\boldsymbol{y}, z]$. Wir setzen

$$A_f := \exists \boldsymbol{y}. A_{g_1}[x, y_1] \wedge \dots \wedge A_{g_m}[x, y_m] \wedge A_h[\boldsymbol{y}, z].$$

Gelte $f(a) = c$. Offenbar haben wir dann $T \vdash A_f[\underline{a}, \underline{c}]$. Zu zeigen bleibt $T \vdash A_f[\underline{a}, z_1] \wedge A_f[\underline{a}, z_2] \to z_1 = z_2$. Wir führen den Beweis wieder informal. Nach Definition von A_f haben wir $y_{11}, \dots, y_{1m}$ und $y_{21}, \dots, y_{2m}$. Die IH für g_i liefert $y_{1i} = y_{2i} = \underline{g_i(a)}$. Mit der IH für h ergibt sich $z_1 = z_2$.

Fall (μg). OBdA sei g zweistellig, also $f(a) = \mu i\, (g(i, a) = 0)$, wobei $\forall a \exists i\, (g(i, a) = 0)$. Nach IH haben wir eine g repräsentierende Formel $A_g[y, x, z]$. Sei

$$A_f[x, y] := Ny \wedge A_g[y, x, 0] \wedge \forall v{\in}N. v < y \to \exists w. w \neq 0 \wedge A_g[v, x, w].$$

Wir zeigen zunächst (4.1). Gelte also $f(a) = b$. Zu zeigen ist $T \vdash A_f[\underline{a}, \underline{b}]$. Aufgrund der Gestalt von A_f folgt dies aber sofort mit $T \vdash v < \underline{b} \to v = \underline{0} \vee \dots \vee v = \underline{b-1}$ aus der IH für g. Wir zeigen jetzt noch (4.3). Gegeben sei also a; setze $b := f(a)$. Es genügt zu zeigen $T \vdash A_f[\underline{a}, y] \to y = \underline{b}$. Den Beweis führen wir wieder informal. Gelte also $A_f[\underline{a}, y]$. Nach Voraussetzung (4.12) genügt es, $y \not< \underline{b}$ und $\underline{b} \not< y$ zu zeigen. Nehmen wir also zunächst $y < \underline{b}$ an. Nach (4.13) folgt $y = \underline{i}$ für ein $i < b$ im Widerspruch zu $A_g[y, \underline{a}, 0]$. Nehmen wir jetzt $\underline{b} < y$ an. Aus $A_f[\underline{a}, y]$ folgt dann $\exists w. w \neq 0 \wedge A_g[\underline{b}, \underline{a}, w]$ im Widerspruch zu $A_g[\underline{b}, \underline{a}, 0]$. □

Wir wollen jetzt noch eine spezielle und besonders einfache arithmetische Theorie betrachten. Sei $\mathcal{L}_1$ die durch 0, S, $+$, $\cdot$ und $=$ bestimmte Sprache und Z_1 die durch $\mathsf{Eq}_{\mathcal{L}_1}$ und die Generalisierungen der folgenden Axiome bestimmte Theorie.

$$\mathsf{S}(x) \neq 0, \tag{4.15}$$
$$\mathsf{S}(x) = \mathsf{S}(y) \to x = y, \tag{4.16}$$
$$x + 0 = x, \tag{4.17}$$
$$x + \mathsf{S}(y) = \mathsf{S}(x + y), \tag{4.18}$$
$$x \cdot 0 = 0, \tag{4.19}$$
$$x \cdot \mathsf{S}(y) = x \cdot y + x, \tag{4.20}$$
$$\exists z\, (x + \mathsf{S}(z) = y) \vee x = y \vee \exists z\, (y + \mathsf{S}(z) = x). \tag{4.21}$$

Satz 4.4.5. *Jede Theorie T mit $T \supseteq Z_1$ erfüllt die Voraussetzungen des Satzes 4.3.3 von* GÖDEL-ROSSER. *Bzgl. $K[x, y] := \exists z\, (x + \mathsf{S}(z) = y)$ ist sogar jede rekursive Funktion in T repräsentierbar.*

Beweis. Wir zeigen, daß T mit $N[x] := (x = x)$ und $K[x, y] := \exists z\, (x + \mathsf{S}(z) = y)$ die Bedingungen von Satz 4.4.4 erfüllt. Für (4.7) und (4.8) ist dies klar. Für (4.9) können wir $x + y = z$ und $x \cdot y = z$ als repräsentierende Formeln wählen. Für (4.10) ist zu zeigen $\neg \exists z\, (x + \mathsf{S}(z) = 0)$. Dies folgt aber aus (4.18) und (4.15). Zum Beweis von (4.11) benötigen wir die Hilfsaussage

$$x = 0 \vee \exists y\, (x = \mathsf{S}(y)), \tag{4.22}$$

deren Beweis wir unten nachtragen. Gelte also $x + \mathsf{S}(z) = \mathsf{S}(\underline{b})$, also auch $\mathsf{S}(x + z) = \mathsf{S}(\underline{b})$ und damit $x + z = \underline{b}$. Wir verwenden jetzt (4.22) für z. Im Fall $z = 0$ folgt $x = \underline{b}$, und im Fall $\exists y\, (z = \mathsf{S}(y))$ haben wir $\exists y\, (x + \mathsf{S}(y) = \underline{b})$. Damit ist (4.11) bewiesen. (4.12) folgt sofort aus (4.21).

Zum Beweis von (4.22) verwenden wir ebenfalls (4.21) und vergleichen x mit 0. Es genügt offenbar, den Fall $\exists z\, (x + \mathsf{S}(z) = 0)$ auszuschließen. Er besagt aber $\mathsf{S}(x + z) = 0$ und widerspricht deshalb (4.15). □

Korollar 4.4.6. *(Starke Unentscheidbarkeit von Z_1). Jede konsistente Theorie T mit $T \supseteq Z_1$ ist nicht rekursiv.*

Beweis. Sätze 4.4.5 und 4.3.1. □

Korollar 4.4.7. *(Unentscheidbarkeit der Prädikatenlogik). Sei $\mathcal{L}_1$ die durch* 0, S, $+$, $\cdot$ *und* $=$ *bestimmte Sprache. Dann ist die Menge der in der klassischen Logik herleitbaren $\mathcal{L}_1$-Formeln nicht rekursiv.*

Beweis. Andernfalls wäre auch Z_1 rekursiv, denn eine Formel A ist herleitbar in Z_1 genau dann, wenn die Implikation aus der Konjunktion der endlich vielen Z_1-Axiome und der Gleichheitsaxiome in der klassischen Logik herleitbar ist. □

In Abschnitt 3.2 hatten wir Σ_1-Formeln der formalen Sprache PR definiert, ausgehend von eventuell negierten PR-Primformeln. Wir wollen jetzt einen entsprechenden Begriff für die durch 0, S, $+$, $\cdot$ und $=$ bestimmte Sprache $\mathcal{L}_1$ definieren, und zwar so, daß immer noch genau die rekursiv aufzählbaren Relationen durch Σ_1-Formeln definierbar sind. In Anbetracht der stark verringerten sprachlichen Ausdrucksmittel ist es jetzt sinnvoll, nur ganz spezielle atomare und negiert atomare Formeln zuzulassen. *Σ_1-Formeln der Sprache $\mathcal{L}_1$* werden wie folgt induktiv definiert.

- Für alle Variablen x, y, z sind $x = y$, $x \neq y$, $0 = x$, $\mathsf{S}(x) = y$, $x + y = z$ und $x \cdot y = z$ Σ_1-Formeln der Sprache $\mathcal{L}_1$.
- Mit A und B sind auch $A \wedge B$ und $A \vee B$ Σ_1-Formeln der Sprache $\mathcal{L}_1$.
- Mit A ist auch $\forall x{<}y\, A$ eine Σ_1-Formel der Sprache $\mathcal{L}_1$, falls y von x verschieden ist.
- Mit A ist auch $\exists x\, A$ eine Σ_1-Formel der Sprache $\mathcal{L}_1$.

Damit jede Σ_1-Formel der Sprache $\mathcal{L}_1$ auch wirklich eine $\mathcal{L}_1$-Formel ist, verstehen wir $\forall x{<}y\, A$ als Abkürzung für $\forall x.\exists z\, (x + \mathsf{S}(z) = y) \to A$.

Satz 4.4.8. *Jede rekursive Funktion ist in Z_1 repräsentierbar durch eine Σ_1-Formel der Sprache $\mathcal{L}_1$.*

Beweis. Dies ergibt sich sofort durch Inspektion des Beweises von Satz 4.4.4. Man hat lediglich zu beachten, daß schon aufgrund der Gleichheitsaxiome $\exists z\, (x + \mathsf{S}(z) = y)$ mit $\exists z \exists w\, (\mathsf{S}(z) = w \wedge x + w = y)$ und $A[0]$ mit $\exists x.x = 0 \wedge A$ äquivalent sind. □

$\mathcal{L}_1$-Formeln heißen auch *arithmetische* Formeln. Eine Relation $R \subseteq \mathbb{N}^n$ heißt *arithmetisch*, wenn sie im *Standardmodell* $\mathcal{N}_1$ der Sprache $\mathcal{L}_1$ definierbar ist.

Korollar 4.4.9. *1. Jede rekursiv ausfzählbare Relation ist arithmetisch.*
2. $\mathsf{Th}(\mathcal{N}_1)$ *ist nicht arithmetisch, insbesondere also auch nicht rekursiv aufzählbar.*

Beweis. Der erste Teil folgt aus Satz 4.4.8 (mit der aus Lemma 3.2.1 bekannten Tatsache, daß jede nichtleere Menge $M \subseteq \mathbb{N}$ Wertebereich einer primitiv rekursiven Funktion ist), und der zweite aus dem Tarskischen Undefinierbarkeitssatz 4.1.2. □

4.5 Unbeweisbarkeit der Widerspruchsfreiheit

Wir haben oben im Satz 4.3.3 von Gödel-Rosser gesehen, wie man zu jeder axiomatisierten konsistenten Theorie T, die gewisse schwache Voraussetzungen erfüllt, einen geschlossenen Satz A konstruieren kann so daß weder A noch $\neg A$ in T beweisbar ist. Die inhaltliche Bedeutung dieses Satzes A war "Zu jedem Beweis von mir gibt es einen kürzeren Beweis meiner Negation". Da also A keinen Beweis besitzt, ist insbesondere A wahr. Wir haben also einen wahren, aber in T unbeweisbaren Satz gefunden.

Diese Aussage wollen wir jetzt noch verschärfen, indem wir zeigen, daß ein besonders interessanter wahrer Satz in T unbeweisbar ist, nämlich der Satz Wf_T, der die *Konsistenz* (oder *Widerspruchsfreiheit*) von T besagt. Dies ist der zweite Unvollständigkeitssatz von Gödel [13].

Wir werden diesen Satz hier in einer von Löb angegebenen verschärften Form beweisen. Zunächst zeigen wir eine Hilfsaussage.

Lemma 4.5.1. *Sei $A[x_1, \dots, x_n]$ eine Σ_1-Formel der durch* 0, S, $+$, $\cdot$ *und* $=$ *bestimmten Sprache $\mathcal{L}_1$. Gilt dann $\mathcal{N}_1 \models A[a_1, \dots, a_n]$ für das Standardmodell $\mathcal{N}_1$ von $\mathcal{L}_1$, so folgt $Z_1 \vdash A[\underline{a_1}, \dots, \underline{a_n}]$.*

Beweis. Durch Induktion über die Σ_1-Formeln der Sprache $\mathcal{L}_1$. Die *Anfangsfälle* hatten wir zu Beginn des Beweises von Satz 4.4.4 erledigt bzw. sie ergeben sich (für $x + y = z$ und $x \cdot y = z$) aus den Rekursionsgleichungen für $+$ und $\cdot$ in (4.17)-(4.20).

Fälle $A \wedge B$, $A \vee B$. Die Behauptung folgt sofort aus der IH.

Fall $\forall x{<}y\, A[x, y, z_1, \ldots, z_n]$; oBdA sei $n = 1$. Gelte also $\mathcal{N}_1 \models (\forall x{<}y\, A)[b, c]$. Dann folgt $\mathcal{N}_1 \models A[i, b, c]$ für jedes $i < b$ und deshalb nach IH $Z_1 \vdash A[\underline{i}, \underline{b}, \underline{c}]$. Nun gilt nach Satz 4.4.5

$$Z_1 \vdash \forall x.\exists z\, (x + \mathsf{S}(z) = \underline{b}) \to x = \underline{0} \vee \cdots \vee x = \underline{b-1},$$

also

$$Z_1 \vdash \forall x{<}\underline{b}\, A[\underline{b}, \underline{c}].$$

Fall $\exists x\, A[x, y_1, \ldots, y_n]$; oBdA sei $n = 1$. Gelte also $\mathcal{N}_1 \models (\exists x\, A)[b]$. Dann ist $\mathcal{N}_1 \models A[a, b]$ für ein $a \in \mathbb{N}$, also nach IH $Z_1 \vdash A[\underline{a}, \underline{b}]$ und deshalb $Z_1 \vdash (\exists x\, A)[\underline{b}]$. □

Sei T eine axiomatisierte konsistente Theorie mit $T \supseteq Z_1$. Sei B_{Abl_T} wie in Abschnitt 4.3 eine $\mathcal{L}_1$-Formel, die die primitiv rekursive Relation $\mathsf{Abl}_T \subseteq \mathbb{N} \times \mathbb{N}$ in Z_1 repräsentiert. Wir definieren dann $\mathcal{L}_1$-Formeln $\mathsf{Thm}_T[x]$ und Wf_T durch

$$\mathsf{Thm}_T[x] := \exists y\, B_{\mathsf{Abl}_T}[y, x],$$
$$\mathsf{Wf}_T := \neg\exists y\, B_{\mathsf{Abl}_T}[y, \underline{\ulcorner \bot \urcorner}].$$

$\mathsf{Thm}_T[x]$ definiert also in $\mathcal{N}_1$ die Menge der in T beweisbaren Formeln, und es gilt $\mathcal{N}_1 \models \mathsf{Wf}_T$ genau dann, wenn T konsistent ist. Wir betrachten die folgenden beiden *Ableitbarkeitsbedingungen* für T.

$$T \vdash A \to \mathsf{Thm}_T[\underline{\ulcorner A \urcorner}] \quad \text{für } A \text{ geschlossene } \Sigma_1\text{-Formel der Sprache } \mathcal{L}_1, \tag{4.23}$$
$$T \vdash \mathsf{Thm}_T[\underline{\ulcorner A \to B \urcorner}] \wedge \mathsf{Thm}_T[\underline{\ulcorner A \urcorner}] \to \mathsf{Thm}_T[\underline{\ulcorner B \urcorner}]. \tag{4.24}$$

(4.23) sagt aus, daß Lemma 4.5.1 (oder genauer die Aussage, daß jede in $\mathcal{N}_1$ gültige geschlossene Σ_1-Formel der Sprache $\mathcal{L}_1$ in T liegt) nicht nur wahr, sondern sogar in T beweisbar ist.

Satz 4.5.2. *(*GÖDEL-LÖB*). Sei T eine axiomatisierte konsistente Erweiterung von Z_1, die die Ableitbarkeitsbedingungen* (4.23) *und* (4.24) *erfüllt. Gilt dann $T \vdash \mathsf{Thm}_T[\underline{\ulcorner C \urcorner}] \to C$ für eine geschlossene $\mathcal{L}_1$-Formel C, so gilt schon $T \vdash C$.*

Beweis. Gelte $T \vdash \mathsf{Thm}_T[\underline{\ulcorner C \urcorner}] \to C$. Wähle A nach dem Fixpunktlemma 4.2.2 so daß

$$Z_1 \vdash A \leftrightarrow (\mathsf{Thm}_T[\underline{\ulcorner A \urcorner}] \to C). \tag{4.25}$$

Zu zeigen ist $T \vdash C$. Wir führen den Beweis informal, müssen jedoch zeigen, daß alle Schlüsse innerhalb von T durchführbar sind. Zunächst zeigen wir

$$\mathsf{Thm}_T[\underline{\ulcorner A \urcorner}] \to C. \tag{4.26}$$

Gelte also

$$\mathsf{Thm}_T[\underline{\ulcorner A \urcorner}].$$

Nach (4.23) folgt

$$\mathsf{Thm}_T[\underline{\ulcorner \mathsf{Thm}_T[\underline{\ulcorner A \urcorner}] \urcorner}].$$

Wegen (4.25) ist die geschlossene Σ_1-Formel $\mathsf{Thm}_T[\underline{\ulcorner A \to (\mathsf{Thm}_T[\underline{\ulcorner A \urcorner}] \to C) \urcorner}]$ wahr (in $\mathcal{N}_1$), und nach Lemma 4.5.1 können wir sie auch in Z_1 beweisen. Also

$$\mathsf{Thm}_T[\underline{\ulcorner A \to (\mathsf{Thm}_T[\underline{\ulcorner A \urcorner}] \to C) \urcorner}].$$

Zweimalige Anwendung von (4.24) ergibt

$$\mathsf{Thm}_T[\underline{\ulcorner C \urcorner}]$$

und damit nach Voraussetzung C. Damit ist (4.26) bewiesen.

Aus (4.26) erhält man jetzt A, nach (4.25).

Wir verlassen nun den informalen, innerhalb von T durchführbaren Beweis und stellen fest, daß wir $T \vdash A$ gezeigt haben. Also ist die geschlossene Σ_1-Formel $\mathsf{Thm}_T[\underline{\ulcorner A \urcorner}]$ wahr (in $\mathcal{N}_1$) und damit nach Lemma 4.5.1 auch in T beweisbar. Da andererseits auch (4.26) in T beweisbar war, ist auch C in T beweisbar. Damit haben wir $T \vdash C$ gezeigt. □

Korollar 4.5.3. *(Zweiter* GÖDEL*scher Unvollständigkeitssatz). Sei T eine axiomatisierte konsistente Erweiterung von Z_1, die die Ableitbarkeitsbedingungen* (4.23) *und* (4.24) *erfüllt. Dann gilt $T \nvdash \mathsf{Wf}_T$.*

Beweis. Setze $C := \bot$ in Satz 4.5.2. □

Korollar 4.5.4. *Sei T eine axiomatisierte konsistente Erweiterung von Z_1, die die Ableitbarkeitsbedingungen* (4.23) *und* (4.24) *erfüllt. Dann ist das Reflexionsschema*

$$\mathsf{Thm}_T[\underline{\ulcorner C \urcorner}] \to C \quad \text{mit } C \text{ geschlossene } \mathcal{L}_1\text{-Formel}$$

unbeweisbar in T.

Beweis. Man wähle in Satz 4.5.2 ein C mit $T \nvdash C$, etwa $C = \bot$. □

Ein wichtiges Beispiel einer axiomatisierten konsistenten Erweiterung von Z_1, die die Ableitbarkeitsbedingungen (4.23) und (4.24) erfüllt, ist die *Peano-Zahlentheorie* Z (oft auch mit PA bezeichnet). Die Sprache von Z ist wieder $\mathcal{L}_1$ (also gegeben durch 0, S, $+$, $\cdot$ und $=$), und die Axiome sind neben $\mathsf{Eq}_{\mathcal{L}_1}$ die ersten sechs Z_1-Axiome und zusätzlich das *Induktionsschema*

$$A[0] \wedge (\forall x.A \to A[\mathsf{S}(x)]) \to \forall x\, A.$$

4.6 Anmerkungen

Die grundlegenden Arbeiten zur Unvollständigkeit stammen von GÖDEL (1930[12] , 1931[13]). GÖDEL fand auch die für den Repäsentationssatz zentrale β-Funktion, und das Fixpunktlemma verwendet er in seiner Argumentation implizit. Sein erster Unvollständigkeitssatz benutzt die Formel "ich bin nicht beweisbar", einen Fixpunkt von $\neg\mathsf{Thm}_T[x]$. Für die Unabhängigkeit dieser Aussage von der zugrunde liegenden Theorie T wird aber die ω-Konsistenz von T benötigt (die allerdings gewährleistet ist, wenn T eine Teiltheorie der Theorie des Standardmodells ist). ROSSER hat dann (1936) die hier behandelte Verschärfung angegeben mit der Formel "jeder Beweis von mir ist kürzer zu widerlegen". Die Undefinierbarkeit des Wahrheitsbegriffs stammt von TARSKI (1939), die Unentscheidbarkeit der Prädikaten-Logik erster Stufe ist ein Resultat von CHURCH (1936). Die arithmetischen Theorien R und Q (aus den Übungsaufgaben 57 und 58) stammen von R. ROBINSON (1950). R ist essentiell unentscheidbar, unvollständig, genügt für den Σ_1 Vollständigkeitssatz und alle rekursiven Prädikate sind in R repäsentierbar. Q ist eine sehr natürliche Theorie und im Gegensatz zu R endlich. Q ist minimal im folgenden Sinne: Streicht man ein Axiom, so ist die verbleibende Theorie nicht mehr essentiell unentscheidbar. (Die erste essentiell unentscheidbare endliche Theorie der Arithmetik wurde von MOSTOWSKI und TARSKI entwickelt (1939). J. ROBINSON hatte bei der Lektüre des Manuskriptes die Idee der Behandlung der rekursiven Funktionen ohne das Schema der primitiven Rekursion). Wichtige Beispiele für unentscheidbare Theorien neben dem Satz von CHURCH sind in historischer Reihenfolge: Die Arithmetik der natürlichen Zahlen (ROSSER, 1936), die Arithmetik der ganzen Zahlen (TARSKI, MOSTOWSKI, 1949), die Arithmetik der rationalen Zahlen und die Theorie der gordneten Körper (J. ROBINSON 1949), die Theorie der Gruppen und die Verbandstheorie (TARSKI 1949). Vielleicht ist es instruktiv, einige entscheidbare Theorien daneben zu stellen: Die Theorie der Addition der natürlichen Zahlen (PRESSBURGER, 1929), die der Multiplikation (MOSTOWSKI, 1952), die Theorie der abelschen Gruppen (SZMIELEW, 1949), der algebraisch abgeschlossenen Körper und die der Booleschen Algebren (TARSKI, 1949), die Theorie der linear geordneten Mengen (EHRENFEUCHT, 1959).

5. Mengenlehre

5.1 Kumulative Typenstrukturen

Die Mengenlehre kann man als einen Rahmen auffassen, innerhalb dessen man Mathematik begründen und betreiben kann. Wir wollen die Mengenlehre hier als eine formale Theorie der mathematischen Logik entwickeln. Zunächst ist es jedoch notwendig, sich eine inhaltliche Vorstellung des durch die Axiome zu beschreibenden Mengenbegriffs zu verschaffen. CANTOR gab 1895 die folgende Definition:

> Unter einer "Menge" verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die Elemente von M genannt werden) zu einem Ganzen.

Man kann versuchen, diese Definition wie folgt zu präzisieren. Sei V die Gesamtheit aller Objekte unserer Anschauung oder unseres Denkens. Mit $A(x)$ bezeichnen wir Eigenschaften von Objekten x aus V. Dann kann man $\{ x \mid A(x) \}$ bilden, die Menge aller Objekte x aus V mit der Eigenschaft $A(x)$. Aufgrund von CANTORs Definiton ist $\{ x \mid A(x) \}$ wieder ein Objekt aus V.

Beispiele für Eigenschaften: (1) x ist natürliche Zahl. (2) x ist Menge. (3) x ist Punkt, y ist Gerade und x liegt auf y. (4) y ist Menge und x ist Element von y, kurz: $\mathsf{Mg}(y) \wedge x \in y$.

CANTORs Definiton ist jedoch in ihrer ursprünglichen Form nicht haltbar, da sie zu Widersprüchen führt. Am bekanntesten ist die sogenannte RUSSELL*sche Antinomie*: Sei $x_0 := \{ x \mid \mathsf{Mg}(x) \wedge x \notin x \}$. Dann gilt

$$x_0 \in x_0 \leftrightarrow \mathsf{Mg}(x_0) \wedge x_0 \notin x_0 \leftrightarrow x_0 \notin x_0,$$

denn x_0 ist Menge.

Der Grund für diesen Widerspruch liegt darin, daß man von der Vorstellung einer fertigen Gesamtheit aller Mengen ausgeht. Dies ist aber weder notwendig noch entspricht es dem Vorgehen in der Mathematik. Es reicht vollkommen aus, wenn man eine Menge nur dann bildet, wenn ihre Elemente bereits "zur Verfügung stehen". Dies führt auf eine Vorstellung einer stufenweisen Konstruktion von Mengen oder genauer auf die *kumulative Typenstruktur*: Man beginnt mit vorgegebenen Urelementen, die die Mengen der Stufe 0 bilden. Auf einer beliebigen Stufe kann man dann alle Mengen bilden, deren Elemente früheren Stufen angehören.

Wählt man zum Beispiel als Urelemente die natürlichen Zahlen, so gehört $\{27, \{5\}\}$ zur Stufe 2.

Es stellen sich die folgenden natürlichen Fragen: (1) Welche Urelemente soll man wählen? (2) Wie weit reichen die Stufen?

Zu (1). Für die Zwecke der Mathematik ist es ausreichend, überhaupt keine Urelemente vorauszusetzen; man spricht dann von *reinen Mengen*. Dies wollen wir im folgenden tun.

$$\begin{array}{ll} \text{Stufe 0:} & - \\ \text{Stufe 1:} & \emptyset \\ \text{Stufe 2:} & \emptyset, \{\emptyset\} \\ \text{Stufe 3:} & \emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\} \\ & \text{und so weiter.} \end{array}$$

Zu (2). SHOENFIELD hat in [34] das folgende Prinzip formuliert.

SHOENFIELD-Prinzip. Man betrachte eine Gesamtheit $\mathcal{S}$ von Stufen. Kann man sich eine Situation vorstellen, in der alle Stufen aus $\mathcal{S}$ konstruiert sind, so soll es eine Stufe geben, die nach allen Stufen aus $\mathcal{S}$ kommt.

Aus diesem vagen Prinzip lassen sich exakte Folgerungen ziehen, die wir als Axiome fixieren werden.

Unter einer *Menge* wollen wir also inhaltlich ein Objekt verstehen, das zu einer Stufe der kumulative Typenstruktur gehört. Unter einer *Klasse* verstehen wir eine beliebige Gesamtheit von Mengen.

Jede Menge ist also eine Klasse. Ferner gibt es Klassen, die keine Mengen sind, z.B. die Klasse V aller Mengen.

5.2 Axiomatische Mengenlehre

Wie in jeder axiomatischen Theorie müssen wir auch in der Mengenlehre sämtliche benötigten Eigenschaften, auch die anscheinend selbstverständlichen, explizit durch Axiome angeben.

Die Sprache der Mengenlehre enthält als einziges nichtlogisches Symbol die Elementbeziehung $\in$. Atomare Formeln sind also nur $x \in y$ (x ist Element von y). Die Gleichheit $x = y$ wird definiert durch

$$x = y := \forall z. z \in x \leftrightarrow z \in y.$$

Um die Verträglichkeit der $\in$-Relation mit der Gleichheit sicherzustellen benötigen wir ein erstes Axiom.

Axiom 1. (Extensionalitätsaxiom).

$$\forall x, y, z. x = y \wedge x \in z \to y \in z.$$

Anmerkung. Will man die Gleichheit als Grundsymbol der Sprache verwenden, so muß man neben den Gleichheitsaxiomen fordern

$$\forall x, y. (\forall z. z \in x \leftrightarrow z \in y) \to x = y.$$

Als Klassen lassen wir in unserer axiomatischen Theorie nur *definierbare* Gesamtheiten von Mengen zu. Unter "definierbar" verstehen wir definierbar durch eine Formel der Sprache der Mengenlehre. Genauer: Ist $A(x)$ eine Formel, in der die hervorgehobene Mengenvariable x und eventuell weitere Mengenvariablen (sog. Parameter) vorkommen können, so heißt

$$\{ x \mid A(x) \}$$

die *Klasse* aller Mengen x mit der Eigenschaft $A(x)$.

Statt mit Klassen könnten wir also auch mit Eigenschaften oder genauer mit Formeln arbeiten. Mit Klassen lassen sich aber viele Aussagen einfacher und suggestiver formulieren.

Ist $A(x)$ die Formel $x = x$, so heißt $\{ x \mid A(x) \}$ die *Allklasse* oder das (mengentheoretische) *Universum*. Ist $A(x)$ die Formel $x \notin x$, so heißt $\{ x \mid A(x) \}$ die RUSSELL*klasse*.

Wir geben jetzt einige im folgenden ständig verwendete Definitionen. Eine Menge b ist genau dann Element der Klasse $\{ x \mid A(x) \}$, wenn $A(b)$ gilt:

$$b \in \{ x \mid A(x) \} := A(b).$$

Zwei Klassen $\mathcal{A}$, $\mathcal{B}$ heißen *gleich*, wenn sie dieselben Elemente enthalten:

$$\mathcal{A} = \mathcal{B} := \forall x. x \in \mathcal{A} \leftrightarrow x \in \mathcal{B}.$$

Ist $\mathcal{A}$ eine Klasse und b eine Menge, so heißen $\mathcal{A}$ und b gleich, wenn sie dieselben Elemente enthalten:

$$\mathcal{A} = b := \forall x. x \in \mathcal{A} \leftrightarrow x \in b.$$

Wir identifizieren dann die Klasse $\mathcal{A}$ mit dieser Menge b. Statt "$\mathcal{A}$ ist Menge" kann man auch schreiben $\mathcal{A} \in V$. Eine Klasse $\mathcal{B}$ ist Element einer Menge a (bzw. einer Klasse $\mathcal{A}$), wenn $\mathcal{B}$ gleich einem Element x von a (bzw. von $\mathcal{A}$) ist:

$$\mathcal{B} \in a := \exists x. x \in a \wedge \mathcal{B} = x,$$
$$\mathcal{B} \in \mathcal{A} := \exists x. x \in \mathcal{A} \wedge \mathcal{B} = x.$$

Eine Klasse $\mathcal{A}$ heißt *echte Klasse*, wenn $\mathcal{A}$ keine Menge ist:

$$\mathcal{A} \text{ echte Klasse} := \forall x (x \neq \mathcal{A}).$$

Anmerkung. Jede Menge b ist eine Klasse, denn

$$b = \{ x \mid x \in b \}.$$

Die RUSSELLKLASSE ist eine echte Klasse, denn wäre $\{ x \mid x \notin x \} = x_0$, so hätte man

$$x_0 \in x_0 \leftrightarrow x_0 \notin x_0.$$

Die RUSSEL-Konstruktion ist also nun keine Antinomie mehr, sondern besagt einfach: Es gibt Mengen und (echte) Klassen.

Seien $\mathcal{A}$, $\mathcal{B}$ Klassen (also echte Klassen oder Mengen) und $a, b, a_1, \dots, a_n$ Mengen. Wir definieren

$\{a_1, \dots, a_n\} := \{ x \mid x = a_1 \vee \dots \vee x = a_n \}$,	
$\emptyset := \{ x \mid x \neq x \}$	leere Klasse,
$V := \{ x \mid x = x \}$	Allklasse,
$\mathcal{A} \subseteq \mathcal{B} := \forall x.x \in \mathcal{A} \to x \in \mathcal{B}$	$\mathcal{A}$ ist Teilklasse von $\mathcal{B}$,
$\mathcal{A} \subsetneq \mathcal{B} := \mathcal{A} \subseteq \mathcal{B} \wedge \mathcal{A} \neq \mathcal{B}$	$\mathcal{A}$ ist echte Teilklasse von $\mathcal{B}$,
$\mathcal{A} \cap \mathcal{B} := \{ x \mid x \in \mathcal{A} \wedge x \in \mathcal{B} \}$	Durchschnitt,
$\mathcal{A} \cup \mathcal{B} := \{ x \mid x \in \mathcal{A} \vee x \in \mathcal{B} \}$	Vereinigung,
$\mathcal{A} \setminus \mathcal{B} := \{ x \mid x \in \mathcal{A} \wedge x \notin \mathcal{B} \}$	Differenz,
$\bigcup \mathcal{A} := \{ x \mid \exists y.y \in \mathcal{A} \wedge x \in y \}$	große Vereinigung,
$\bigcap \mathcal{A} := \{ x \mid \forall y.y \in \mathcal{A} \to x \in y \}$	großer Durchschnitt,
$\mathcal{P}(\mathcal{A}) := \{ x \mid x \subseteq \mathcal{A} \}$	Potenzklasse von $\mathcal{A}$,
$(a, b) := \{ x \mid x = \{a\} \vee x = \{ a, b \mid \} \}$	(geordnetes) Paar (KURATOWSKI-Paar.)

Insbesondere ist $a \cup b = \bigcup\{a, b\}$ und $a \cap b = \bigcap\{a, b\}$, und $\bigcap \emptyset$ ist die Allklasse. Ferner ist $\mathcal{P}(\mathcal{A})$ die Klasse aller Teilklassen von $\mathcal{A}$, die Mengen sind.

Um sicherzustellen, daß (a, b) nicht die leere Klasse ist, müssen wir axiomatisch fordern, daß $\{a\}$ und $\{a, b\}$ Mengen sind. Also

Axiom 2. (Paarmengenaxiom).

$$\forall x, y(\{x, y\} \text{ ist Menge}).$$

In der kumulativen Typenstruktur ist das Paarmengenaxiom offenbar gültig, da es zu je zwei Stufen S_1 und S_2 nach dem SHOENFIELD-Prinzip eine Stufe S geben muß, die nach S_1 und S_2 kommt.

Ausgeschrieben lautet das Paarmengenaxiom: $\forall x \forall y \exists z \forall u.u \in z \leftrightarrow u = x \vee u = y$. Insbesondere folgt aus dem Paarmengenaxiom, daß für jede Menge x die Einermenge $\{x\}$ eine Menge ist. Ferner folgt, daß $(a, b) = \{\{a\}, \{a, b\}\}$ Menge ist.

Weiter definieren wir

$\{ (x, y) \mid A(x, y) \} := \{ z \mid \exists x, y.A(x, y) \wedge z = (x, y) \}$	
$\mathcal{A} \times \mathcal{B} := \{ (x, y) \mid x \in \mathcal{A} \wedge y \in \mathcal{B} \}$	kartesisches Produkt von $\mathcal{A}$ und $\mathcal{B}$,
$\mathsf{dom}(\mathcal{A}) := \{ x \mid \exists y\, ((x, y) \in \mathcal{A}) \}$	Definitionsbereich von $\mathcal{A}$,
$\mathsf{rng}(\mathcal{A}) := \{ y \mid \exists x\, ((x, y) \in \mathcal{A}) \}$	Wertebereich von $\mathcal{A}$,
$\mathcal{A} \restriction \mathcal{B} := \{ (x, y) \mid (x, y) \in \mathcal{A} \wedge x \in \mathcal{B} \}$	Einschränkung von $\mathcal{A}$ auf $\mathcal{B}$,
$\mathcal{A}[\mathcal{B}] := \{ y \mid \exists x.x \in \mathcal{B} \wedge (x, y) \in \mathcal{A} \}$	Bild von $\mathcal{B}$ unter $\mathcal{A}$,
$\mathcal{A}^{-1} := \{ (y, x) \mid (x, y) \in \mathcal{A} \}$,	Inverses von $\mathcal{A}$,
$\mathcal{A} \circ \mathcal{B} := \{ (x, z) \mid \exists y.(x, y) \in \mathcal{B} \wedge (y, z) \in \mathcal{A} \}$	Verkettung von $\mathcal{A}$ und $\mathcal{B}$.

Jetzt kann man ohne Mühe die üblichen Begriffe betreffend Relationen und Funktionen einführen. Für Klassen $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$ definieren wir

1. $\mathcal{A}$ ist *Relation* genau dann, wenn $\mathcal{A} \subseteq V \times V$. Eine Relation ist also eine Klasse von Paaren. Statt $(a,b) \in \mathcal{A}$ schreiben wir auch $a\mathcal{A}b$.
2. $\mathcal{A}$ ist *Relation auf* $\mathcal{B}$ genau dann, wenn $\mathcal{A} \subseteq \mathcal{B} \times \mathcal{B}$.
3. $\mathcal{A}$ ist *Funktion* genau dann, wenn $\mathcal{A}$ Relation ist und

$$\forall x, y, z.(x,y) \in \mathcal{A} \wedge (x,z) \in \mathcal{A} \to y = z.$$

   Eine Funktion ist also eine rechtseindeutige Relation.
4. $\mathcal{A}\colon \mathcal{B} \to \mathcal{C}$ genau dann, wenn $\mathcal{A}$ Funktion ist mit $\mathsf{dom}(\mathcal{A}) = \mathcal{B}$ und $\mathcal{A}[\mathcal{B}] \subseteq \mathcal{C}$. $\mathcal{A}$ heißt dann *Funktion von* $\mathcal{B}$ *nach* $\mathcal{C}$.
5. $\mathcal{A}\colon \mathcal{B} \to_{\text{auf}} \mathcal{C}$ genau dann, wenn $\mathcal{A}\colon \mathcal{B} \to \mathcal{C}$ und $\mathcal{A}[\mathcal{B}] = \mathcal{C}$. $\mathcal{A}$ heißt dann *surjektive Funktion von* $\mathcal{B}$ *auf* $\mathcal{C}$.
6. $\mathcal{A}$ ist *injektiv* genau dann, wenn $\mathcal{A}$ und $\mathcal{A}^{-1}$ Funktionen sind.
7. $\mathcal{A}\colon \mathcal{B} \leftrightarrow \mathcal{C}$ genau dann, wenn $\mathcal{A}\colon \mathcal{B} \to_{\text{auf}} \mathcal{C}$ und $\mathcal{A}$ injektiv ist. $\mathcal{A}$ heißt dann *bijektive Funktion von* $\mathcal{B}$ *auf* $\mathcal{C}$.

Für den weiteren Aufbau der Mengenlehre sind die folgenden Axiome notwendig.

Axiom 3. (Vereinigungsmengenaxiom).

$$\forall x\, (\bigcup x \text{ ist Menge}).$$

Das Vereinigungsmengenaxiom ist gültig in der kumulativen Typenstruktur. Um dies zu sehen, betrachte man eine Stufe S, in der x gebildet ist. Ein beliebiges Element $v \in x$ steht dann bereits in einer früheren Stufe S_v zur Verfügung. Ebenso ist jedes Element $u \in v$ in einer vor S_v liegenden Stufe $S_{v,u}$ vorhanden. Alle diese u bilden aber $\bigcup x$. Damit kann auch $\bigcup x$ auf der Stufe S gebildet werden.

Wir können jetzt die obige Definition fortsetzen durch

8. $\mathcal{A}(x) := \bigcup\{\, y \mid (x,y) \in \mathcal{A}\,\}$.

Ist $\mathcal{A}$ eine Funktion und $(x,y) \in \mathcal{A}$, so ist $\mathcal{A}(x) = \bigcup\{y\} = y$ und wir schreiben $\mathcal{A}\colon x \mapsto y$.

Axiom 4. (Aussonderungsschema). Für jede Klasse $\mathcal{A}$ gilt (die Generalisierung von)

$$\mathcal{A} \subseteq x \to \exists y\, (\mathcal{A} = y).$$

Das Aussonderungsschema sagt also aus, daß jede Teilklasse $\mathcal{A}$ einer Menge x selbst eine Menge ist. Es ist gültig in der kumulativen Typenstruktur, denn auf derselben Stufe, auf der die Menge x gebildet ist, kann man auch die Menge y bilden, deren Elemente gerade die Elemente der Klasse $\mathcal{A}$ sind. – Man beachte, daß das Aussonderungsschema aus unendlich vielen Axiomen besteht.

Axiom 5. (Potenzmengenaxiom).

$$\forall x\, (\mathcal{P}(x) \text{ ist Menge}).$$

Das Potenzmengenaxiom ist gültig in der kumulativen Typenstruktur. Um dies zu sehen, betrachte man eine Stufe S, auf der x gebildet ist. Dann ist auch jede Teilmenge $y \subseteq x$ auf der Stufe S gebildet. Auf der (nach dem Shoenfield-Prinzip existierenden) nächsten Stufe S' kann man also $\mathcal{P}(x)$ bilden.

Lemma 5.2.1. $\forall a, b\, (a \times b$ *ist Menge*)..

Beweis. Wir zeigen $a \times b \subseteq \mathcal{P}(\mathcal{P}(a \cup b))$. Seien also $x \in a$ und $y \in b$. Dann gilt

$$\begin{aligned} \{x\}, \{x,y\} &\subseteq a \cup b \\ \{x\}, \{x,y\} &\in \mathcal{P}(a \cup b) \\ \{\, \{x\}, \{x,y\}\,\} &\subseteq \mathcal{P}(a \cup b) \\ (x,y) = \{\, \{x\}, \{x,y\}\,\} &\in \mathcal{P}(\mathcal{P}(a \cup b)) \end{aligned}$$

Die Behauptung folgt mit dem Vereinigungsmengenaxiom, dem Paarmengenaxiom, dem Potenzmengenaxiom und dem Aussonderungsschema. □

Axiom 6. (Ersetzungsschema). Für jede Klasse $\mathcal{A}$ gilt

$$\mathcal{A} \text{ ist Funktion } \to \forall x\, (\mathcal{A}[x] \text{ ist Menge})$$

Auch das Ersetzungsschema ist gültig in der kumulativen Typenstruktur; dies ist allerdings nur mit etwas mehr Mühe einzusehen. Man betrachte alle Elemente u der Menge $x \cap \mathsf{dom}(\mathcal{A})$. Für jedes solche u ist $\mathcal{A}(u)$ eine Menge und deshalb auf einer Stufe S_u der kumulativen Typenstruktur gebildet. Da $x \cap \mathsf{dom}(\mathcal{A})$ eine Menge ist, kann man sich eine Situation vorstellen, in der alle S_u für $u \in x \cap \mathsf{dom}(\mathcal{A})$ konstruiert sind. Nach dem Shoenfield-Prinzip gibt es also auch eine Stufe S nach allen S_u. In S kann $\mathcal{A}[x]$ gebildet werden.

Lemma 5.2.2. *Aus dem Ersetzungsschema folgt das Aussonderungsschema.*

Beweis. Sei $\mathcal{A} \subseteq x$ und $\mathcal{B} := \{\, (u, v) \mid u = v \wedge u \in \mathcal{A} \,\}$. Dann ist $\mathcal{B}$ Funktion und es gilt $\mathcal{B}[x] = \mathcal{A}$. □

Damit sind die Axiome der Mengenlehre noch nicht vollständig angegeben: wir werden später noch das Unendlichkeitsaxiom, das Regularitätsaxiom und das Auswahlaxiom fordern.

5.3 Rekursion, Induktion, Ordinalzahlen

Wir beginnen damit, einen möglichst allgemeinen Rahmen für rekursive Definitionen und induktive Beweise abzustecken. Es wird sich zeigen, daß beides über sog. fundierten Relationen möglich ist. Um dies zu zeigen, führen wir als Hilfsbegriff den einer transitiv fundierten Relation ein; er wird sich später als äquivalent zu dem einer fundierten Relation erweisen. Wir definieren dann die natürlichen Zahlen im Rahmen der Mengenlehre, wobei wir Induktion und Rekursion als Spezialfälle der entsprechenden allgemeinen Sätze für transitiv fundierte Relationen erhalten. Durch Rekursion über natürliche Zahlen können wir dann die transitive Hülle einer Menge einführen, und mit Hilfe dieses Begriffs zeigen wir, daß die fundierten Relationen mit den transitiv fundierten Relationen übereinstimmen.

Anschließend untersuchen wir spezielle fundierte Relationen. Zunächst zeigen wir, daß beliebige Klassen mit der $\in$-Relation bis auf Isomorphie die einzigen fundierten extensionalen Relationen sind (Isomorphiesatz von Mostowski). Dann betrachten wir lineare fundierte Ordnungen, kurz Wohlordnungen. Da sie stets extensional sind, sind sie alle zu gewissen Klassen mit der $\in$-Relation isomorph, die wir ordinale Klassen nennen. Ordinalzahlen sind dann die ordinalen Mengen.

Rekursion über transitiv fundierte Relationen

Mit $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ bezeichnen wir immer Klassen. Für eine beliebige Relation $\mathcal{R}$ auf $\mathcal{A}$ definieren wir

1. $\hat{x}^{\mathcal{R}} = \{\, y \mid y\mathcal{R}x \,\}$ heißt die Klasse der *$\mathcal{R}$-Vorgänger* von x. Im folgenden schreiben wir $\hat{x}$ statt $\hat{x}^{\mathcal{R}}$, wenn $\mathcal{R}$ aus dem Zusammenhang klar ist.
2. $\mathcal{B} \subseteq \mathcal{A}$ heißt *$\mathcal{R}$-transitiv*, wenn
$$\forall x . x \in \mathcal{B} \to \hat{x} \subseteq \mathcal{B}.$$
$\mathcal{B} \subseteq \mathcal{A}$ ist also $\mathcal{R}$-transitiv genau dann, wenn aus $y\mathcal{R}x$ und $x \in \mathcal{B}$ stets folgt $y \in \mathcal{B}$.
3. Sei $\mathcal{B} \subseteq \mathcal{A}$. $x \in \mathcal{B}$ heißt *$\mathcal{R}$-minimales Element* von $\mathcal{B}$, wenn $\hat{x} \cap \mathcal{B} = \emptyset$.
4. $\mathcal{R}$ heißt *transitiv fundierte Relation* auf $\mathcal{A}$, wenn gilt
   a. Jede nichtleere Teilmenge von $\mathcal{A}$ besitzt ein $\mathcal{R}$-minimales Element, d.h.
   $$\forall a . a \subseteq \mathcal{A} \wedge a \neq \emptyset \to \exists x . x \in a \wedge \hat{x} \cap a = \emptyset.$$
   b. Zu jedem $x \in \mathcal{A}$ gibt es eine $\mathcal{R}$-transitive Menge $b \subseteq \mathcal{A}$ mit $\hat{x} \subseteq b$.

Bei allen diesen Begriffen wird im folgenden $\mathcal{R}$ weggelassen, wie bereits oben für $\hat{x}$ geschehen.

Anmerkung. Sei $\mathcal{R}$ eine Relation auf $\mathcal{A}$. $\mathcal{R}$ heißt *transitive Relation auf* $\mathcal{A}$, wenn für alle $x, y, z \in \mathcal{A}$ gilt

$$x\mathcal{R}y \wedge y\mathcal{R}z \to x\mathcal{R}z.$$

Zum Begriff der $\mathcal{R}$-Transitivität von Klassen besteht folgender Zusammenhang. Sei $\mathcal{R}$ eine Relation auf $\mathcal{A}$. Dann gilt

$$\mathcal{R} \text{ ist transitive Relation auf } \mathcal{A} \leftrightarrow \text{für jedes } y \in \mathcal{A} \text{ gilt: } y \text{ ist } \mathcal{R}\text{-transitiv.}$$

Beweis. : $\to$. Sei $\mathcal{R}$ transitive Relation auf $\mathcal{A}$, $y \in \mathcal{A}$ und $x \in \hat{y}$, also $x\mathcal{R}y$. Zu zeigen ist $\hat{x} \subseteq \hat{y}$. Sei also $z\mathcal{R}x$. Zu zeigen ist $z\mathcal{R}y$. Dies folgt aber aus der Transitivität von $\mathcal{R}$. $\leftarrow$. Seien $x, y, z \in \mathcal{A}$, $x\mathcal{R}y$ und $y\mathcal{R}z$. Zu zeigen ist $x\mathcal{R}z$. Es gilt also $x\mathcal{R}y$ und $y \in \hat{z}$. Da $\hat{z}$ $\mathcal{R}$-transitiv ist, folgt $x \in \hat{z}$, also $x\mathcal{R}z$. □

Lemma 5.3.1. *Sei $\mathcal{R}$ transitiv fundierte Relation auf $\mathcal{A}$. Dann gilt*

1. *Jede nichtleere Teilklasse $\mathcal{B} \subseteq \mathcal{A}$ hat ein $\mathcal{R}$-minimales Element.*
2. *$\forall x . x {\in} \mathcal{A} \to \hat{x}$ ist Menge.*

Beweis. 1. Sei $\mathcal{B} \subseteq \mathcal{A}$ und $z \in \mathcal{B}$. OBdA ist z nicht $\mathcal{B}$-minimal, d.h. $\hat{z} \cap \mathcal{B} \neq \emptyset$. Nach 4b existiert eine $\mathcal{R}$-transitive Obermenge $b \subseteq \mathcal{A}$ von $\hat{z}$. Wegen $\hat{z} \cap \mathcal{B} \neq \emptyset$ ist $b \cap \mathcal{B} \neq \emptyset$. Nach 4a existiert ein $\mathcal{R}$-minimales $x \in b \cap \mathcal{B}$, d.h. $\hat{x} \cap b \cap \mathcal{B} = \emptyset$. Da b $\mathcal{R}$-transitiv ist, folgt aus $x \in b$ sofort $\hat{x} \subseteq b$. Also ist $\hat{x} \cap \mathcal{B} = \emptyset$ und damit x $\mathcal{R}$-minimales Element von $\mathcal{B}$.

2. Dies folgt wegen 4b aus dem Aussonderungsschema. □

Wir schreiben $\forall x{\in}\mathcal{A} \dots$ für $\forall x . x \in \mathcal{A} \to \dots$ und entsprechend $\exists x{\in}\mathcal{A} \dots$ für $\exists x . x \in \mathcal{A} \wedge \dots$.

Satz 5.3.2. *(Induktionssatz). Sei $\mathcal{R}$ eine transitiv fundierte Relation auf $\mathcal{A}$ und $\mathcal{B}$ eine beliebige Klasse. Gilt dann*

$$\forall x{\in}\mathcal{A} . \hat{x} \subseteq \mathcal{B} \to x \in \mathcal{B},$$

so folgt $\mathcal{A} \subseteq \mathcal{B}$.

Beweis. Annahme: $\mathcal{A} \setminus \mathcal{B} \neq \emptyset$. Sei x minimales Element von $\mathcal{A} \setminus \mathcal{B}$. Es genügt, $\hat{x} \subseteq \mathcal{B}$ zu zeigen, denn daraus folgt nach Annahme $x \in \mathcal{B}$, also ein Widerspruch. Sei $z \in \hat{x}$. Nach Wahl von x folgt $z \notin \mathcal{A} \setminus \mathcal{B}$, also $z \in \mathcal{B}$ (denn $z \in \mathcal{A}$ gilt, da $\mathcal{R}$ eine Relation auf $\mathcal{A}$ ist). □

Satz 5.3.3. *(Rekursionssatz). Sei $\mathcal{R}$ eine transitiv fundierte Relation auf $\mathcal{A}$ und $\mathcal{G}\colon V \to V$. Dann gibt es genau eine Funktion $\mathcal{F}\colon \mathcal{A} \to V$ mit*

$$\forall x{\in}\mathcal{A}\big(\mathcal{F}(x) = \mathcal{G}(\mathcal{F}{\restriction}\hat{x})\big).$$

Beweis. Man beachte zunächst, daß für $\mathcal{F}\colon \mathcal{A} \to V$ gilt $\mathcal{F}{\restriction}\hat{x} \subseteq \hat{x} \times \mathcal{F}[\hat{x}]$ und damit $\mathcal{F}{\restriction}\hat{x}$ eine Menge ist.

Eindeutigkeit. Gegeben $\mathcal{F}_1$, $\mathcal{F}_2$. Betrachte

$$\{\, x \mid x \in \mathcal{A} \wedge \mathcal{F}_1(x) = \mathcal{F}_2(x) \,\} =: \mathcal{B}.$$

Nach dem Induktionssatz genügt es zu zeigen $\forall x{\in}\mathcal{A} . \hat{x} \subseteq \mathcal{B} \to x \in \mathcal{B}$. Sei also $x \in \mathcal{A}$ und $\hat{x} \subseteq \mathcal{B}$. Dann gilt

$$\begin{aligned}
&\mathcal{F}_1{\restriction}\hat{x} = \mathcal{F}_2{\restriction}\hat{x} \\
&\mathcal{G}\left(\mathcal{F}_1{\restriction}\hat{x}\right) = \mathcal{G}\left(\mathcal{F}_2{\restriction}\hat{x}\right) \\
&\mathcal{F}_1(x) = \mathcal{F}_2(x) \\
&x \in \mathcal{B}.
\end{aligned}$$

Existenz. Sei

$$\mathcal{B} := \{\, f \mid f \text{ Funktion, } \mathsf{dom}(f)\ \mathcal{R}\text{-transitive Teilmenge von } \mathcal{A} \text{ und } \forall x{\in}\mathsf{dom}(f)\,\big(f(x) = \mathcal{G}(f{\restriction}\hat{x})\big) \,\}$$

und

$$\mathcal{F} := \bigcup \mathcal{B}.$$

Wir beweisen zunächst, daß

$$f, g \in \mathcal{B} \wedge x \in \mathsf{dom}(f) \cap \mathsf{dom}(g) \to f(x) = g(x).$$

Seien also $f, g \in \mathcal{B}$. Wir zeigen die Behauptung durch Induktion über x, d.h. durch Anwendung des Induktionssatzes auf

$$\{\, x \mid x \in \mathsf{dom}(f) \cap \mathsf{dom}(g) \to f(x) = g(x) \,\}.$$

Sei also $x \in \mathsf{dom}(f) \cap \mathsf{dom}(g)$. Dann gilt

$$\begin{aligned}
&\hat{x} \subseteq \mathsf{dom}(f) \cap \mathsf{dom}(g), && \text{da } \mathsf{dom}(f), \mathsf{dom}(g)\ \mathcal{R}\text{-transitiv} \\
&f{\restriction}\hat{x} = g{\restriction}\hat{x} && \text{nach IH} \\
&\mathcal{G}\,(f{\restriction}\hat{x}) = \mathcal{G}\,(g{\restriction}\hat{x}) \\
&f(x) = g(x).
\end{aligned}$$

Damit ist gezeigt, daß $\mathcal{F}$ eine Funktion ist.

Hieraus ergibt sich unmittelbar $f \in \mathcal{B} \wedge x \in \mathsf{dom}(f) \to \mathcal{F}(x) = f(x)$; wir haben also gezeigt, daß

$$\mathcal{F}(x) = \mathcal{G}\,(\mathcal{F}{\restriction}\hat{x}) \quad \text{für alle } x \in \mathsf{dom}(\mathcal{F}). \tag{5.1}$$

Wir beweisen jetzt

$$\mathsf{dom}(\mathcal{F}) = \mathcal{A}.$$

$\subseteq$ ist klar. $\supseteq$. Wir verwenden den Induktionssatz. Sei also $\hat{y} \subseteq \mathsf{dom}(\mathcal{F})$. Zu zeigen ist $y \in \mathsf{dom}(\mathcal{F})$. Wir führen den Beweis indirekt und nehmen $y \notin \mathsf{dom}(\mathcal{F})$ an. Sei b $\mathcal{R}$-transitiv mit $\hat{y} \subseteq b \subseteq \mathcal{A}$. Setze

$$g := \mathcal{F}{\restriction}b \cup \{\, (y, \mathcal{G}(\mathcal{F}{\restriction}\hat{y})) \,\}.$$

Es genügt offenbar, $g \in \mathcal{B}$ zu zeigen, denn wegen $y \in \mathsf{dom}(g)$ folgt daraus $y \in \mathsf{dom}(\mathcal{F})$ und damit der gesuchte Widerspruch.

g Funktion: Dies ist klar, da $y \notin \mathsf{dom}(\mathcal{F})$ nach Annahme.

$\mathsf{dom}(g)$ $\mathcal{R}$-transitiv: Es ist $\mathsf{dom}(g) = (b \cap \mathsf{dom}(\mathcal{F})) \cup \{y\}$. Man beachte zunächst, daß $\mathsf{dom}(\mathcal{F})$ als Vereinigung $\mathcal{R}$-transitiver Mengen selbst $\mathcal{R}$-transitiv ist. Da ferner b $\mathcal{R}$-transitiv ist, ist auch $b \cap \mathsf{dom}(\mathcal{F})$ $\mathcal{R}$-transitiv. Sei nun $z\mathcal{R}x$ und $x \in \mathsf{dom}(g)$. Zu zeigen: $z \in \mathsf{dom}(g)$. Im Fall $x \in b \cap \mathsf{dom}(\mathcal{F})$ ist auch $z \in b \cap \mathsf{dom}(\mathcal{F})$ (da wie eben bemerkt $b \cap \mathsf{dom}(\mathcal{F})$ $\mathcal{R}$-transitiv ist), also $z \in \mathsf{dom}(g)$. Im Fall $x = y$ ist $z \in \hat{y}$, also $z \in b$ und $z \in \mathsf{dom}(\mathcal{F})$ nach Wahl von y, also wieder $z \in \mathsf{dom}(g)$.

$\forall x{\in}\mathsf{dom}(g)\,\big(g(x) = \mathcal{G}(g{\restriction}\hat{x})\big)$: Im Fall $x \in b \cap \mathsf{dom}(\mathcal{F})$ ist

$$\begin{aligned}
g(x) &= \mathcal{F}(x) \\
&= \mathcal{G}(\mathcal{F}{\restriction}\hat{x}) && \text{nach (5.1)} \\
&= \mathcal{G}(g{\restriction}\hat{x}) && \text{da } \hat{x} \subseteq b \cap \mathsf{dom}(\mathcal{F}), \text{ denn } b \cap \mathsf{dom}(\mathcal{F}) \text{ ist } \mathcal{R}\text{-transitiv.}
\end{aligned}$$

Im Fall $x = y$ ist $g(x) = \mathcal{G}(\mathcal{F}{\restriction}\hat{x}) = \mathcal{G}(g{\restriction}\hat{x})$, da $\hat{x} = \hat{y} \subseteq b \cap \mathsf{dom}(\mathcal{F})$ nach Wahl von y. □

Natürliche Zahlen

Zermelo hat die natürlichen Zahlen wie folgt in der Mengenlehre definiert: $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{\{\emptyset\}\}$, $3 = \{\{\{\emptyset\}\}\}$ und so weiter. Ein gravierender Nachteil dieser Definition ist, daß man sie nicht ins Transfinite verallgemeinern kann. Deshalb hat J. von Neumann vorgeschlagen, die Zahl n darzustellen durch eine Menge aus genau n Elementen, und zwar durch

$$n := \{\, 0, 1, \dots, n-1 \,\}.$$

Also

$$\begin{aligned}
0 &= \emptyset \\
n+1 &= \{\, 0, 1, \dots, n \,\} \\
&= \{\, 0, 1, \dots, n-1 \,\} \cup \{n\}
\end{aligned}$$

Wir definieren deshalb allgemein

$$0 := \emptyset, \quad x + 1 := x \cup \{x\}.$$

Speziell setzen wir $1 := 0 + 1$, $2 := 1 + 1$, $3 := 2 + 1$.

Die Klasse aller so konstruierten natürlichen Zahlen soll eine Menge sein. Dazu brauchen wir ein weiteres Axiom.

Axiom 7. (Unendlichkeitsaxiom).

$$\exists x.\emptyset \in x \wedge \forall y.y \in x \to y \cup \{y\} \in x.$$

Wir nennen eine Klasse $\mathcal{A}$ *induktiv*, wenn

$$\emptyset \in \mathcal{A} \wedge \forall y.y \in \mathcal{A} \to y \cup \{y\} \in \mathcal{A}.$$

Das Unendlichkeitsaxiom sagt also aus: Es gibt eine induktive Menge. Ferner setzen wir

$$\omega := \bigcap \{\, x \mid x \text{ ist induktiv} \,\}.$$

Offenbar ist ω ist eine Menge und es gilt $0 \in \omega$ und $y \in \omega \to y + 1 \in \omega$. ω heißt die Menge der *natürlichen Zahlen*.

Mit n, m bezeichnen wir natürliche Zahlen. Insbesondere steht $\forall n\, A(n)$ für $\forall x.x \in \omega \to A(x)$, ebenso $\exists n\, A(n)$ für $\exists x.x \in \omega \wedge A(x)$ und $\{\, n \mid A(n) \,\}$ für $\{\, x \mid x \in \omega \wedge A(x) \,\}$.

Satz 5.3.4. *(Induktion über ω).*

1. $x \subseteq \omega \wedge 0 \in x \wedge (\forall n.n \in x \to n + 1 \in x) \to x = \omega$.
2. *Für jede Formel $A(x)$ gilt*

$$A(0) \wedge (\forall n.A(n) \to A(n+1)) \to \forall n A(n).$$

Beweis. 1. x ist induktiv, also $\omega \subseteq x$. 2. Sei $\mathcal{A} := \{\, n \mid A(n) \,\}$. Dann gilt $\mathcal{A} \subseteq \omega$ (also $\mathcal{A}$ ist Menge),

$$\begin{aligned} &0 \in \mathcal{A}, \\ &n \in \mathcal{A} \to n + 1 \in \mathcal{A}. \end{aligned}$$

Nach Teil 1 folgt $\mathcal{A} = \omega$. □

Wir zeigen jetzt, daß für natürliche Zahlen die Relation $\in$ die Eigenschaften von $<$ und die Relation $\subseteq$ die Eigenschaften von $\leq$ hat.

Eine Klasse $\mathcal{A}$ heißt *transitiv*, wenn $\mathcal{A}$ bezüglich der speziellen Relation $\mathcal{E} := \{\, (x, y) \mid x \in y \,\}$ auf V $\mathcal{E}$-transitiv ist, d.h. wenn gilt $\forall x.x \in \mathcal{A} \to x \subseteq \mathcal{A}$. $\mathcal{A}$ ist also transitiv genau dann, wenn

$$\forall x, y.y \in x \in \mathcal{A} \to y \in \mathcal{A}.$$

Lemma 5.3.5. *1. n ist transitiv.*
2. ω ist transitiv.

Beweis. 1. Induktion nach n. 0 ist transitiv. $n \to n + 1$. Nach IH ist n transitiv. Zu zeigen: $n + 1$ ist transitiv. Sei

$$\begin{aligned} &y \in x \in n + 1 \\ &y \in x \in n \cup \{n\} \\ &y \in x \in n \vee y \in x = n \\ &y \in n \vee y \in n \\ &y \in n \cup \{n\} = n + 1. \end{aligned}$$

2. Gezeigt wird $\forall x.x \in n \to x \in \omega$ durch Induktion nach n. 0: Klar. $n \to n+1$. Nach IH gilt $\forall x.x \in n \to x \in \omega$. Sei

$$
\begin{aligned}
&x \in n+1 \\
&x \in n \vee x = n \\
&x \in \omega.
\end{aligned}
$$

□

Lemma 5.3.6. *$n \notin n$.*

Beweis. Induktion nach n. 0. Klar. $n \to n+1$: Nach IH ist $n \notin n$. Annahme:

$$
\begin{aligned}
&n+1 \in n+1 \\
&n+1 \in n \vee n+1 = n \\
&n \in n+1 \in n \vee n \in n+1 = n \\
&n \in n && \text{denn } n \text{ ist nach Lemma 5.3.5(1)transitiv.}
\end{aligned}
$$

Dies ist ein Widerspruch zur IH. □

Lemma 5.3.7. *1. $n \subseteq m+1 \leftrightarrow n \subseteq m \vee n = m+1$.*
2. $n \subseteq m \leftrightarrow n \in m \vee n = m$.
3. $n \subseteq m \vee m \subseteq n$.
4. $n \in m \vee n = m \vee m \in n$.

Beweis. 1. $\leftarrow$ folgt aus $m \subseteq m+1$. $\to$. *Fall* $m \in n$. Wir zeigen $n = m+1$. $\subseteq$ gilt nach Voraussetzung. $\supseteq$.

$$
\begin{aligned}
&p \in m+1 \\
&p \in m \vee p = m \\
&p \in n.
\end{aligned}
$$

Fall $m \notin n$. Wir zeigen $n \subseteq m$.

$$
\begin{aligned}
&p \in n \\
&p \in m+1 \\
&p \in m \vee p = m,
\end{aligned}
$$

aber $p = m$ ist unmöglich wegen $m \notin n$.

2. $\leftarrow$ folgt aus der Transitivität von m. $\to$. Induktion nach m. 0. Klar. $m \to m+1$.

$$
\begin{aligned}
&n \subseteq m+1 \\
&n \subseteq m \vee n = m+1 && \text{nach (1)} \\
&n \in m \vee n = m \vee n = m+1 && \text{nach IH} \\
&n \in m+1 \vee n = m+1.
\end{aligned}
$$

3. Induktion nach n. 0. Klar. $n \to n+1$: *Fall* $m \subseteq n$. Klar. *Fall* $n \subseteq m$. Dann gilt

$$
\begin{aligned}
&n \in m \vee n = m && \text{nach (2)} \\
&n, \{n\} \subseteq m \vee m \subseteq n+1 \\
&n+1 \subseteq m \vee m \subseteq n+1
\end{aligned}
$$

4. Folgt aus (3) und (2). □

Satz 5.3.8. *(Peano-Axiome).*

1. *$n+1 \neq \emptyset$.*
2. *$n+1 = m+1 \to n = m$.*
3. *$x \subseteq \omega \wedge 0 \in x \wedge (\forall n.n \in x \to n+1 \in x) \to x = \omega$.*

Beweis. 1. Klar. 3. Dies wurde eben als Satz 5.3.4(1) bewiesen. 2.

$$\begin{aligned}
&n+1 = m+1\\
&n \in m+1 \wedge m \in n+1\\
&(n \in m \wedge m \in n) \vee n = m\\
&n \in n \vee n = m\\
&n = m.
\end{aligned}$$
□

Wir behandeln jetzt noch verschiedene Formen der Induktion.

Satz 5.3.9. *(Induktion über ω mit Rückgriff auf sämtliche Vorgänger).*

1. $x \subseteq \omega \wedge [\forall n.(\forall m.m \in n \to m \in x) \to n \in x] \to x = \omega$.
2. $[\forall n.(\forall m.m \in n \to A(m)) \to A(n)] \to \forall n A(n)$.

Beweis. 2. Gelte $\forall n.(\forall m.m \in n \to A(m)) \to A(n)$; man sagt in diesem Fall, daß $A(n)$ *progressiv* ist. Gezeigt wird $\forall m.m \in n \to A(m)$ durch Induktion nach n. 0. Klar. $n \to n+1$. Nach IH gilt $\forall m.m \in n \to A(m)$. Sei also $m \in n+1$. Dann gilt $m \in n \vee m = n$. Im Fall $m \in n$ folgt $A(m)$ nach IH, und im Fall $m = n$ folgt $A(n)$ aus der Progressivität von A mit der IH.

1. Aus (2) mit $A(y) := y \in x$. □

Satz 5.3.10. *(Prinzip vom kleinsten Element für ω).*

1. $\emptyset \neq x \subseteq \omega \to \exists n.n \in x \wedge n \cap x = \emptyset$.
2. $\exists n A(n) \to \exists n.A(n) \wedge \neg\exists m.m \in n \wedge A(m)$.

Beweis. 2. Es gilt nach Satz 5.3.9(2)

$$[\forall n.(\forall m.m \in n \to \neg A(m)) \to \neg A(n)] \to \forall n \neg A(n).$$

Kontraposition ergibt

$$\begin{aligned}
&\exists n A(n) \to \exists n.A(n) \wedge \forall m.m \in n \to \neg A(m)\\
&\exists n A(n) \to \exists n.A(n) \wedge \neg\exists m.m \in n \wedge A(m).
\end{aligned}$$

1. Aus (2) mit $A(y) := y \in x$. □

Wir kommen jetzt zur Rekursion über den natürlichen Zahlen, die wir als Spezialfall des Rekursionssatzes 5.3.3 behandeln können. Wir identifizieren dazu $\in$ mit der Relation $\mathcal{E} = \{\, (x, y) \mid x \in y \,\}$ und zeigen folgendes Lemma.

Lemma 5.3.11. $\in \cap (\omega \times \omega)$ *ist eine transitiv fundierte Relation auf* ω.

Beweis. Wir zeigen beide Bedingungen aus der Definition transitiv fundierter Relationen. a. Sei $\emptyset \neq a \subseteq \omega$. Zu zeigen $\exists n.n \in a \wedge n \cap a = \emptyset$. Das ist obiges Prinzip vom kleinsten Element. b. Klar, da n transitiv ist. □

Satz 5.3.12. *(Rekursion über ω mit Rückgriff auf sämtliche Vorgänger). Sei $\mathcal{G}\colon V \to V$. Dann gibt es genau eine Funktion $f\colon \omega \to V$ mit*

$$\forall n\, (f(n) = \mathcal{G}(f{\restriction}n)).$$

Beweis. Nach dem Rekursionssatz 5.3.3 gibt es ein eindeutig bestimmtes $\mathcal{F}\colon \omega \to V$ mit $\forall n\,(\mathcal{F}(n) = \mathcal{G}(\mathcal{F}{\restriction}n))$. Nach dem Ersetzungsschema ist $\mathsf{rng}(\mathcal{F}) = \mathcal{F}[\omega]$ eine Menge. Nach Lemma 5.2.1 und dem Aussonderungsschema ist damit auch $\mathcal{F} \subseteq \omega \times \mathcal{F}[\omega]$ eine Menge. □

Korollar 5.3.13. *(Rekursion über ω). Sei $\mathcal{G}\colon V \to V$ und a eine Menge. Dann gibt es genau eine Funktion $f\colon \omega \to V$ mit*

$$\begin{aligned}
&f(0) = a,\\
&\forall n\, (f(n+1) = \mathcal{G}(f(n))).
\end{aligned}$$

Beweis. Man beachte zunächst, daß $\bigcup(n+1) = n$ ist wegen

$$\begin{aligned} x \in \bigcup(n+1) &\leftrightarrow \exists y. x \in y \in n+1 \\ &\leftrightarrow \exists m. x \in m \in n+1 \\ &\leftrightarrow \exists m. x \in m \subseteq n \\ &\leftrightarrow x \in n. \end{aligned}$$

Ansatz: Zu dem gegebenen $\mathcal{G}$ finde man ein $\mathcal{G}'$ mit $\mathcal{G}'(f \restriction n+1) = \mathcal{G}(f(n))$. Wir definieren eine Funktion $\mathcal{G}' \colon V \to V$ mit

$$\mathcal{G}'(x) = \begin{cases} \mathcal{G}(x(\bigcup \mathsf{dom}(x))), & \text{falls } x \neq \emptyset; \\ a, & \text{falls } x = \emptyset, \end{cases}$$

und zwar durch

$$\mathcal{G}' = \{\, (x,y) \mid (x \neq \emptyset \to y = \mathcal{G}(x \bigcup \mathsf{dom}(x))) \wedge (x = \emptyset \to y = a) \,\}.$$

Dann gibt es genau eine Funktion $f \colon \omega \to V$ mit

$$\begin{aligned} f(n+1) &= \mathcal{G}'(f \restriction n+1) \\ &= \mathcal{G}((f \restriction n+1)(\underbrace{\bigcup(n+1)}_{n})) \\ &= \mathcal{G}(f(n)), \\ f(0) &= \mathcal{G}'(\underbrace{f \restriction 0}_{\emptyset}) \\ &= a. \end{aligned}$$

□

Wir definieren jetzt

$$s_m(0) = m, \quad s_m(n+1) = s_m(n) + 1.$$

Nach Korollar 5.3.13 existiert für jedes m eine solche Funktion, und sie ist eindeutig bestimmt. Weiter definieren wir

$$m + n := s_m(n).$$

Wegen $s_m(1) = s_m(0+1) = s_m(0)+1 = m+1$ verträgt sich diese Definition für $n = 1$ mit der bisherigen Terminologie. Ferner gilt $m + 0 = m$ und $m + (n+1) = (m+n) + 1$.

Lemma 5.3.14. *1.* $m + n \in \omega$.
2. $(m+n) + p = m + (n+p)$.
3. $m + n = n + m$.

Beweis. 1. Induktion nach n. 0. Klar. $n \to n+1$. $m + (n+1) = (m+n)+1$, und nach IH ist $m + n \in \omega$.

2. Induktion nach p. 0. Klar. $p \to p+1$.

$$\begin{aligned} (m+n) + (p+1) &= [(m+n)+p] + 1 && \text{nach Definition} \\ &= [m + (n+p)] + 1 && \text{nach IH} \\ &= m + [(n+p)+1] \\ &= m + [n + (p+1)]. \end{aligned}$$

3. Zunächst zeigen wir zwei Hilfsaussagen.
(i) $0 + n = n$. Beweis durch Induktion nach n. 0. Klar. $n \to n+1$. $0 + (n+1) = (0+n)+1 = n+1$.
(ii) $(m+1) + n = (m+n) + 1$. Beweis durch Induktion nach n. 0. Klar. $n \to n+1$.

$$\begin{aligned} (m+1) + (n+1) &= [(m+1)+n] + 1 \\ &= [(m+n)+1] + 1 && \text{nach IH} \\ &= [m + (n+1)] + 1. \end{aligned}$$

Jetzt ergibt sich die Behauptung $m + n = n + m$ durch Induktion nach m. 0. Nach (i). $m \to m + 1$.

$$\begin{aligned}(m+1)+n &= (m+n)+1 && \text{nach (ii)}\\ &= (n+m)+1 && \text{nach IH}\\ &= n+(m+1). \end{aligned} \qquad \square$$

Wir definieren

$$p_m(0) = 0, \quad p_m(n+1) = p_m(n) + m.$$

Nach Korollar 5.3.13 existiert wieder für jedes m eine eindeutig bestimmte solche Funktion. Gebraucht wird hier

$$\mathcal{G}\colon V \to V,$$
$$\mathcal{G}(x) = \begin{cases} x + m, & \text{falls } x \in \omega;\\ \emptyset, & \text{sonst.}\end{cases}$$

Schließlich definieren wir $m \cdot n := p_m(n)$. Man beachte, daß hieraus folgt $m \cdot 0 = 0$, $m \cdot (n+1) = m \cdot n + m$.

Lemma 5.3.15. *1.* $m \cdot n \in \omega$.
2. $m \cdot (n + p) = m \cdot n + m \cdot p$.
3. $(n + p) \cdot m = n \cdot m + p \cdot m$.
4. $(m \cdot n) \cdot p = m \cdot (n \cdot p)$.
5. $0 \cdot n = 0$, $1 \cdot n = n$, $m \cdot n = n \cdot m$.

Beweis. Übung. $\square$

Anmerkung. n^m, $m - n$ lassen sich ähnlich behandeln; später (in der Ordinalzahlarithmetik) werden wir dies allgemeiner durchführen. - Man kann jetzt leicht auf die bekannte Weise ganze, rationale, reelle und komplexe Zahlen definieren und ihre elementaren Eigenschaften beweisen.

Transitive Hülle

Wir definieren die $\mathcal{R}$-transitive Hülle einer Menge a, und zwar bezüglich einer Relation $\mathcal{R}$ mit der Eigenschaft, daß die $\mathcal{R}$-Vorgänger eines beliebigen Elements ihres Bereichs eine Menge bilden.

Satz 5.3.16. *Sei $\mathcal{R}$ eine Relation auf $\mathcal{A}$ derart daß $\hat{x}^{\mathcal{R}}$ $(:= \{\, y \mid y\mathcal{R}x \,\})$ für jedes $x \in \mathcal{A}$ eine Menge ist. Dann gibt es zu jeder Teilmenge $a \subseteq \mathcal{A}$ eine eindeutig bestimmte Menge b mit*

1. $a \subseteq b \subseteq \mathcal{A}$.
2. b *ist* $\mathcal{R}$*-transitiv.*
3. $\forall c.a \subseteq c \subseteq \mathcal{A} \land c$ $\mathcal{R}$*-transitiv* $\to b \subseteq c$.

b heißt die $\mathcal{R}$-transitive Hülle *von a.*

Beweis. Eindeutigkeit. Klar nach (3). Existenz. Wir wollen ein $f\colon \omega \to V$ durch Rekursion über ω so definieren, daß gilt

$$f(0) = a,$$
$$f(n+1) = \{\, y \mid \exists x{\in}f(n)(y\mathcal{R}x) \,\}.$$

Um den Rekursionssatz für ω anwenden zu können, müssen wir $f(n+1)$ definieren in der Form $\mathcal{G}(f(n))$. Dazu wählen wir $\mathcal{G}\colon V \to V$, $z \mapsto \bigcup \mathsf{rng}(\hat{}\restriction z)$ mit $\hat{}\colon V \to V$, $x \mapsto \hat{x}$; nach Annahme ist $\hat{}$ eine Funktion. Dann gilt

$$\begin{aligned} y \in \mathcal{G}(f(n)) &\leftrightarrow y \in \bigcup \mathsf{rng}(\hat{}\restriction f(n))\\ &\leftrightarrow \exists z.z \in \mathsf{rng}(\hat{}\restriction f(n)) \land y \in z\\ &\leftrightarrow \exists z, x.x \in f(n) \land z = \hat{x} \land y \in z\\ &\leftrightarrow \exists x.x \in f(n) \land y\mathcal{R}x. \end{aligned}$$

Durch Induktion über n sieht man leicht, daß $f(n)$ eine Menge ist. Für 0 ist dies klar, und im Schritt $n \to n+1$ folgt dies wegen $f(n+1) = \bigcup\{\, \hat{x} \mid x \in f(n) \,\}$ aus der IH, dem Ersetzungsschema und dem Vereinigungsmengenaxiom. – Wir setzen jetzt $b := \bigcup \mathsf{rng}(f) = \bigcup\{\, f(n) \mid n \in \omega \,\}$. Dann erhält man folgendes.

1. $a = f(0) \subseteq b \subseteq \mathcal{A}$.
2.

$$\begin{aligned}
&y \mathcal{R} x \in b \\
&y \mathcal{R} x \in f(n) \\
&y \in f(n+1) \\
&y \in b.
\end{aligned}$$

3. Sei $a \subseteq c \subseteq \mathcal{A}$ und c $\mathcal{R}$-transitiv. Wir zeigen $f(n) \subseteq c$ durch Induktion über n. 0. $a \subseteq c$. $n \to n+1$.

$$\begin{aligned}
&y \in f(n+1) \\
&y \mathcal{R} x \in f(n) \\
&y \mathcal{R} x \in c \\
&y \in c.
\end{aligned}$$

□

Anmerkung. Im Spezialfall der Relation $\in$ auf V ist die Bedingung $\forall x(\hat{x} = \{\, y \mid y \in x \,\}$ ist Menge) offenbar erfüllt. Also gibt es zu jeder Menge a eine eindeutig bestimmte $\in$-transitive Hülle von a. Sie heißt die *transitive Hülle* von a.

Rekursion über fundierte Relationen

Mit Hilfe des Begriffs der $\mathcal{R}$-transitiven Hülle können wir jetzt zeigen, daß die transitiv fundierten Relationen auf $\mathcal{A}$ mit den fundierten Relationen auf $\mathcal{A}$ übereinstimmen.

Sei $\mathcal{R}$ eine Relation auf $\mathcal{A}$. $\mathcal{R}$ heißt *fundierte Relation* auf $\mathcal{A}$, wenn gilt

1. Jede nichtleere Teilmenge von $\mathcal{A}$ besitzt ein $\mathcal{R}$-minimales Element, d.h.

$$\forall a. a \subseteq \mathcal{A} \land a \neq \emptyset \to \exists x{\in}a. \hat{x} \cap a = \emptyset.$$

2. Für jedes $x \in \mathcal{A}$ ist $\hat{x}$ eine Menge.

Satz 5.3.17. *Die transitiv fundierten Relationen auf $\mathcal{A}$ stimmen mit den fundierten Relationen auf $\mathcal{A}$ überein.*

Beweis. Jede transitiv fundierte Relation auf $\mathcal{A}$ ist fundiert nach Lemma 5.3.1(2). Umgekehrt ist jede fundierte Relation auf $\mathcal{A}$ transitiv fundiert, denn zu jedem $x \in \mathcal{A}$ ist die $\mathcal{R}$-transitive Hülle von $\hat{x}$ ein $\mathcal{R}$-transitives $b \subseteq \mathcal{A}$ mit $\hat{x} \subseteq b$. □

Der Induktionssatz 5.3.2 und der Rekursionssatz 5.3.3 gelten für also fundierte Relationen. Ebenso gilt nach Lemma 5.3.1(1), daß jede nichtleere Teil*klasse* einer fundierten Relation $\mathcal{R}$ ein $\mathcal{R}$-minimales Element hat.

Später werden wir ein sogenanntes Regularitätsaxiom fordern, welches aussagt, daß die Relation $\in$ auf V fundiert ist, d.h. daß gilt

$$\forall a. a \neq \emptyset \to \exists x{\in}a. x \cap a = \emptyset.$$

Damit werden wir dann ein wichtiges Beispiel einer fundierten Relation zur Verfügung haben.

Wohlordnungen

Wir betrachten jetzt extensionale fundierte Relationen. Aus dem Regularitätsaxiom wird später folgen, daß die $\in$-Relation auf einer beliebigen Klasse $\mathcal{A}$ eine fundierte extensionale Relation ist. Hier zeigen wir noch ohne das Regularitätsaxiom die Unkehrung, daß nämlich jede fundierte extensionale Relation zu der $\in$-Relation auf einer transitiven Klasse isomorph ist. Dies ist der Isomorphiesatz von MOSTOWSKI. Dann betrachten wir lineare fundierte Ordnungen, kurz Wohlordnungen. Sie sind stets extensional und deshalb zur $\in$-Relation auf gewissen Klassen isomorph, die wir ordinale Klassen nennen. Ordinalzahlen sind dann die ordinalen Mengen.

Eine Relation $\mathcal{R}$ auf $\mathcal{A}$ heißt *extensional*, wenn für alle $x, y \in \mathcal{A}$ gilt

$$(\forall z{\in}\mathcal{A}. z\mathcal{R}x \leftrightarrow z\mathcal{R}y) \to x = y.$$

Zum Beispiel ist für eine transitive Klasse $\mathcal{A}$ die Relation $\in \cap(\mathcal{A} \times \mathcal{A})$ extensional auf $\mathcal{A}$. Dies sieht man wie folgt. Seien $x, y \in \mathcal{A}$. Für $\mathcal{R} :=\in \cap(\mathcal{A} \times \mathcal{A})$ gilt $z\mathcal{R}x \leftrightarrow z \in x$, da $\mathcal{A}$ transitiv ist. Man erhält

$$\begin{aligned}
&\forall z{\in}\mathcal{A}. z\mathcal{R}x \leftrightarrow z\mathcal{R}y \\
&\forall z. z \in x \leftrightarrow z \in y \\
&x = y
\end{aligned}$$

Aus dem Regularitätsaxiom wird folgen, daß alle diese Relationen fundiert sind. Auch ohne das Regularitätsaxiom haben diese Relationen eine ausgezeichnete Bedeutung; vgl. Korollar 5.3.19.

Satz 5.3.18. *(Isomorphiesatz von* MOSTOWSKI*). Sei $\mathcal{R}$ eine fundierte extensionale Relation auf $\mathcal{A}$. Dann gibt es einen eindeutig bestimmten Isomorphismus $\mathcal{F}$ von $\mathcal{A}$ auf eine transitive Klasse $\mathcal{B}$, d.h.*

$$\exists^{=1}\mathcal{F}.\mathcal{F}\colon \mathcal{A} \leftrightarrow \mathsf{rng}(\mathcal{F}) \wedge \mathsf{rng}(\mathcal{F}) \text{ transitiv } \wedge \forall x, y{\in}\mathcal{A}. y\mathcal{R}x \leftrightarrow \mathcal{F}(y) \in \mathcal{F}(x).$$

Beweis. Existenz. Wir definieren nach dem Rekursionssatz

$$\begin{aligned}
&\mathcal{F}\colon \mathcal{A} \to V, \\
&\mathcal{F}(x) = \mathsf{rng}(\mathcal{F}{\restriction}\hat{x}) \quad (= \{\, \mathcal{F}(y) \mid y\mathcal{R}x \,\}).
\end{aligned}$$

$\mathcal{F}$ injektiv. Wir zeigen $\forall x, y{\in}\mathcal{A}. \mathcal{F}(x) = \mathcal{F}(y) \to x = y$ durch $\mathcal{R}$-Induktion über x. Seien also $x, y \in \mathcal{A}$ gegeben mit $\mathcal{F}(x) = \mathcal{F}(y)$. Nach IH gilt

$$\forall z{\in}\mathcal{A}. z\mathcal{R}x \to \forall u{\in}\mathcal{A}. \mathcal{F}(z) = \mathcal{F}(u) \to z = u.$$

Es genügt zu zeigen, daß für alle $z \in \mathcal{A}$ gilt $z\mathcal{R}x \leftrightarrow z\mathcal{R}y$. $\to$.

$$\begin{aligned}
&z\mathcal{R}x \\
&\mathcal{F}(z) \in \mathcal{F}(x) = \mathcal{F}(y) = \{\, \mathcal{F}(u) \mid u\mathcal{R}y \,\} \\
&\mathcal{F}(z) = \mathcal{F}(u) && \text{für ein } u\mathcal{R}y \\
&z = u && \text{nach IH, da } z\mathcal{R}x \\
&z\mathcal{R}y
\end{aligned}$$

$\leftarrow$.

$$\begin{aligned}
&z\mathcal{R}y \\
&\mathcal{F}(z) \in \mathcal{F}(y) = \mathcal{F}(x) = \{\, \mathcal{F}(u) \mid u\mathcal{R}x \,\} \\
&\mathcal{F}(z) = \mathcal{F}(u) && \text{für ein } u\mathcal{R}x \\
&z = u && \text{nach IH, da } u\mathcal{R}x \\
&z\mathcal{R}x.
\end{aligned}$$

$\mathsf{rng}(\mathcal{F})$ ist transitiv. Gelte $u \in v \in \mathsf{rng}(\mathcal{F})$. Dann ist $v = \mathcal{F}(x)$, also $u = \mathcal{F}(y)$ für ein $y\mathcal{R}x$.

$y\mathcal{R}x \leftrightarrow \mathcal{F}(y) \in \mathcal{F}(x)$. $\to$. Gelte $y\mathcal{R}x$. Dann ist $\mathcal{F}(y) \in \mathcal{F}(x)$ nach Definition von $\mathcal{F}$. $\leftarrow$.

$$\begin{array}{ll} \mathcal{F}(y) \in \mathcal{F}(x) = \{\, \mathcal{F}(z) \mid z\mathcal{R}x \,\} & \\ \mathcal{F}(y) = \mathcal{F}(z) & \text{für ein } z\mathcal{R}x \\ y = z & \text{da } \mathcal{F} \text{ injektiv ist} \\ y\mathcal{R}x. & \end{array}$$

Eindeutigkeit. Gegeben seien $\mathcal{F}_i$, $i = 1, 2$. Wir zeigen $\forall x{\in}\mathcal{A}\,(\mathcal{F}_1(x) = \mathcal{F}_2(x))$ durch $\mathcal{R}$-Induktion über x. Aus Symmetriegründen genügt $u \in \mathcal{F}_1(x) \to u \in \mathcal{F}_2(x)$.

$$\begin{array}{ll} u \in \mathcal{F}_1(x) & \\ u = \mathcal{F}_1(y) & \text{für ein } y \in \mathcal{A}, \text{ da } \mathsf{rng}(\mathcal{F}_1) \text{ transitiv} \\ y\mathcal{R}x & \text{nach der Isomorphiebedingung für } \mathcal{F}_1 \\ u = \mathcal{F}_2(y) & \text{nach IH} \\ \mathcal{F}_2(y) \in \mathcal{F}_2(x) & \text{nach der Isomorphiebedingung für } \mathcal{F}_2 \\ u \in \mathcal{F}_2(x). & \end{array}$$

□

Eine Relation $\mathcal{R}$ auf $\mathcal{A}$ heißt *lineare Ordnung*, wenn für alle $x, y, z \in \mathcal{A}$ gilt

$$\begin{array}{ll} \neg x\mathcal{R}x & \text{Irreflexivität,} \\ x\mathcal{R}y \wedge y\mathcal{R}z \to x\mathcal{R}z & \text{Transitivität,} \\ x\mathcal{R}y \vee x = y \vee y\mathcal{R}x & \text{Trichotomie (oder Vergleichbarkeit).} \end{array}$$

$\mathcal{R}$ heißt *Wohlordnung*, wenn $\mathcal{R}$ eine fundierte lineare Ordnung ist.

Anmerkung. Jede Wohlordnung $\mathcal{R}$ auf $\mathcal{A}$ ist extensional. Gilt nämlich

$$\forall z{\in}\mathcal{A}.z\mathcal{R}x \leftrightarrow z\mathcal{R}y,$$

so folgt $x = y$ aus der Trichotomie, denn aus $x\mathcal{R}y$ folgt nach Annahme $x\mathcal{R}x$ im Widerspruch zur Irreflexivität, und ebenso erhält man aus $y\mathcal{R}x$ einen Widerspruch.

Korollar 5.3.19. *Zu jeder Wohlordnung $\mathcal{R}$ auf $\mathcal{A}$ gibt es einen eindeutig bestimmten Isomorphismus $\mathcal{F}$ von $\mathcal{A}$ auf eine transitive Klasse $\mathcal{B}$.* □

Ordinale Klassen und Ordinalzahlen

Wir wollen jetzt die transitiven Klassen, die als Bilder von Wohlordnungen auftreten, genauer untersuchen.

$\mathcal{A}$ heißt *ordinale Klasse*, wenn $\mathcal{A}$ transitiv ist und $\in \cap(\mathcal{A} \times \mathcal{A})$ eine Wohlordnung auf $\mathcal{A}$ ist. Ordinale Klassen, die Mengen sind, heißen *Ordinalzahlen*. Wir setzen

$$\mathsf{On} := \{\, x \mid x \text{ ist Ordinalzahl} \,\}.$$

Zunächst geben wir eine bequeme Charakterisierung ordinaler Klassen an. $\mathcal{A}$ heißt *konnex*, wenn für alle $x, y \in \mathcal{A}$ gilt

$$x \in y \vee x = y \vee y \in x.$$

Zum Beispiel ist ω konnex nach Lemma 5.3.7(4). Auch jedes n ist konnex, da ω nach Lemma 5.3.5(2) transitiv ist.

Eine Klasse $\mathcal{A}$ heißt *fundiert*, wenn $\in \cap(\mathcal{A} \times \mathcal{A})$ eine fundierte Relation auf $\mathcal{A}$ ist, d.h. wenn gilt

$$\forall a.a \subseteq \mathcal{A} \wedge a \neq \emptyset \to \exists x{\in}a.x \cap a = \emptyset.$$

Wir zeigen jetzt, daß es in fundierten Klassen keine endlichen $\in$-Zykel geben kann.

Lemma 5.3.20. *Sei $\mathcal{A}$ fundiert. Dann kann für beliebige $x_1, \ldots, x_n \in \mathcal{A}$ niemals gelten*

$$x_1 \in x_2 \in \cdots \in x_n \in x_1.$$

Beweis. Annahme: $x_1 \in x_2 \in \cdots \in x_n \in x_1$. Betrachte $\{x_1, \dots, x_n\}$. Da $\mathcal{A}$ fundiert ist, gilt oBdA $x_1 \cap \{x_1, \dots, x_n\} = \emptyset$. Dies widerspricht aber $x_n \in x_1$. □

Korollar 5.3.21. *$\mathcal{A}$ ist ordinale Klasse genau dann, wenn $\mathcal{A}$ transitiv, konnex und fundiert ist.*

Beweis. $\to$ ist klar; die Konnexität von $\mathcal{A}$ folgt aus der Trichotomieeigenschaft. $\leftarrow$. Zu zeigen ist für alle $x, y, z \in \mathcal{A}$

$$x \notin x,$$
$$x \in y \wedge y \in z \to x \in z.$$

Da $\mathcal{A}$ konnex ist, folgen beide Aussagen aus Lemma 5.3.20. □

Anmerkung. 1. ω ist transitiv nach Lemma 5.3.5(2), konnex wie oben bemerkt und fundiert nach dem Prinzip vom kleinsten Element (Satz 5.3.10). Also ist ω eine ordinale Klasse. Da ω nach dem Unendlichkeitsaxiom eine Menge ist, ist ω sogar eine Ordinalzahl.

2. n ist transitiv nach Lemma 5.3.5(1), konnex wie oben bemerkt und fundiert; letzteres folgt mit der Transitivität von ω aus dem Prinzip vom kleinsten Element (Satz 5.3.10).

$\mathsf{Ord}(\mathcal{A})$ steht für "$\mathcal{A}$ ist ordinale Klasse". Wir zeigen jetzt, daß ordinale Klassen ähnliche Eigenschaften haben wie die natürlichen Zahlen: auch für ordinale Klassen hat die Relation $\in$ die Eigenschaften von $<$ und die Relation $\subseteq$ die Eigenschaften von $\leq$.

Lemma 5.3.22. *1.* $\mathsf{Ord}(\mathcal{A}) \wedge \mathsf{Ord}(\mathcal{B}) \to \mathsf{Ord}(\mathcal{A} \cap \mathcal{B})$.
2. $\mathsf{Ord}(\mathcal{A}) \wedge x \in \mathcal{A} \to \mathsf{Ord}(x)$.
3. $\mathsf{Ord}(\mathcal{A}) \wedge \mathsf{Ord}(\mathcal{B}) \to (\mathcal{A} \subseteq \mathcal{B} \leftrightarrow \mathcal{A} \in \mathcal{B} \vee \mathcal{A} = \mathcal{B})$.
4. $\mathsf{Ord}(\mathcal{A}) \wedge \mathsf{Ord}(\mathcal{B}) \to (\mathcal{A} \in \mathcal{B} \vee \mathcal{A} = \mathcal{B} \vee \mathcal{B} \in \mathcal{A})$.

Beweis. 1. $\mathcal{A} \cap \mathcal{B}$ transitiv. Man erhält

$$x \in y \in \mathcal{A} \cap \mathcal{B}$$
$$x \in y \in \mathcal{A} \quad \text{und} \quad x \in y \in \mathcal{B}$$
$$x \in \mathcal{A} \quad \text{und} \quad x \in \mathcal{B}$$
$$x \in \mathcal{A} \cap \mathcal{B}.$$

$\mathcal{A} \cap \mathcal{B}$ konnex, fundiert. Klar.

2. x transitiv. Man erhält

$$u \in v \in x \in \mathcal{A}$$
$$u \in v \in \mathcal{A}$$
$$u \in \mathcal{A}$$
$$u \in x \vee u = x \vee x \in u.$$

Aus $u = x$ folgt $u \in v \in u$ und damit ein Widerspruch zu Lemma 5.3.20, und aus $x \in u$ folgt $u \in v \in x \in u$ und damit ebenfalls ein Widerspruch zu Lemma 5.3.20.

x konnex, fundiert. Klar, da $x \subseteq \mathcal{A}$.

3. $\leftarrow$. Klar, da $\mathcal{B}$ transitiv ist. $\to$. Sei $\mathcal{A} \subseteq \mathcal{B}$. OBdA $\mathcal{A} \subsetneq \mathcal{B}$. Man wähle ein $x \in \mathcal{B} \setminus \mathcal{A}$ mit $x \cap (\mathcal{B} \setminus \mathcal{A}) = \emptyset$ (dies ist möglich, da $\mathcal{B}$ fundiert ist). Es genügt zu zeigen, daß $x = \mathcal{A}$.

$x \subseteq \mathcal{A}$. Gelte $y \in x$, also $y \in x \in \mathcal{B}$. Dann folgt $y \in \mathcal{A}$, denn $x \cap (\mathcal{B} \setminus \mathcal{A}) = \emptyset$.

$\mathcal{A} \subseteq x$. Gelte $y \in \mathcal{A}$. Dann ist auch $y \in \mathcal{B}$. Es folgt $x \in y \vee x = y \vee y \in x$. Die ersten beiden Fälle sind aber unmöglich, denn jedesmal erhält man $x \in \mathcal{A}$.

4. Gelte $\mathsf{Ord}(\mathcal{A})$ und $\mathsf{Ord}(\mathcal{B})$. Dann folgt $\mathsf{Ord}(\mathcal{A} \cap \mathcal{B})$ nach (1). Aus (3) ergibt sich

$$[(\mathcal{A} \cap \mathcal{B} \in \mathcal{A}) \vee (\mathcal{A} \cap \mathcal{B} = \mathcal{A})] \wedge [(\mathcal{A} \cap \mathcal{B} \in \mathcal{B}) \vee (\mathcal{A} \cap \mathcal{B} = \mathcal{B})].$$

Ausdistribuieren liefert

$$(\mathcal{A} \cap \mathcal{B} \in \mathcal{A} \cap \mathcal{B}) \vee (\mathcal{A} \in \mathcal{B}) \vee (\mathcal{B} \in \mathcal{A}) \vee (\mathcal{A} = \mathcal{B}).$$

Der erste Fall $\mathcal{A} \cap \mathcal{B} \in \mathcal{A} \cap \mathcal{B}$ ist aber unmöglich nach Lemma 5.3.20. □

Lemma 5.3.23. *1.* $\mathsf{Ord}(\mathsf{On})$.
2. On *ist keine Menge.*
3. On *ist die einzige echte ordinale Klasse.*

Beweis. 1. On ist transitiv nach Lemma 5.3.22(2) und konnex nach Lemma 5.3.22(4). On ist auch fundiert. Sei nämlich $a \subseteq \mathsf{On}$, $a \neq \emptyset$. Wähle $x \in a$. OBdA ist $x \cap a \neq \emptyset$. Da x fundiert ist, gibt es ein $y \in x \cap a$ mit $y \cap x \cap a = \emptyset$. Es folgt $y \in a$ und $y \cap a = \emptyset$; letzteres da $y \subseteq x$ wegen $y \in x$, x transitiv.

2. Annahme: On ist Menge. Dann folgt $\mathsf{On} \in \mathsf{On}$ und damit ein Widerspruch zu Lemma 5.3.20.

3. Sei $\mathsf{Ord}(\mathcal{A})$, $\mathcal{A}$ keine Menge. Nach Lemma 5.3.22(4) folgt

$$\mathcal{A} \in \mathsf{On} \vee \mathcal{A} = \mathsf{On} \vee \mathsf{On} \in \mathcal{A}.$$

Der erste und der dritte Fall scheiden aber aus, da dann $\mathcal{A}$ bzw. On eine Menge wäre. □

Lemma 5.3.24. *1.* On *ist induktiv,*
2. $n, \omega \in \mathsf{On}$.

Beweis. 1. $0 \in \mathsf{On}$ ist klar. Sei nun $x \in \mathsf{On}$. Zu zeigen ist $x + 1 \in \mathsf{On}$, also $x \cup \{x\} \in \mathsf{On}$.

$x \cup \{x\}$ transitiv. Gelte $u \in v \in x \cup \{x\}$, also $u \in v \in x$ oder $u \in v = x$. In beiden Fällen folgt $u \in x$.

$x \cup \{x\}$ konnex. Gelte $u, v \in x \cup \{x\}$. Dann folgt

$$\begin{aligned}
&u, v \in x \vee (u \in x \wedge v = x) \vee (u = x \wedge v \in x) \vee (u = v = x)\\
&u \in v \vee u = v \vee v \in u.
\end{aligned}$$

$x \cup \{x\}$ fundiert. Sei $a \subseteq x \cup \{x\}$, $a \neq \emptyset$. Zu zeigen ist $\exists y{\in}a\, (y \cap a = \emptyset)$. *Fall* $a \cap x \neq \emptyset$. Dann folgt die Behauptung aus der Fundiertheit von x. *Fall* $a \cap x = \emptyset$. Dann ist $a = \{x\}$, und wir haben $x \cap \{x\} = \emptyset$.

2. Dies wurde schon oben im Anschluß an Korollar 5.3.21 gezeigt. □

Lemma 5.3.25. $x, y \in \mathsf{On} \wedge x + 1 = y + 1 \to x = y$.

Beweis. Dies zeigt man wie das zweite Peano-Axiom in Lemma 5.3.8(2).

$$\begin{aligned}
&x + 1 = y + 1\\
&x \in y + 1 \wedge y \in x + 1\\
&(x \in y \wedge y \in x) \vee x = y.
\end{aligned}$$

Da der erste Fall nach Lemma 5.3.22 unmöglich ist, folgt $x = y$. □

Lemma 5.3.26. $\mathcal{A} \subseteq \mathsf{On} \to \bigcup \mathcal{A} \in \mathsf{On} \vee \bigcup \mathcal{A} = \mathsf{On}$.

Beweis. Genügt: $\mathsf{Ord}(\bigcup \mathcal{A})$. $\bigcup \mathcal{A}$ transitiv. Sei $x \in y \in \bigcup \mathcal{A}$, also $x \in y \in z \in \mathcal{A}$ für ein z. Dann folgt $x \in z \in \mathcal{A}$, da $\mathcal{A} \subseteq \mathsf{On}$. Also ist $x \in \bigcup \mathcal{A}$.

$\bigcup \mathcal{A}$ konnex und fundiert. Genügt: $\bigcup \mathcal{A} \subseteq \mathsf{On}$. Sei also $x \in \bigcup \mathcal{A}$, also $x \in y \in \mathcal{A}$ für ein y. Dann ist $x \in y$ und $y \in \mathsf{On}$, also $x \in \mathsf{On}$. □

Anmerkung. Gilt $\mathcal{A} \subseteq \mathsf{On}$, so ist $\bigcup \mathcal{A}$ die kleinste obere Schranke von $\mathcal{A}$ bzgl. der Wohlordnung $\in \cap (\mathsf{On} \times \mathsf{On})$ von On, denn nach Definition von $\bigcup \mathcal{A}$ gilt

$$\begin{aligned}
&x \in \mathcal{A} \to x \subseteq \bigcup \mathcal{A},\\
&(\forall x{\in}\mathcal{A}.x \subseteq y) \to \bigcup \mathcal{A} \subseteq y.
\end{aligned}$$

Wir schreiben deshalb auch $\sup \mathcal{A}$ für $\bigcup \mathcal{A}$.

Beispiele für Ordinalzahlen

0	
$1 = 0 + 1$	
$2 = 1 + 1$	
$\vdots$	
ω	Menge nach dem Unendlichkeitsaxiom
$\omega + 1$	
$\omega + 2$	
$\vdots$	
$\omega \cdot 2 := \bigcup\{\, \omega + n \mid n \in \omega \,\}$	Def. mit Rekursion über ω; Menge nach Ersetzungsschema
$\omega \cdot 2 + 1$	
$\omega \cdot 2 + 2$	
$\vdots$	
$\omega \cdot 3 := \bigcup\{\, \omega \cdot 2 + n \mid n \in \omega \,\}$	
$\vdots$	
$\omega \cdot 4$	
$\vdots$	
$\omega \cdot \omega := \omega^2 := \bigcup\{\, \omega \cdot n \mid n \in \omega \,\}$	
$\omega^2 + 1$	
$\omega^2 + 2$	
$\vdots$	
$\omega^2 + \omega$	
$\omega^2 + \omega + 1$	
$\omega^2 + \omega + 2$	
$\vdots$	
$\omega^2 + \omega \cdot 2$	
$\vdots$	
$\omega^2 + \omega \cdot 3$	
$\vdots$	
ω^3	
$\vdots$	
ω^4	
$\vdots$	
ω^ω	
$\omega^\omega + 1$	

$\vdots$

und so weiter.

α, β, γ stehen im folgenden für Ordinalzahlen.

α heißt eine *Nachfolgerzahl*, wenn $\exists\beta(\alpha = \beta + 1)$. α heißt *Limeszahl*, wenn α weder 0 noch Nachfolgerzahl ist. Wir schreiben

$$\mathsf{Lim}(\alpha) \text{ für } \alpha \neq 0 \wedge \neg\exists\beta(\alpha = \beta + 1).$$

Offenbar gilt für beliebige α entweder $\alpha = 0$ oder α ist Nachfolgerzahl oder α ist Limeszahl.

Lemma 5.3.27. *1.* $\mathsf{Lim}(\alpha) \leftrightarrow \alpha \neq 0 \wedge \forall\beta.\beta \in \alpha \to \beta + 1 \in \alpha$.
2. $\mathsf{Lim}(\omega)$.
3. $\mathsf{Lim}(\alpha) \to \omega \subseteq \alpha$.

Beweis. 1. $\to$: Sei $\beta \in \alpha$. Dann gilt $\beta + 1 \in \alpha \vee \beta + 1 = \alpha \vee \alpha \in \beta + 1$. Der zweite Fall $\beta + 1 = \alpha$ ist ausgeschlossen nach Annahme. Im dritten Fall folgt $\alpha \in \beta \vee \alpha = \beta$; beides ist aber wegen $\beta \in \alpha$ unmöglich nach Lemma 5.3.20. $\leftarrow$. Sei $\alpha \neq 0$ und es gelte $\forall\beta.\beta \in \alpha \to \beta + 1 \in \alpha$. Ist dann α keine Limeszahl, so muß $\alpha = \beta + 1$ sein. Dann folgt $\beta \in \alpha$, also nach Annahme auch $\beta + 1 \in \alpha$ und damit $\alpha \in \alpha$, was nicht sein kann.

2. Folgt aus (1), da ω induktiv ist.

3. Gelte $\mathsf{Lim}(\alpha)$. Wir zeigen $n \in \alpha$ durch Induktion über n. 0. Es gilt $0 \in \alpha \vee 0 = \alpha \vee \alpha \in 0$, wobei die Fälle zwei und drei offenbar unmöglich sind. $n + 1$. Es ist $n \in \alpha$ nach IH, also $n + 1 \in \alpha$ nach a. □

Lemma 5.3.28. *1.* $\alpha = \bigcup_{\beta\in\alpha}(\beta + 1)$.
2. Für Limeszahlen α gilt $\alpha = \bigcup_{\beta\in\alpha}\beta$.

Beweis. 1. $\subseteq$. Sei $\beta \in \alpha$. Die Behauptung folgt aus $\beta \in \beta + 1$. $\supseteq$. Sei $\beta \in \alpha$. Dann ist $\beta + 1 \subseteq \alpha$.

2. $\subseteq$. Sei $\gamma \in \alpha$. Dann folgt $\gamma \in \gamma + 1 \in \alpha$. $\supseteq$. Sei $\gamma \in \beta \in \alpha$. Es folgt $\gamma \in \alpha$. □

Satz 5.3.29. *(Transfinite Induktion über* On *mit Rückgriff auf sämtliche Vorgänger; Klassenform).*

$$(\forall\alpha.\alpha \subseteq \mathcal{B} \to \alpha \in \mathcal{B}) \to \mathsf{On} \subseteq \mathcal{B}.$$

Beweis. Dies ist ein Spezialfall des Induktionssatzes 5.3.2 □

Korollar 5.3.30. *(Verschiedene Formen der transfiniten Induktion über* On*). Erste Form:*

$$A(0) \wedge (\forall\alpha.A(\alpha) \to A(\alpha + 1)) \wedge (\forall\alpha.\mathsf{Lim}(\alpha) \wedge (\forall\beta.\beta \in \alpha \to A(\beta)) \to A(\alpha)) \to \forall\alpha A(\alpha).$$

Zweite Form: (Transfinite Induktion über On *mit Rückgriff auf sämtliche Vorgänger).*

$$[\forall\alpha.(\forall\beta.\beta \in \alpha \to A(\beta)) \to A(\alpha)] \to \forall\alpha A(\alpha).$$

Dritte Form: (Prinzip vom kleinsten Element für On).

$$\exists\alpha A(\alpha) \to \exists\alpha.A(\alpha) \wedge \neg\exists\beta.\beta \in \alpha \wedge A(\beta).$$

Beweis. Die dritte Form ergibt sich aus der zweiten durch Kontraposition. Ferner ergibt sich die erste Form ebenfalls leicht aus der zweiten. Die zweite Form erhält man aus Satz 5.3.29 mit $\mathcal{B} := \{\, \alpha \mid A(\alpha) \,\}$. □

Satz 5.3.31. *(Transfinite Rekursion über* On *mit Rückgriff auf sämtliche Vorgänger). Sei $\mathcal{G}\colon V \to V$. Dann gibt es genau eine Funktion $\mathcal{F}\colon \mathsf{On} \to V$ so daß für alle α gilt*

$$\mathcal{F}(\alpha) = \mathcal{G}(\mathcal{F}{\restriction}\alpha).$$

Beweis. Dies ist ein Spezialfall des Rekursionssatzes 5.3.3. □

Korollar 5.3.32. *(Transfinite Rekursion über* On*). Sei* $\mathcal{G}\colon V \to V$*,* $\mathcal{H}\colon V \to V$ *und* a *eine Menge. Dann gibt es genau eine Funktion* $\mathcal{F}\colon \mathsf{On} \to V$ *mit*

$$\begin{aligned} &\mathcal{F}(0) = a, \\ &\mathcal{F}(\alpha+1) = \mathcal{G}(\mathcal{F}(\alpha)), \\ &\mathcal{F}(\alpha) = \mathcal{H}(\mathcal{F}\restriction\alpha) \qquad \textit{für } \alpha \textit{ Limeszahl.} \end{aligned}$$

Beweis. Man beachte zunächst, daß $\bigcup(\alpha+1) = \alpha$ ist wegen

$$\begin{aligned} \gamma \in \bigcup(\alpha+1) &\leftrightarrow \exists\beta.\gamma \in \beta \in \alpha+1 \\ &\leftrightarrow \exists\beta.\gamma \in \beta \subseteq \alpha \\ &\leftrightarrow \gamma \in \alpha. \end{aligned}$$

Ansatz: Zu dem gegebenen a, $\mathcal{G}$ und $\mathcal{H}$ finde man ein $\mathcal{G}'$ mit

$$\begin{aligned} &\mathcal{G}'(0) = a, \\ &\mathcal{G}'(\mathcal{F}\restriction\alpha+1) = \mathcal{G}(\mathcal{F}(\alpha)), \\ &\mathcal{G}'(\mathcal{F}\restriction\alpha) = \mathcal{H}(\mathcal{F}\restriction\alpha) \qquad \text{für } \alpha \text{ Limeszahl.} \end{aligned}$$

Wir definieren eine Funktion $\mathcal{G}'\colon V \to V$ durch

$$\mathcal{G}'(x) = \begin{cases} a, & \text{sonst;} \\ \mathcal{G}(x(\bigcup \mathsf{dom}(x))), & \text{falls } \exists\beta(\mathsf{dom}(x) = \beta+1); \\ \mathcal{H}(x), & \text{falls } \mathsf{Lim}(\mathsf{dom}(x)). \end{cases}$$

Nach dem Rekursionssatz 5.3.3 existiert genau ein $\mathcal{F}\colon \mathsf{On} \to V$ so daß für alle α gilt

$$\mathcal{F}(\alpha) = \mathcal{G}'(\mathcal{F}\restriction\alpha).$$

Offenbar ist diese Eigenschaft von $\mathcal{F}$ zu den obigen Gleichungen äquivalent. □

Regularitätsaxiom, von Neumannsche Stufen, Rang

Wir erinnern uns zunächst an die kumulative Typenstruktur.

Stufe 0:	–
Stufe 1:	$\emptyset$
Stufe 2:	$\emptyset, \{\emptyset\}$
Stufe 3:	$\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}$
	und so weiter.

Unter Verwendung von Ordinalzahlen können wir jetzt auch transfinite Stufen betrachten. Die Stufe ω besteht dann aus allen Mengen, deren Elemente auf endlichen Stufen gebildet waren, und die Stufe $\omega+1$ aus allen Mengen, deren Elemente auf endlichen Stufen oder der Stufe ω gebildet waren, und so weiter. Allgemein definieren wir die von Neumann*schen Stufen* V_α wie folgt durch transfinite Rekursion über On.

$$\begin{aligned} &V_0 = \emptyset, \\ &V_{\alpha+1} = \mathcal{P}(V_\alpha), \\ &V_\alpha = \bigcup_{\beta\in\alpha} V_\beta \qquad \text{für } \alpha \text{ Limeszahl.} \end{aligned}$$

Anmerkung. Genauer ist $V_\alpha := \mathcal{F}(\alpha)$, wobei $\mathcal{F}\colon \mathsf{On} \to V$ wie folgt definiert wird durch transfinite Rekursion über On.

$$\begin{aligned}
&\mathcal{F}(0) = \emptyset,\\
&\mathcal{F}(\alpha+1) = \mathcal{P}(\mathcal{F}(\alpha)),\\
&\mathcal{F}(\alpha) = \bigcup \mathsf{rng}(\mathcal{F}{\restriction}\alpha) \qquad \text{für } \alpha \text{ Limeszahl.}
\end{aligned}$$

Lemma 5.3.33. *1. V_α ist transitiv.*
2. $\alpha \in \beta \to V_\alpha \in V_\beta$.
3. $\alpha \subseteq \beta \to V_\alpha \subseteq V_\beta$.
4. $V_\alpha \cap \mathsf{On} = \alpha$.

Beweis. 1. (Transfinite) Induktion nach α. 0. $\emptyset$ ist transitiv. $\alpha + 1$.

$$\begin{aligned}
&x \in y \in V_{\alpha+1} = \mathcal{P}(V_\alpha)\\
&x \in y \subseteq V_\alpha\\
&x \in V_\alpha\\
&x \subseteq V_\alpha && \text{nach IH}\\
&x \in V_{\alpha+1}.
\end{aligned}$$

α Limeszahl.

$$\begin{aligned}
&x \in y \in V_\alpha = \bigcup_{\beta\in\alpha} V_\beta\\
&x \in y \in V_\beta && \text{für ein } \beta \in \alpha\\
&x \in V_\beta && \text{nach IH}\\
&x \in V_\alpha.
\end{aligned}$$

2. Induktion nach β. 0. Klar. $\beta + 1$.

$$\begin{aligned}
&\alpha \in \beta + 1\\
&\alpha \in \beta \text{ oder } \alpha = \beta\\
&V_\alpha \in V_\beta \text{ oder } V_\alpha = V_\beta && \text{nach IH}\\
&V_\alpha \subseteq V_\beta && \text{nach (1)}\\
&V_\alpha \in V_{\beta+1}.
\end{aligned}$$

β Limeszahl.

$$\begin{aligned}
&\alpha \in \beta\\
&\alpha + 1 \in \beta\\
&V_\alpha \in V_{\alpha+1} \subseteq \bigcup_{\gamma\in\beta} V_\gamma = V_\beta.
\end{aligned}$$

3. Mit $\alpha \subseteq \beta \leftrightarrow \alpha \in \beta \vee \alpha = \beta$ folgt die Behauptung aus (1) und (2).
4. Induktion nach α. 0. Klar. $\alpha + 1$.

$$\begin{aligned}
\beta \in V_{\alpha+1} &\leftrightarrow \beta \subseteq V_\alpha\\
&\leftrightarrow \beta \subseteq V_\alpha \cap \mathsf{On} = \alpha \quad \text{nach IH}\\
&\leftrightarrow \beta \in \alpha + 1.
\end{aligned}$$

α Limeszahl.

$$\begin{aligned} V_\alpha \cap \mathsf{On} &= (\bigcup_{\beta\in\alpha} V_\beta) \cap \mathsf{On} \\ &= \bigcup_{\beta\in\alpha} (V_\beta \cap \mathsf{On}) \\ &= \bigcup_{\beta\in\alpha} \beta \qquad\qquad \text{nach IH} \\ &= \alpha. \end{aligned}$$

□

Wir wollen jetzt zeigen, daß die VON NEUMANNschen Stufen das Universum ausschöpfen, also daß gilt $V = \bigcup_{\alpha\in\mathsf{On}} V_\alpha$. Dazu brauchen wir das bereits erwähnte *Regularitätsaxiom*, welches aussagt, daß die Relation $\in$ auf V fundiert ist, d.h. daß gilt

Axiom 8.

$$\forall a. a \neq \emptyset \to \exists x{\in}a\, (x \cap a = \emptyset).$$

Wir können dann jeder Menge x eine Ordinalzahl α als ihren Rang zuordnen, und zwar als kleinstes α mit $x \subseteq V_\alpha$. Dazu brauchen wir den Begriff des *Rangs* $\mathsf{rn}(x)$ einer Menge x, den wir rekursiv definieren durch

$$\mathsf{rn}(x) := \bigcup\{\, \mathsf{rn}(y) + 1 \mid y \in x \,\}.$$

Genauer setzen wir $\mathsf{rn}(x) := \mathcal{F}(x)$, wobei $\mathcal{F}\colon V \to V$ wie folgt definiert wird (mittels des Rekursionssatzes 5.3.3 für fundierte Relationen).

$$\mathcal{F}(x) := \bigcup \mathsf{rng}(\mathcal{H}(\mathcal{F}{\restriction}x))$$

mit

$$\mathcal{H}(z) := \{\, (u, v+1) \mid (u,v) \in z \,\}.$$

Zunächst zeigen wir, daß $\mathsf{rn}(x)$ die oben formulierte Eigenschaft hat.

Lemma 5.3.34. *1.* $\mathsf{rn}(x) \in \mathsf{On}$.
2. $x \subseteq V_{\mathsf{rn}(x)}$.
3. $x \subseteq V_\alpha \to \mathsf{rn}(x) \subseteq \alpha$.

Beweis. 1. $\in$-Induktion nach x. Es ist $\mathsf{rn}(x) = \bigcup\{\, \mathsf{rn}(y)+1 \mid y \in x \,\} \in \mathsf{On}$, da nach IH $\mathsf{rn}(y) \in \mathsf{On}$ für jedes $y \in x$.

2. $\in$-Induktion nach x. Sei $y \in x$. Dann gilt $y \subseteq V_{\mathsf{rn}(y)}$ nach IH, also $y \in \mathcal{P}(V_{\mathsf{rn}(y)}) = V_{\mathsf{rn}(y)+1} \subseteq V_{\mathsf{rn}(x)}$ wegen $\mathsf{rn}(y)+1 \subseteq \mathsf{rn}(x)$.

3. Induktion nach α. Sei $x \subseteq V_\alpha$. Zu zeigen $\mathsf{rn}(x) = \bigcup\{\, \mathsf{rn}(y)+1 \mid y \in x \,\} \subseteq \alpha$. Sei also $y \in x$. Zu zeigen ist dann $\mathsf{rn}(y)+1 \subseteq \alpha$. Wegen $x \subseteq V_\alpha$ ist $y \in V_\alpha$. Daraus folgt $y \subseteq V_\beta$ für ein $\beta \in \alpha$, denn im Fall $\alpha = \alpha'+1$ hat man $y \in V_{\alpha'+1} = \mathcal{P}(V'_\alpha)$ und damit $y \subseteq V'_\alpha$, und im Fall α Limeszahl hat man $y \in V_\alpha = \bigcup_{\beta\in\alpha} V_\beta$, also $y \in V_\beta$ und damit $y \subseteq V_\beta$ für ein $\beta \in \alpha$. - Nach IH folgt jetzt $\mathsf{rn}(y) \subseteq \beta$ und damit $\mathsf{rn}(y) \in \alpha$. □

Jetzt folgt leicht die oben als Ziel formulierte Aussage.

Korollar 5.3.35. $V = \bigcup_{\alpha\in\mathsf{On}} V_\alpha$.

Beweis. $\supseteq$ ist klar. $\subseteq$. Für jedes x gilt $x \subseteq V_{\mathsf{rn}(x)}$ nach Lemma 5.3.34(2), also $x \in V_{\mathsf{rn}(x)+1}$. □

V_α läßt sich jetzt charakterisieren als die Menge aller Mengen vom Rang kleiner als α.

Lemma 5.3.36. $V_\alpha = \{\, x \mid \mathsf{rn}(x) \in \alpha \,\}$.

Beweis. $\supseteq$. Sei $\mathsf{rn}(x) \in \alpha$. Wegen $x \subseteq V_{\mathsf{rn}(x)}$ folgt $x \in V_{\mathsf{rn}(x)+1} \subseteq V_\alpha$.

$\subseteq$. Induktion nach α. *Fall* 0. Klar. *Fall* $\alpha+1$. Sei $x \in V_{\alpha+1}$. Dann ist $x \in \mathcal{P}(V_\alpha)$, also $x \subseteq V_\alpha$. Für jedes $y \in x$ gilt also $y \in V_\alpha$ und damit $\mathsf{rn}(y) \in \alpha$ nach IH, also $\mathsf{rn}(y)+1 \subseteq \alpha$. Daher gilt $\mathsf{rn}(x) = \bigcup\{\, \mathsf{rn}(y)+1 \mid y \in x \,\} \subseteq \alpha$. *Fall* α Limeszahl. Sei $x \in V_\alpha$. Dann ist $x \in V_\beta$ für ein $\beta \in \alpha$, also $\mathsf{rn}(x) \in \beta$ nach IH, also $\mathsf{rn}(x) \in \alpha$. □

Aus $x \in y$ bzw. $x \subseteq y$ kann man auf die entsprechenden Beziehungen zwischen den Rängen schließen.

Lemma 5.3.37. *1.* $x \in y \to \mathsf{rn}(x) \in \mathsf{rn}(y)$.
2. $x \subseteq y \to \mathsf{rn}(x) \subseteq \mathsf{rn}(y)$.

Beweis. 1. Wegen $\mathsf{rn}(y) = \bigcup\{\, \mathsf{rn}(x) + 1 \mid x \in y \,\}$ ist dies klar. 2. Für jedes $z \in x$ gilt $\mathsf{rn}(z) \in \mathsf{rn}(y)$ nach (1), also $\mathsf{rn}(x) = \bigcup\{\, \mathsf{rn}(z) + 1 \mid z \in x \,\} \subseteq \mathsf{rn}(y)$. □

Weiter läßt sich zeigen, daß die Mengen α und V_α beide den Rang α haben.

Lemma 5.3.38. *1.* $\mathsf{rn}(\alpha) = \alpha$.
2. $\mathsf{rn}(V_\alpha) = \alpha$.

Beweis. 1. Induktion über α. Es gilt $\mathsf{rn}(\alpha) = \bigcup\{\, \mathsf{rn}(\beta) + 1 \mid \beta \in \alpha \,\}$, also nach IH $\mathsf{rn}(\alpha) = \bigcup\{\, \beta + 1 \mid \beta \in \alpha \,\} = \alpha$ nach Lemma 5.3.28(1).

2. Es ist

$$\begin{aligned} \mathsf{rn}(V_\alpha) &= \bigcup\{\, \mathsf{rn}(x) + 1 \mid x \in V_\alpha \,\} \\ &= \bigcup\{\, \mathsf{rn}(x) + 1 \mid \mathsf{rn}(x) \in \alpha \,\} \quad \text{nach Lemma 5.3.36} \\ &\subseteq \alpha. \end{aligned}$$

Sei umgekehrt $\beta \in \alpha$. Nach (1) ist $\mathsf{rn}(\beta) = \beta \in \alpha$, also

$$\beta = \mathsf{rn}(\beta) \in \bigcup\{\, \mathsf{rn}(x) + 1 \mid \mathsf{rn}(x) \in \alpha \,\} = \mathsf{rn}(V_\alpha).$$ □

Schließlich zeigen wir noch, daß eine Klasse $\mathcal{A}$ genau dann eine Menge ist, wenn die Ränge ihrer Elemente durch eine Ordinalzahl beschränkt werden können.

Lemma 5.3.39. *$\mathcal{A}$ ist Menge genau dann, wenn es ein α gibt mit $\forall y{\in}\mathcal{A}\,(\mathsf{rn}(y) \in \alpha)$.*

Beweis. $\to$. Sei $\mathcal{A} = x$. Aus Lemma 5.3.37(1) folgt dann, daß $\mathsf{rn}(x)$ ist das gesuchte α ist.
$\leftarrow$. Gelte $\mathsf{rn}(x) \in \alpha$ für alle $y \in \mathcal{A}$. Dann folgt $\mathcal{A} \subseteq \{\, y \mid \mathsf{rn}(y) \in \alpha \,\} = V_\alpha$. □

5.4 Kardinalzahlen

Wir behandeln jetzt die wichtigsten Eigenschaften von Kardinalzahlen. Sie werden in der Mathematik häufig benutzt.

Größenvergleiche zwischen Mengen

Wir definieren

$$\begin{aligned} |a| \le |b| &:\leftrightarrow \exists f.f\colon a \to b \text{ und } f \text{ injektiv}, \\ |a| = |b| &:\leftrightarrow \exists f.f\colon a \leftrightarrow b, \\ |a| < |b| &:\leftrightarrow |a| \le |b| \wedge |a| \ne |b|, \\ {}^b a &:= \{\, f \mid f\colon b \to a \,\}. \end{aligned}$$

Zwei Mengen a und b heißen *gleichmächtig*, wenn $|a| = |b|$ gilt. Man beachte, daß wir hier nicht $|a|$ definieren, sondern nur Relationen $|a| \le |b|$, $|a| = |b|$ und $|a| < |b|$.

Lemma 5.4.1. *1.* $|a \times b| = |b \times a|$.
2. $|{}^a({}^b c)| = |{}^{a \times b} c|$.
3. $|\mathcal{P}(a)| = |{}^a\{0, 1\}|$.
4. *(*CANTOR*).* $|a| < |\mathcal{P}(a)|$.

Beweis. (1) – (3) sind klar. 4. $f\colon a \to \mathcal{P}(a)$, $x \mapsto \{x\}$ ist injektiv. Angenommen, wir hätten ein $g\colon a \leftrightarrow \mathcal{P}(a)$. Betrachte

$$b := \{\, x \mid x \in a \wedge x \notin g(x) \,\}.$$

Dann ist $b \subseteq a$, also $b = g(x_0)$ für ein $x_0 \in a$. Es folgt $x_0 \in g(x_0) \leftrightarrow x_0 \notin g(x_0)$ und damit ein Widerspruch. □

Satz 5.4.2. *(*Cantor*,* Schröder*,* Bernstein*) Ist $a \subseteq b \subseteq c$ und $|a| = |c|$, so folgt $|b| = |c|$.*

Beweis. Sei $f\colon c \to a$ bijektiv und $r := c \setminus b$. Wir definieren rekursiv $g\colon \omega \to V$ durch

$$\begin{aligned} g(0) &= r, \\ g(n+1) &= f[g(n)]. \end{aligned}$$

Ferner setzen wir

$$\overline{r} := \bigcup_n g(n)$$

und definieren $i\colon c \to b$ durch

$$i(x) := \begin{cases} f(x), & \text{falls } x \in \overline{r}, \\ x, & \text{falls } x \notin \overline{r}. \end{cases}$$

Es genügt zu zeigen, daß (1) $\mathsf{rng}(i) = b$ und (2) i injektiv ist. Zu (1). Sei $x \in b$. Zu zeigen ist $x \in \mathsf{rng}(i)$. OBdA sei $x \in \overline{r}$. Wegen $x \in b$ ist dann $x \notin g(0)$. Also gibt es ein n mit $x \in g(n+1) = f[g(n)]$, also $x = f(y) = i(y)$ für ein $y \in \overline{r}$. Zu (2). Sei $x \neq y$. OBdA ist $x \in \overline{r}$, $y \notin \overline{r}$. Dann hat man aber $i(x) \in \overline{r}$, $i(y) \notin \overline{r}$, also $i(x) \neq i(y)$. □

Anmerkung. Der Satz von Cantor, Schröder und Bernstein läßt sich auch als Anwendung des Fixpunktsatzes von Knaster-Tarski auffassen; siehe dazu Gunter [14].

Korollar 5.4.3. $|a| \leq |b| \wedge |b| \leq |a| \to |a| = |b|$.

Beweis. Seien $f\colon a \to b$ und $g\colon b \to a$ injektiv. Dann gilt $(g \circ f)[a] \subseteq g[b] \subseteq a$ und $|(g \circ f)[a]| = |a|$. Nach dem Satz von Cantor, Schröder und Bernstein folgt $|b| = |g[b]| = |a|$. □

Kardinalzahlen, Alephfunktion

Unter einer Kardinalzahl verstehen wir eine Ordinalzahl, die zu keiner kleineren Ordinalzahl gleichmächtig ist, also

$$\alpha \text{ heißt } \textit{Kardinalzahl}\text{, wenn gilt } \forall \beta{<}\alpha\, (|\beta| \neq |\alpha|).$$

Hier und im folgenden schreiben wir aufgrund von Lemma 5.3.22 $\alpha < \beta$ für $\alpha \in \beta$ und $\alpha \leq \beta$ für $\alpha \subseteq \beta$.

Lemma 5.4.4. $|n| = |m| \to n = m$.

Beweis. Induktion über n. 0. Klar. $n+1$. Sei $f\colon n+1 \leftrightarrow m+1$. OBdA können wir $f(n) = m$ annehmen. Also gilt $f{\restriction}n\colon n \leftrightarrow m$ und damit $n = m$ nach IH, also auch $n+1 = m+1$. □

Korollar 5.4.5. *n ist Kardinalzahl.* □

Lemma 5.4.6. $|n| \neq |\omega|$.

Beweis. Annahme: $|n| = |\omega|$. Es ist $n \subseteq n+1 \subseteq \omega$, also nach dem Satz von Cantor, Schröder und Bernstein $|n| = |n+1|$, was nicht sein kann. □

Korollar 5.4.7. *ω ist Kardinalzahl.* □

Lemma 5.4.8. $\omega \leq \alpha \to |\alpha + 1| = |\alpha|$.

Beweis. Definiere $f\colon \alpha \to \alpha + 1$ durch

$$f(x) := \begin{cases} \alpha, & \text{falls } x = 0; \\ n, & \text{falls } x = n + 1; \\ x, & \text{sonst.} \end{cases}$$

Dann ist $f\colon \alpha \leftrightarrow \alpha + 1$. □

Korollar 5.4.9. *Ist $\omega \leq \alpha$ und α eine Kardinalzahl, so ist α eine Limeszahl.*

Beweis. Annahme: $\alpha = \beta + 1$. Dann ist $\omega \leq \beta < \alpha$, also $|\beta| = |\beta + 1|$ im Widerspruch zur Voraussetzung, daß α eine Kardinalzahl ist. □

Lemma 5.4.10. *Ist a eine Menge von Kardinalzahlen, so ist* $\sup(a)$ *($:= \bigcup a$) eine Kardinalzahl.*

Beweis. Andernfalls gäbe es ein $\alpha < \sup(a)$ mit $|\alpha| = |\sup(a)|$. Also $\alpha \in \bigcup a$ und damit $\alpha \in \beta \in a$ für eine Kardinalzahl β. Nach dem Satz von Cantor, Schröder und Bernstein folgt aber aus $\alpha \subseteq \beta \subseteq \bigcup a$ und $|\alpha| = |\bigcup a|$, daß $|\alpha| = |\beta|$. Wegen $\alpha \in \beta$ und β Kardinalzahl ist dies aber nicht möglich. □

Wir wollen jetzt zeigen, daß es zu jeder Ordinalzahl eine größere Kardinalzahl gibt. Es gilt sogar allgemeiner folgendes.

Satz 5.4.11. $\forall a \exists! \alpha . \forall \beta{<}\alpha (|\beta| \leq |a|) \wedge |\alpha| \not\leq |a|$. *$\alpha$ heißt* Hartogszahl *von a; sie wird mit $\mathcal{H}(a)$ bezeichnet.*

Beweis. Eindeutigkeit. Klar. Existenz. Sei $w := \{\, (b, r) \mid b \subset a \wedge r \text{ Wohlordnung auf } b \,\}$ und $\gamma_{(b,r)}$ die eindeutig bestimmte Ordinalzahl isomorph zu (b, r). Dann ist $\{\, \gamma_{(b,r)} \mid (b, r) \in w \,\}$ eine transitive Teilmenge von On, also eine Ordinalzahl α. Zu zeigen ist

1. $\beta < \alpha \to |\beta| \leq |a|$,
2. $|\alpha| \not\leq |a|$.

1. Sei $\beta < \alpha$. Dann ist β isomorph einem $\gamma_{(b,r)}$ mit $(b, r) \in w$, also existiert $f\colon \beta \leftrightarrow b$.

2. Annahme: $f\colon \alpha \to a$ injektiv. Dann ist $\alpha = \gamma_{(b,r)}$ für ein $b \subseteq a$ ($b := \mathsf{rng}(f)$), also $\alpha \in \alpha$, was nicht sein kann. □

Anmerkung. 1. Die Hartogszahl von a ist eine Kardinalzahl. Denn sei α Hartogszahl von a, $\beta < \alpha$. Wäre $|\beta| = |\alpha|$, so hätte man $|\alpha| = |\beta| \leq |a|$, was nicht sein kann.

2. Die Hartogszahl von β ist die kleinste Kardinalzahl α mit $\alpha > \beta$.

$\aleph\colon \mathsf{On} \to V$ wird rekursiv definiert durch

$$\begin{aligned} &\aleph_0 := \omega, \\ &\aleph_{\alpha+1} := \mathcal{H}(\aleph_\alpha), \\ &\aleph_\alpha := \sup\{\, \aleph_\beta \mid \beta < \alpha \,\} \text{ für } \alpha \text{ Limeszahl.} \end{aligned}$$

Lemma 5.4.12. *(Eigenschaften von $\aleph$).*

1. *$\aleph_\alpha$ ist Kardinalzahl.*
2. *$\alpha < \beta \to \aleph_\alpha < \aleph_\beta$.*
3. *$\forall \beta . \beta$ Kardinalzahl $\wedge \omega \leq \beta \to \exists \alpha (\beta = \aleph_\alpha)$.*

Beweis. 1. Induktion über α; klar. 2. Induktion über β. 0. Klar. $\beta + 1$.

$$\begin{aligned} &\alpha < \beta + 1 \\ &\alpha < \beta \vee \alpha = \beta \\ &\aleph_\alpha < \aleph_\beta \vee \aleph_\alpha = \aleph_\beta \\ &\aleph_\alpha < \aleph_{\beta+1}. \end{aligned}$$

β Limeszahl.

$$\begin{aligned}
&\alpha < \beta \\
&\alpha < \gamma \text{ für ein } \gamma < \beta \\
&\aleph_\alpha < \aleph_\gamma \le \aleph_\beta.
\end{aligned}$$

3. Sei α minimal mit $\beta \le \aleph_\alpha$. Ein solches α existiert, da sonst $\aleph\colon \mathsf{On} \to \beta$ injektiv wäre. Wir zeigen $\aleph_\alpha \le \beta$ durch Fallunterscheidung nach α. 0. Klar. $\alpha = \alpha' + 1$. Nach Wahl von α ist $\aleph_{\alpha'} < \beta$, also $\aleph_\alpha \le \beta$. α Limeszahl. Nach Wahl von α ist $\aleph_\gamma < \beta$ für alle $\gamma < \alpha$, also $\aleph_\alpha = \sup\{\, \aleph_\gamma \mid \gamma < \alpha \,\} \le \beta$. □

Wir zeigen noch, daß jede unendliche Ordinalzahl zu einer Kardinalzahl gleichmächtig ist.

Lemma 5.4.13. $(\forall \beta{\ge}\omega)\exists\alpha(|\beta| = |\aleph_\alpha|)$.

Beweis. Betrachte $\delta := \min\{\, \gamma \mid \gamma < \beta \wedge |\gamma| = |\beta| \,\}$. δ ist offenbar Kardinalzahl. Ferner ist $\delta \ge \omega$, denn andernfalls hätte man

$$\begin{aligned}
&\delta = n \\
&|n| = |\beta| \\
&n \subseteq n + 1 \subseteq \beta \\
&|n| = |n+1|,
\end{aligned}$$

was nicht sein kann. Also ist $\delta = \aleph_\alpha$ für ein α, und wir haben $|\delta| = |\beta| = |\aleph_\alpha|$. □

Produkte von Kardinalzahlen

Wir zeigen jetzt, daß stets gilt $|\aleph_\alpha \times \aleph_\alpha| = |\aleph_\alpha|$.

Auf $\mathsf{On} \times \mathsf{On}$ definieren wir eine Relation $\prec$ durch

$$\begin{aligned}
(\alpha,\beta) \prec (\gamma,\delta) :\leftrightarrow\ & \max\{\alpha,\beta\} < \max\{\gamma,\delta\} \vee \\
& (\max\{\alpha,\beta\} = \max\{\gamma,\delta\} \wedge \alpha < \gamma) \vee \\
& (\max\{\alpha,\beta\} = \max\{\gamma,\delta\} \wedge \alpha = \gamma \wedge \beta < \delta).
\end{aligned}$$

Lemma 5.4.14. *$\prec$ ist eine Wohlordnung auf* $\mathsf{On} \times \mathsf{On}$.

Beweis. : $\prec$ ist offenbar eine lineare Ordnung. Zum Beweis der Fundiertheit von $\prec$ betrachten wir ein $a \subseteq \mathsf{On} \times \mathsf{On}$ mit $a \ne \emptyset$. Dann gilt

$$\emptyset \ne \mathcal{A} := \{\, \alpha \mid \exists\rho,\mu((\rho,\mu) \in a \wedge \max\{\rho,\mu\} = \alpha) \,\} \subseteq \mathsf{On}.$$

Sei $\alpha_0 := \min(\mathcal{A})$. Dann gilt

$$\emptyset \ne \mathcal{A}_1 := \{\, \rho \mid \exists\mu((\rho,\mu) \in a \wedge \max\{\rho,\mu\} = \alpha_0) \,\} \subseteq \mathsf{On}.$$

Sei $\rho_0 := \min(\mathcal{A}_1)$. Dann gilt

$$\emptyset \ne \mathcal{A}_2 := \{\, \mu \mid (\rho_0,\mu) \in a \wedge \max\{\rho_0,\mu\} = \alpha_0) \,\} \subseteq \mathsf{On}.$$

Sei $\mu_0 := \min(\mathcal{A}_2)$. Dann gilt offenbar $(\rho_0,\mu_0) = \min_\prec(a)$. Schließlich ist stets $\widehat{(\alpha,\beta)}$ eine Menge, denn es ist $\widehat{(\alpha,\beta)} \subseteq \gamma \times \gamma$ mit $\gamma := \max\{\alpha,\beta\} + 1$. □

Korollar 5.4.15. $\mathsf{On} \times \mathsf{On}$ *ist isomorph zu* On *(bzgl. $\prec$ und $\in \restriction \mathsf{On}$).*

Beweis. Nach Lemma 5.4.14 ist $\prec$ eine Wohlordnung auf $\mathsf{On} \times \mathsf{On}$. Also gibt es nach Korollar 5.3.19 einen Isomorphismus auf eine transitive und damit auch ordinale Klasse. Diese Klasse kann keine Menge sein, da sonst $\mathsf{On} \times \mathsf{On}$ eine Menge wäre. Nach Lemma 5.3.23(3) ist aber On die einzige echte ordinale Klasse. □

Satz 5.4.16. $\aleph_\alpha \times \aleph_\alpha$ *ist isomorph zu* $\aleph_\alpha$. *(bzgl.* $\prec \restriction \aleph_\alpha \times \aleph_\alpha$ *und* $\in \restriction \aleph_\alpha$*).*

Beweis. Annahme: $\exists\alpha(\aleph_\alpha \times \aleph_\alpha$ nicht isomorph zu $\aleph_\alpha)$. Sei

$$\alpha_0 := \min\{\, \alpha \mid \aleph_\alpha \times \aleph_\alpha \text{ nicht isomorph zu } \aleph_\alpha \,\}.$$

Offenbar ist $\alpha_0 \neq 0$. Da $\aleph_{\alpha_0} \times \aleph_{\alpha_0}$ und $\aleph_{\alpha_0}$ wohlgeordnete Mengen sind, muß eine von ihnen isomorph zu einem echten Anfangsstück der anderen sein. Wir unterscheiden deshalb zwei Fälle.

Fall 1. $\aleph_{\alpha_0}$ isomorph zu $\widehat{(\beta,\gamma)}$ mit $\beta,\gamma < \aleph_{\alpha_0}$. Wähle $\delta < \aleph_{\alpha_0}$ mit $\beta,\gamma < \delta$. Dann ist $\widehat{(\beta,\gamma)} \subseteq \delta \times \delta$. Man erhält

$$\begin{aligned} |\aleph_{\alpha_0}| = |\widehat{(\beta,\gamma)}| \leq |\delta \times \delta| &= |\aleph_\tau \times \aleph_\tau| && \text{für ein } \tau < \alpha_0 \\ &= |\aleph_\tau| && \text{nach Wahl von } \alpha_0 \end{aligned}$$

und damit einen Widerspruch zu Lemma 5.4.12(2).

Fall 2. $\aleph_{\alpha_0} \times \aleph_{\alpha_0}$ ist isomorph zu $\beta < \aleph_{\alpha_0}$. Dann folgt

$$\begin{aligned} &|\aleph_{\alpha_0}| \leq |\aleph_{\alpha_0} \times \aleph_{\alpha_0}| = |\beta| \leq |\aleph_{\alpha_0}| \\ &|\aleph_{\alpha_0}| = |\beta| \end{aligned}$$

und damit ein Widerspruch zu $\aleph_{\alpha_0}$ Kardinalzahl, $\beta < \aleph_{\alpha_0}$. □

Korollar 5.4.17. *1.* $|\aleph_\alpha \times \aleph_\beta| = |\max\{\aleph_\alpha, \aleph_\beta\}|$.
2. $n \neq 0 \to |{}^n\aleph_\alpha| = |\aleph_\alpha|$.

Beweis. 1. OBdA sei $\alpha \leq \beta$. Dann gilt

$$|\aleph_\beta| \leq |\aleph_\alpha \times \aleph_\beta| \leq |\aleph_\beta \times \aleph_\beta| = |\aleph_\beta|.$$

2. Dies folgt durch Induktion nach n leicht aus Satz 5.4.16. □

5.5 Das Auswahlaxiom

Auswahlaxiom, Wohlordnungssatz und Zornsches Lemma

Eine Relation $\mathcal{R}$ auf $\mathcal{A}$ heißt *partielle Ordnung*, wenn für alle $x, y, z \in \mathcal{A}$ gilt

$$\begin{aligned} &\neg x\mathcal{R}x, && \text{Irreflexivität} \\ &x\mathcal{R}y \wedge y\mathcal{R}z \to x\mathcal{R}z, && \text{Transitivität.} \end{aligned}$$

Ein $x \in \mathcal{A}$ heißt *maximales Element*, wenn es kein $y \in \mathcal{A}$ gibt mit $x\mathcal{R}y$. Sei noch $\mathcal{B} \subseteq \mathcal{A}$. Ein $x \in \mathcal{A}$ heißt *obere Schranke* von $\mathcal{B}$, wenn gilt

$$\forall y{\in}\mathcal{B}.\, y\mathcal{R}x \vee y = x.$$

Satz 5.5.1. *Die folgenden Aussagen sind äquivalent.*

1. *Das Auswahlaxiom (AC)*
$$\forall x.\emptyset \notin x \to \exists f.f\colon x \to \bigcup x \wedge (\forall y{\in}x)(f(y) \in y).$$
2. *Der Wohlordnungssatz (WO)*
$$\forall a \exists r(r \text{ ist Wohlordnung auf } a).$$
3. *Das* ZORN*sche Lemma (ZL): Für jede nichtleere partielle Ordnung* $(P, <)$ *mit der Eigenschaft, daß jedes durch* $<$ *linear geordnete* $L \subseteq P$ *eine obere Schranke in* P *hat, gilt:* P *hat ein maximales Element.*

Beweis. (ZL) $\to$ (WO). Gegeben sei ein a. Sei

$$P := \{\, f \mid \exists\alpha(f\colon \alpha \to a \text{ injektiv}) \,\} \subseteq \mathcal{P}(\mathcal{H}(a) \times a).$$

P wird durch die echte Inklusion $\subsetneq$ partiell geordnet. Sei $L \subseteq P$ linear geordnet. Dann gilt $\bigcup L \in P$. Also ist $\bigcup L$ obere Schranke von L. Mit dem ZORNschen Lemma erhält man ein maximales Element $f_0 \in P$. Offenbar ist dann f_0 eine Bijektion einer Ordinalzahl α_0 auf a, und f_0 induziert eine Wohlordnung auf a.

(WO) $\to$ (AC). Sei $\emptyset \notin x$. Nach (WO) existiert eine Wohlordnung $<$ auf $\bigcup x$. Für jedes $y \in x$ induziert $<$ eine Wohlordnung auf y. Definiere

$$\begin{aligned} f\colon x &\to \bigcup x, \\ y &\mapsto \min{}_<(y) \in y. \end{aligned}$$

(AC) $\to$ (ZL). Sei $<$ partielle Ordnung auf $P \neq \emptyset$. Wir nehmen an, daß jedes durch $<$ linear geordnete $L \subseteq P$ eine obere Schranke in P hat. Nach (AC) existiert eine Auswahlfunktion f auf $\mathcal{P}(P) \setminus \{\emptyset\}$. Sei $z \notin P$ beliebig. Definiere

$$\mathcal{F}\colon \mathsf{On} \to V$$
$$\mathcal{F}(\alpha) = \begin{cases} f(\{\, y \mid y \in P \wedge y \text{ obere Schranke von } \mathcal{F}[\alpha] \wedge y \notin \mathcal{F}[\alpha] \,\}), & \text{falls } \{\dots\} \neq \emptyset; \\ z, & \text{sonst.} \end{cases}$$

Dann gibt es ein ρ mit $\mathcal{F}(\rho) = z$, denn sonst wäre $\mathcal{F}\colon \mathsf{On} \to P$ injektiv, im Widerspruch zu der Voraussetzung, daß P eine Menge ist. Setze $\rho_0 := \min\{\, \rho \mid \mathcal{F}(\rho) = z \,\}$. $\mathcal{F}[\rho_0]$ ist linear geordnet, und es gilt $\mathcal{F}[\rho_0] \subseteq P$. Nach Voraussetzung existiert eine obere Schranke $y_0 \in P$ von $\mathcal{F}[\rho_0]$. Wir zeigen jetzt, daß y_0 maximales Element in P ist. Annahme: $y_0 < y$ für ein $y \in P$. Dann ist y obere Schranke von $\mathcal{F}[\rho_0]$ und $y \notin \mathcal{F}[\rho_0]$. Dies widerspricht aber der Definition von ρ_0. □

Ab jetzt wird das Auswahlaxiom immer vorausgesetzt. Wir markieren jedoch jeden Satz und jede Definition mit (AC), wenn darin das Auswahlaxiom verwendet wird.

(AC) ist offenbar äquivalent zu dem Spezialfall, in dem je zwei $y_1, y_2 \in x$ disjunkt sind. Wir notieren noch die folgende äquivalente Fassung des Auswahlaxioms.

Lemma 5.5.2. *Folgende Aussagen sind äquivalent.*

1. *Das Auswahlaxiom (AC).*
2. *Zu jedem surjektiven $g\colon a \to b$ gibt es ein injektives $f\colon b \to a$ mit $(\forall x \in b)(g(fx) = x)$.*

Beweis. Übung. □

Kardinalität

α heißt *Kardinalität* (oder *Mächtigkeit*) von a, wenn α Kardinalzahl ist und es eine Bijektion $f\colon a \to \alpha$ gibt.

Satz 5.5.3. *(AC). Jede Menge hat genau eine Kardinalität.*

Beweis. Eindeutigkeit. Klar. Existenz. Sei $<$ eine Wohlordnung auf a. Dann existiert ein γ mit a isomorph zu γ. Also ist $\{\, \tau \mid |\tau| = |a| \,\} \neq \emptyset$ und damit $\min\{\, \tau \mid |\tau| = |a| \,\}$ Kardinalzahl. □

Offenbar ist $|a| = |b|$ genau dann, wenn die Kardinalität von a gleich der Kardinalität von b ist, und $|a| \leq |b|$ genau dann, wenn die Kardinalität von a kleiner oder gleich der Kardinalität von b ist. Deshalb können wir $|a|$ als Bezeichnung der Kardinalität von a wählen.

Eine Menge a heißt *endlich*, wenn sich a bijektiv auf eine natürliche Zahl abbilden läßt, und *unendlich* sonst. Mit (AC) folgt dann, daß a genau dann endlich ist, wenn $|a| < \omega$ gilt.

Lemma 5.5.4. *(AC). Sind $a, b \neq \emptyset$ und ist a oder b unendlich, so ist*

$$|a \times b| = \max\{|a|, |b|\}.$$

Beweis. Sei etwa $|a| = \max\{|a|, |b|\}$. Dann gilt

$$|a| \leq |a \times b| = ||a| \times |b|| \leq ||a| \times |a|| = |a|. \qquad \square$$

Satz 5.5.5. *(AC). Sei I unendlich oder $\sup_{i \in I} |A_i|$ unendlich. Dann gilt*

1. $|\bigcup_{i \in I} A_i| \leq \max\{|I|, \sup_{i \in I} |A_i|\}$.
2. *Hat man zusätzlich $(\forall i \in I)(A_i \neq \emptyset)$ und $(\forall i, j \in I)(i \neq j \to A_i \cap A_j = \emptyset)$, so gilt $=$.*

Beweis. 1. OBdA $\kappa := \sup_{i \in I} |A_i| \neq 0$. Wir wählen uns eine Wohlordnung $<$ von I und definieren mit Bezug auf diese Wohlordnung

$$f \colon \bigcup_{i \in I} A_i \to \bigcup_{i \in I} (\{i\} \times A_i),$$
$$f(x) = (\min\{\, i \in I \mid x \in A_i \,\}, x).$$

f ist offenbar injektiv. Also

$$\begin{aligned} |\bigcup_{i \in I} A_i| &\leq |\bigcup_{i \in I} (\{i\} \times A_i)| \\ &\leq |\bigcup_{i \in I} (\{i\} \times \kappa)| \\ &= |I \times \kappa| \\ &= \max\{|I|, \kappa\}. \end{aligned}$$

2. Wegen (1) genügt es zu zeigen, daß $|I|, |A_i| \leq |\bigcup_{i \in I} A_i|$. Die zweite Abschätzung ist klar. Zum Beweis der ersten wählen wir uns eine Wohlordnung $<$ von $\bigcup_{i \in I} A_i$ und definieren $f \colon I \to \bigcup_{i \in I} A_i$ durch $f(i) := \min_<\{\, x \mid x \in A_i \,\}$. Aufgrund der Voraussetzungen ist f injektiv. $\square$

a heißt *Dedekind-endlich*, wenn sich a nicht bijektiv auf eine echte Teilmenge b von a abbilden läßt, und andernfalls *Dedekind-unendlich*.

Satz 5.5.6. *(AC). a ist Dedekind-unendlich genau dann, wenn a unendlich ist.*

Beweis. $\to$. Sei $b \subsetneq a$ und $f \colon a \leftrightarrow b$. Annahme: $|a| < \omega$, etwa $|a| = n$. Dann existiert $c \subsetneq n$ und ein $g \colon n \leftrightarrow c$. Wir zeigen durch Induktion über n, daß dies nicht möglich ist, also daß gilt

$$\forall n \neg(\exists c \subsetneq n) \exists g (g \colon n \leftrightarrow c).$$

0. Klar. $n+1$. Sei $g \colon n+1 \leftrightarrow c$ und $c \subsetneq n+1$. OBdA ist $n \notin \mathsf{rng}(g{\restriction}n)$. Es folgt $g{\restriction}n \colon n \leftrightarrow c \setminus \{n\} \subsetneq n$ und damit ein Widerspruch zur IH.

$\leftarrow$. Sei $g \colon \omega \to a$ injektiv und $h \colon g[\omega] \leftrightarrow g[\omega \setminus 1]$ definiert durch

$$h = \{\, (g(n), g(n+1)) \mid n \in \omega \,\}.$$

Ein $f \colon a \leftrightarrow (a \setminus \{g(0)\})$ ist definiert durch

$$f(x) = \begin{cases} x, & \text{falls } x \in a \setminus g[\omega]; \\ h(x), & \text{sonst.} \end{cases} \qquad \square$$

Reguläre und singuläre Kardinalzahlen

κ, λ bezeichnen im folgenden Kardinalzahlen $\geq \omega$. Das Auswahlaxiom (AC) werden wir in diesem Abschnitt stets voraussetzen.

Definition 5.5.7. (AC).

1. $x \subseteq \kappa$ heißt *konfinal* in κ, wenn $\sup(x) = \kappa$.
2. $\mathsf{cf}(\kappa) := \min\{\, |x| \mid x \subseteq \kappa$ und x konfinal mit $\kappa \,\}$ heißt *Konfinalität* von κ.
3. κ heißt *regulär*, wenn $\mathsf{cf}(\kappa) = \kappa$.
4. κ heißt *singulär*, wenn $\mathsf{cf}(\kappa) < \kappa$.

Satz 5.5.8. *(AC).*

1. *$\omega = \aleph_0$ ist regulär.*
2. *$\aleph_{\alpha+1}$ ist regulär.*
3. *Ist β Limeszahl und gilt $\beta < \aleph_\beta$, so ist $\aleph_\beta$ singulär.*

Beweis. 1. Angenommen, ω ist singulär, also $\mathsf{cf}(\omega) < \omega$. Dann gibt es ein $x \subseteq \omega$ mit $|x| = n$ und $\sup(x) = \omega$. Dies ist aber nicht möglich (Beweis durch Induktion über n).

2. Annahme: $\aleph_{\alpha+1}$ singulär. Dann ist $\mathsf{cf}(\aleph_{\alpha+1}) \leq \aleph_a$. Es gibt also ein $x \subseteq \aleph_{\alpha+1}$ mit $|x| \leq \aleph_\alpha$ und $\sup(x) = \aleph_{\alpha+1}$. Dann gilt aber

$$\begin{aligned} \aleph_{\alpha+1} &= |\bigcup x| \\ &\leq \max\{|x|, \sup\{\, |y| \mid y \in x \,\}\} \quad \text{nach Satz 5.5.5(1)} \\ &\leq \aleph_\alpha, \end{aligned}$$

was nicht sein kann.

3. Sei β eine Limeszahl mit $\beta < \aleph_\beta$. Dann hat man $\aleph_\beta = \sup\{\, \aleph_\gamma \mid \gamma < \beta \,\}$ und es gilt $|\{\, \aleph_\gamma \mid \gamma < \beta \,\}| = |\beta| < \aleph_\beta$. Also ist $\aleph_\beta$ singulär. □

Nach Definition hat man zu jeder unendlichen Kardinalzahl eine Teilmenge $x \subseteq \kappa$, deren Kardinalität gleich $\mathsf{cf}(\kappa)$ ist, die sich also bijektiv auf $\mathsf{cf}(\kappa)$ abbilden läßt. Wir zeigen jetzt, daß man sogar annehmen kann, daß diese Bijektion ein Isomorphismus ist.

Lemma 5.5.9. *(AC). Sei κ unendliche Kardinalzahl. Dann gibt es eine in κ konfinale Teilmenge $x \subseteq \kappa$, die isomorph zu $\mathsf{cf}(\kappa)$ ist.*

Beweis. Sei $y \subseteq \kappa$, $\sup(y) = \kappa$, $|y| = \mathsf{cf}(\kappa)$ und $g\colon \mathsf{cf}(\kappa) \leftrightarrow y$. Durch transfinite Rekursion definieren wir

$$\begin{aligned} &\mathcal{F}\colon \mathsf{On} \to V, \\ &\mathcal{F}(\alpha) := \sup(\mathcal{F}[\alpha] \cup g[\alpha]) + 1. \end{aligned}$$

Sei $f := \mathcal{F}{\restriction}\mathsf{cf}(\kappa)$. Man sieht leicht:

1. $\alpha < \beta < \mathsf{cf}(\kappa) \to f(\alpha) < f(\beta) \wedge g(\alpha) < f(\beta)$.
2. $\mathsf{rng}(f) \subseteq \kappa$
3. $\mathsf{rng}(f)$ ist konfinal mit κ.

$\mathsf{rng}(f)$ ist dann das gesuchte x. □

Korollar 5.5.10. *(AC). Ist κ eine unendliche Kardinalzahl, so ist $\mathsf{cf}(\kappa)$ eine reguläre Kardinalzahl.*

Beweis. $\mathsf{cf}(\mathsf{cf}(\kappa)) \leq \mathsf{cf}(\kappa)$ ist klar. Zu zeigen ist also $\mathsf{cf}(\kappa) \leq \mathsf{cf}(\mathsf{cf}(\kappa))$. Nach dem vorigen Lemma existieren x, f mit $x \subseteq \kappa$, $\sup(x) = \kappa$ und $f\colon \mathsf{cf}(\kappa) \leftrightarrow x$ Isomorphismus. Weiter existiert $y \subseteq \mathsf{cf}(\kappa)$ mit $\sup(y) = \mathsf{cf}(\kappa)$ und $|y| = \mathsf{cf}(\mathsf{cf}(\kappa))$. Man sieht leicht, daß $\{\, f(\alpha) \mid \alpha \in y \,\}$ konfinal mit κ ist. Daraus folgt

$$\begin{aligned} \mathsf{cf}(\kappa) &\leq |\{\, f(\alpha) \mid \alpha \in y \,\}| \\ &= |y| \\ &= \mathsf{cf}(\mathsf{cf}(\kappa)) \end{aligned}$$

□

Satz 5.5.11. *(AC). (König). Sei κ unendliche Kardinalzahl. Dann ist $\kappa < |^{\mathsf{cf}(\kappa)}\kappa|$.*

Beweis. $\kappa = |^1\kappa| \leq |^{\mathsf{cf}(\kappa)}\kappa|$ ist klar. Es genügt also, aus der Annahme der Existenz einer Bijektion $f\colon \kappa \leftrightarrow {}^{\mathsf{cf}(\kappa)}\kappa$ einen Widerspruch herzuleiten. Gemäß Lemma 5.5.9 gibt es $x \subseteq \kappa$ mit $\sup(x) = \kappa$ und einen Isomorphismus $g\colon \mathsf{cf}(\kappa) \leftrightarrow x$. Für jedes $\alpha < \mathsf{cf}(\kappa)$ ist also $g(\alpha) < \kappa$ und deshalb

$$|\{\, f(\gamma)(\alpha) \mid \gamma < g(\alpha)\,\}| \leq |g(\alpha)| < \kappa,$$

also $\{\, f(\gamma)(\alpha) \mid \gamma < g(\alpha)\,\} \subsetneq \kappa$. Setze

$$\begin{aligned} &h\colon \mathsf{cf}(\kappa) \to \kappa, \\ &h(\alpha) := \min(\kappa \setminus \{\, f(\gamma)(\alpha) \mid \gamma < g(\alpha)\,\}). \end{aligned}$$

Den gesuchten Widerspruch erhalten wir, indem wir zeigen, daß für alle $\gamma < \kappa$ gilt $f(\gamma) \neq h$. Sei also $\gamma < \kappa$. Wählt man jetzt ein $\alpha < \mathsf{cf}(\kappa)$ mit $\gamma < g(\alpha)$, so ist $h(\alpha) \neq f(\gamma)(\alpha)$ nach Konstruktion von h. □

Kardinalzahlpotenzen, Kontinuumshypothese

Auch in diesem Abschnitt wird (AC) stets vorausgesetzt. Wir definieren

$$\aleph_\alpha^{\aleph_\beta} := |^{\aleph_\beta}\aleph_\alpha|.$$

Später werden wir auch eine Ordinalzahlpotenz einführen. Aus Kontext sollte jeweils klar sein, ob die Ordinalzahlpotenz oder die Kardinalzahlpotenz gemeint ist.

Satz 5.5.12. *(AC).*

1. $\aleph_\beta < \mathsf{cf}(\aleph_\alpha) \to \aleph_\alpha \leq \aleph_\alpha^{\aleph_\beta} \leq |\mathcal{P}(\aleph_\alpha)|$
2. $\mathsf{cf}(\aleph_\alpha) \leq \aleph_\beta \leq \aleph_\alpha \to \aleph_\alpha < \aleph_\alpha^{\aleph_\beta} \leq |\mathcal{P}(\aleph_\alpha)|$
3. $\aleph_\alpha \leq \aleph_\beta \to \aleph_\alpha^{\aleph_\beta} = |\mathcal{P}(\aleph_\beta)|$.

Beweis. 1.

$$\begin{aligned} \aleph_\alpha &\leq |^{\aleph_\beta}\aleph_\alpha| \\ &\leq |^{\aleph_\beta}(^{\aleph_\alpha}\{0,1\})| \\ &= |^{\aleph_\beta\times\aleph_\alpha}\{0,1\}| \\ &= |^{\aleph_\alpha}\{0,1\}| \quad \text{da } \aleph_\beta \leq \aleph_\alpha \\ &= |\mathcal{P}(\aleph_\alpha)|. \end{aligned}$$

2.

$$\begin{aligned} \aleph_\alpha &< |^{\mathsf{cf}(\aleph_\alpha)}\aleph_\alpha| \quad \text{Satz von König} \\ &\leq |^{\aleph_\beta}\aleph_\alpha| \\ &\leq |\mathcal{P}(\aleph_\alpha)| \qquad \text{wie bei (1).} \end{aligned}$$

3.

$$\begin{aligned} |\mathcal{P}(\aleph_\beta)| &= |^{\aleph_\beta}\{0,1\}| \\ &\leq |^{\aleph_\beta}\aleph_\alpha| \\ &\leq |^{\aleph_\beta\times\aleph_\alpha}\{0,1\}| \\ &= |^{\aleph_\beta}\{0,1\}| \\ &= |\mathcal{P}(\aleph_\beta)|. \end{aligned}$$

□

Wesentlich mehr kann man über Kardinalzahlpotenzen aussagen, wenn man die sogenannte *Kontinuumshypothese* annimmt; darunter versteht man die Aussage

$$|\mathcal{P}(\aleph_0)| = \aleph_1. \tag{CH}$$

Die naheliegende Erweiterung auf alle Kardinalzahlen ist die *verallgemeinerte Kontinuumshypothese*:

$$|\mathcal{P}(\aleph_\alpha)| = \aleph_{\alpha+1}. \tag{GCH}$$

Es ist ein offenes Problem, ob die Kontinuumshypothese in der kumulativen Typenstruktur gilt (Nr. 1 in HILBERTS Liste von mathematischen Problemen, die er in seinem Vortrag vor dem internationalen Mathematikerkongress in Paris 1900 aufgestellt hat). Bekannt ist jedoch, daß die Kontinuumshypothese von den restlichen Axiomen der Mengenlehre unabhängig ist. Die Verwendung von (CH), (GCH) wird im folgenden immer angezeigt.

Satz 5.5.13. *(GCH).*

1. $\aleph_\beta < \mathsf{cf}(\aleph_\alpha) \to \aleph_\alpha = \aleph_\alpha^{\aleph_\beta}$.
2. $\mathsf{cf}(\aleph_\alpha) \le \aleph_\beta \le \aleph_\alpha \to \aleph_\alpha^{\aleph_\beta} = \aleph_{\alpha+1}$.
3. $\aleph_\alpha \le \aleph_\beta \to \aleph_\alpha^{\aleph_\beta} = \aleph_{\beta+1}$.

Beweis. (2) und (3) folgen mit (GCH) aus dem vorigen Satz.

1. Sei $\aleph_\beta < \mathsf{cf}(\aleph_\alpha)$. Man beachte zunächst, daß gilt

$${}^{\aleph_\beta}\aleph_\alpha = \bigcup\{\, {}^{\aleph_\beta}\gamma \mid \gamma < \aleph_\alpha \,\}$$

Dies sieht man wie folgt. $\supseteq$ ist klar. $\subseteq$. Sei $f\colon \aleph_\beta \to \aleph_\alpha$. Wegen $|f[\aleph_\beta]| \le \aleph_\beta < \mathsf{cf}(\aleph_\alpha)$ ist $\sup(f[\aleph_\beta]) < \gamma < \aleph_\alpha$ für ein γ, also $f\colon \aleph_\beta \to \gamma$.

Man erhält

$$\begin{aligned} \aleph_\alpha &\le |{}^{\aleph_\beta}\aleph_\alpha| && \text{voriger Satz} \\ &= |\bigcup\{\, {}^{\aleph_\beta}\gamma \mid \gamma < \aleph_\alpha \,\}| && \text{nach der Vorbemerkung} \\ &\le \max\{|\aleph_\alpha|, \sup_{\gamma<\aleph_\alpha} |{}^{\aleph_\beta}\gamma|\,\} && \text{nach Satz 5.5.5(1)} \end{aligned}$$

Es genügt also zu zeigen, daß $|{}^{\aleph_\beta}\gamma| \le \aleph_\alpha$ ist für $\gamma < \alpha$. Sei also $\gamma < \alpha$.

$$\begin{aligned} |{}^{\aleph_\beta}\gamma| &\le |{}^{\aleph_\beta\times\gamma}\{0,1\}| \\ &\le |{}^{\aleph_\beta\times\aleph_\delta}\{0,1\}| \quad \text{für ein } \delta \text{ mit } |\gamma| \le \aleph_\delta < \aleph_\alpha \\ &\le \begin{cases} |\mathcal{P}(\aleph_\delta)| & \text{falls } \beta < \delta \\ |\mathcal{P}(\aleph_\beta)| & \text{falls } \delta \le \beta \end{cases} \\ &= \begin{cases} \aleph_{\delta+1} & \text{falls } \beta < \delta \\ \aleph_{\beta+1} & \text{falls } \delta \le \beta \end{cases} \\ &\le \aleph_\alpha. \end{aligned}$$

□

5.6 Ordinalzahlarithmetik

Wir definieren Addition, Multiplikation und Exponentiation für Ordinalzahlen und beweisen ihre grundlegenden Eigenschaften. Ferner behandeln wir die CANTORsche Normalform.

$$\begin{aligned} &\alpha + 0 := \alpha, \\ &\alpha + (\beta + 1) := (\alpha + \beta) + 1, \\ &\alpha + \beta := \sup\{\, \alpha + \gamma \mid \gamma < \beta \,\} \quad \text{falls } \beta \text{ Limeszahl.} \end{aligned}$$

Genauer definiert man $s_\alpha \colon \mathsf{On} \to V$ durch

$$\begin{aligned} &s_\alpha(0) := \alpha, \\ &s_\alpha(\beta + 1) := s_\alpha(\beta) + 1, \\ &s_\alpha(\beta) := \bigcup \mathsf{rng}(s_\alpha \restriction \beta) \quad \text{falls } \beta \text{ Limeszahl} \end{aligned}$$

und setzt dann $\alpha + \beta := s_\alpha(\beta)$.

Lemma 5.6.1. *(Eigenschaften der Ordinalzahladdition).*

1. $\alpha + \beta \in \mathsf{On}$.
2. $0 + \beta = \beta$.
3. $\exists \alpha, \beta (\alpha + \beta \neq \beta + \alpha)$.
4. $\beta < \gamma \to \alpha + \beta < \alpha + \gamma$.
5. *Es gibt* α, β, γ *mit* $\alpha < \beta$, *aber* $\alpha + \gamma \not< \beta + \gamma$.
6. $\alpha \leq \beta \to \alpha + \gamma \leq \beta + \gamma$.
7. *Ist* $\alpha \leq \beta$, *so gibt es genau ein* γ *mit* $\alpha + \gamma = \beta$.
8. *Ist* β *Limeszahl, so auch* $\alpha + \beta$.
9. $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

Beweis. 1. Induktion über β. *Fall* 0. Klar. *Fall* $\beta + 1$. Dann ist $\alpha + (\beta + 1) = (\alpha + \beta) + 1 \in \mathsf{On}$, da nach IH $\alpha + \beta \in \mathsf{On}$. *Fall* β Limeszahl. Dann ist $\alpha + \beta = \sup\{\, \alpha + \gamma \mid \gamma < \beta \,\} \in \mathsf{On}$, da nach IH $\alpha + \gamma \in \mathsf{On}$ für alle $\gamma < \beta$.

2. Induktion über β. *Fall* 0. Klar. *Fall* $\beta + 1$. Dann ist $0 + (\beta + 1) = (0 + \beta) + 1 = \beta + 1$, da nach IH $0 + \beta = \beta$. *Fall* β Limeszahl. Dann ist

$$\begin{aligned} 0 + \beta &= \sup\{\, 0 + \gamma \mid \gamma < \beta \,\} \\ &= \sup\{\, \gamma \mid \gamma < \beta \,\} && \text{nach IH} \\ &= \bigcup \beta \\ &= \beta, && \text{da } \beta \text{ Limeszahl.} \end{aligned}$$

3. $1 + \omega = \sup\{\, 1 + n \mid n \in \omega \,\} = \omega \neq \omega + 1$.

4. Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma + 1$. Dann ist

$$\begin{aligned} &\beta < \gamma + 1, \\ &\beta < \gamma \lor \beta = \gamma, \\ &\alpha + \beta < \alpha + \gamma \lor \alpha + \beta = \alpha + \gamma && \text{nach IH,} \\ &\alpha + \beta \leq \alpha + \gamma < (\alpha + \gamma) + 1 = \alpha + (\gamma + 1). \end{aligned}$$

Fall γ Limeszahl. Sei $\beta < \gamma$, also $\beta < \delta$ für ein $\delta < \gamma$. Dann gilt $\alpha + \beta < \alpha + \delta$ nach IH, also $\alpha + \beta < \sup\{\, \alpha + \delta \mid \delta < \gamma \,\} = \alpha + \gamma$.

5. $0 < 1$, aber $0 + \omega = \omega = 1 + \omega$

6. Wir bemerken zunächst, daß es niemals ein β geben kann mit $\alpha < \beta < \alpha + 1$, denn andernfalls hätte man im Fall $\beta \in \alpha$ den Widerspruch $\beta \in \alpha \in \beta$ und im Fall $\beta = \alpha$ den Widerspruch $\alpha \in \alpha$. Als zweite Vorbemerkung notieren wir, daß stets gilt

$$\alpha \leq \beta \to \alpha + 1 \leq \beta + 1,$$

denn im Fall $\beta + 1 < \alpha + 1$ hätte man $\alpha < \beta + 1 < \alpha + 1$, was wie gerade bewiesen nicht sein kann. – Wir zeigen jetzt die Behauptung $\alpha \leq \beta \to \alpha + \gamma \leq \beta + \gamma$ durch Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma + 1$. Dann ist

$$\begin{aligned} &\alpha + \gamma \leq \beta + \gamma && \text{nach IH,} \\ &(\alpha + \gamma) + 1 \leq (\beta + \gamma) + 1 && \text{nach der zweiten Vorbemerkung,} \\ &\alpha + (\gamma + 1) \leq \beta + (\gamma + 1) && \text{nach Definition.} \end{aligned}$$

Fall γ Limeszahl. Dann ist

$$\begin{aligned}
&\alpha + \delta \leq \beta + \delta && \text{für alle } \delta < \gamma \text{, nach IH,}\\
&\alpha + \delta \leq \sup\{\, \beta + \delta \mid \delta < \gamma \,\}\\
&\sup\{\, \alpha + \delta \mid \delta < \gamma \,\} \leq \sup\{\, \beta + \delta \mid \delta < \gamma \,\}\\
&\alpha + \gamma \leq \beta + \gamma && \text{nach Definition.}
\end{aligned}$$

7. Die Eindeutigkeit von γ folgt aus (4). Existenz: Sei $\alpha \leq \beta$. Nach (2) und (6) gilt $\beta = 0 + \beta \leq \alpha + \beta$. Sei γ die kleinste Ordinalzahl mit $\beta \leq \alpha + \gamma$. Wir zeigen jetzt $\beta = \alpha + \gamma$. *Fall* $\gamma = 0$. Dann ist $\beta \leq \alpha + \gamma = \alpha + 0 = \alpha \leq \beta$, also $\beta = \alpha + \gamma$. *Fall* $\gamma = \gamma' + 1$. Dann ist $\alpha + \gamma' < \beta$, also $(\alpha + \gamma') + 1 \leq \beta$ nach der ersten Vorbemerkung zu (6) und damit $\alpha + \gamma = \beta$. *Fall* γ Limeszahl. Dann ist $\alpha + \delta < \beta$ für alle $\delta < \gamma$, also $\alpha + \gamma = \sup\{\, \alpha + \delta \mid \delta < \gamma \,\} \leq \beta$ und damit $\alpha + \gamma = \beta$.

8. Sei β Limeszahl. Wir verwenden die in Lemma 5.3.27(1) gegebene Charakterisierung von Limeszahlen. $\alpha + \beta \neq 0$: Wegen $0 \leq \alpha$ ist $0 < \beta = 0 + \beta \leq \alpha + \beta$ nach (6). $\gamma < \alpha + \beta \to \gamma + 1 < \alpha + \beta$: Sei $\gamma < \alpha + \beta = \sup\{\, \alpha + \delta \mid \delta < \beta \,\}$, also $\gamma < \alpha + \delta$ für ein $\delta < \beta$, also $\gamma + 1 < (\alpha + \delta) + 1$ (nach der ersten Vorbemerkung zu (6)) und damit $\gamma + 1 < \alpha + (\delta + 1)$ mit $\delta + 1 < \beta$, also $\gamma + 1 < \sup\{\, \alpha + \delta \mid \delta < \beta \,\}$.

9. Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma + 1$. Dann ist

$$\begin{aligned}
(\alpha + \beta) + (\gamma + 1) &= [(\alpha + \beta) + \gamma] + 1\\
&= [\alpha + (\beta + \gamma)] + 1 \quad \text{nach IH}\\
&= \alpha + [(\beta + \gamma) + 1]\\
&= \alpha + [\beta + (\gamma + 1)]
\end{aligned}$$

Fall γ Limeszahl. Nach (8) ist dann auch $\beta + \gamma$ Limeszahl. Man erhält

$$\begin{aligned}
(\alpha + \beta) + \gamma &= \sup\{\, (\alpha + \beta) + \delta \mid \delta < \gamma \,\}\\
&= \sup\{\, \alpha + (\beta + \delta) \mid \delta < \gamma \,\} && \text{nach IH}\\
&= \sup\{\, \alpha + \varepsilon \mid \varepsilon < \beta + \gamma \,\} && \text{siehe unten}\\
&= \alpha + (\beta + \gamma).
\end{aligned}$$

Die Gleichheit der beiden Suprema sieht man wie folgt. Ist $\varepsilon < \beta + \gamma$, so ist $\varepsilon < \beta + \delta$ für ein $\delta < \gamma$ (nach Definition von $\beta + \gamma$) und damit $\alpha + \varepsilon < \alpha + (\beta + \delta)$. Ist umgekehrt $\delta < \gamma$, so ist $\beta + \delta < \beta + \gamma$, also $\alpha + (\beta + \delta) = \alpha + \varepsilon$ für ein $\varepsilon < \beta + \gamma$. $\square$

Die Ordinalzahlmultiplikation definieren wir durch

$$\begin{aligned}
&\alpha \cdot 0 := 0,\\
&\alpha \cdot (\beta + 1) := (\alpha \cdot \beta) + \alpha,\\
&\alpha \cdot \beta := \sup\{\, \alpha \cdot \gamma \mid \gamma < \beta \,\} \quad \text{falls } \beta \text{ Limeszahl.}
\end{aligned}$$

Wir schreiben $\alpha\beta$ für $\alpha \cdot \beta$.

Lemma 5.6.2. *(Eigenschaften der Ordinalzahlmultiplikation).*

1. $\alpha\beta \in \mathsf{On}$.
2. $0\beta = 0$, $1\beta = \beta$.
3. $\exists \alpha, \beta(\alpha\beta \neq \beta\alpha)$.
4. $0 < \alpha \wedge \beta < \gamma \to \alpha\beta < \alpha\gamma$.
5. *Es gibt* α, β, γ *mit* $0 < \gamma$ *und* $\alpha < \beta$, *aber* $\alpha\gamma \not< \beta\gamma$.
6. $\alpha \leq \beta \to \alpha\gamma \leq \beta\gamma$.
7. *Ist* $0 < \alpha$ *und* β *Limeszahl, so ist auch* $\alpha\beta$ *Limeszahl.*
8. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.
9. *Es gibt* α, β, γ *mit* $(\alpha + \beta)\gamma \neq \alpha\gamma + \alpha\gamma$.
10. $\alpha\beta = 0 \to \alpha = 0 \vee \beta = 0$.
11. $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.

12. Ist $0 < \beta$*, so gibt es genau ein* γ, ρ *mit* $\alpha = \beta\gamma + \rho$ *und* $\rho < \beta$*.*

Beweis. 1. Induktion über β. *Fall* 0. Klar. *Fall* $\beta + 1$. Dann ist $\alpha(\beta + 1) = (\alpha\beta) + \alpha \in \mathsf{On}$, da nach IH $\alpha\beta \in \mathsf{On}$. *Fall* β Limeszahl. Dann ist $\alpha\beta = \sup\{\, \alpha\gamma \mid \gamma < \beta \,\} \in \mathsf{On}$, da nach IH $\alpha\gamma \in \mathsf{On}$ für alle $\gamma < \beta$.

2. $0\beta = 0$: Induktion über β. *Fall* 0. Klar. *Fall* $\beta+1$. Dann ist $0(\beta+1) = (0\beta)+0 = 0$ nach IH. *Fall* β Limeszahl. $0\beta = \sup\{\, 0\gamma \mid \gamma < \beta \,\} = 0$ nach IH. – $1\beta = \beta$: Induktion über β. *Fall* 0. Klar. *Fall* $\beta+1$. Dann ist $1(\beta + 1) = (1\beta) + 1 = \beta + 1$ nach IH. *Fall* β Limeszahl. $1\beta = \sup\{\, 1\gamma \mid \gamma < \beta \,\} = \sup\{\, \gamma \mid \gamma < \beta \,\} = \beta$ nach IH.

3. Man beachte zunächst, daß für alle $n \in \omega$ gilt $n\omega = \sup\{\, nm \mid m < \omega \,\} = \omega$. Damit folgt $2\omega = \omega$, aber $\omega 2 = \omega(1 + 1) = \omega 1 + \omega = \omega + \omega > \omega$.

4. Sei $0 < \alpha$. Wir zeigen $\beta < \gamma \to \alpha\beta < \alpha\gamma$ durch Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma + 1$. Dann ist

$$\begin{array}{ll} \beta < \gamma + 1, & \\ \beta < \gamma \vee \beta = \gamma, & \\ \alpha\beta < \alpha\gamma \vee \alpha\beta = \alpha\gamma & \text{nach IH,} \\ \alpha\beta \leq \alpha\gamma < (\alpha\gamma) + \alpha = \alpha(\gamma + 1). & \end{array}$$

Fall γ Limeszahl. Sei $\beta < \gamma$, also $\beta < \delta$ für ein $\delta < \gamma$. Dann gilt $\alpha\beta < \alpha\delta$ nach IH, also $\alpha\beta < \sup\{\, \alpha\delta \mid \delta < \gamma \,\} = \alpha\gamma$.

5. Es ist $0 < \omega$ und $1 < 2$, aber $1\omega = \omega = 2\omega$.

6. Wir zeigen die Behauptung $\alpha \leq \beta \to \alpha\gamma \leq \beta\gamma$ durch Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma + 1$. Dann ist

$$\begin{array}{ll} \alpha\gamma \leq \beta\gamma \quad \text{nach IH,} & \\ (\alpha\gamma) + \alpha \leq (\beta\gamma) + \alpha \leq (\beta\gamma) + \beta & \text{nach Lemma 5.6.1(6) und (4)} \\ \alpha(\gamma + 1) \leq \beta(\gamma + 1) & \text{nach Definition.} \end{array}$$

Fall γ Limeszahl. Dann ist

$$\begin{array}{ll} \alpha\delta \leq \beta\delta & \text{für alle } \delta < \gamma \text{, nach IH,} \\ \alpha\delta \leq \sup\{\, \beta\delta \mid \delta < \gamma \,\}, & \\ \sup\{\, \alpha\delta \mid \delta < \gamma \,\} \leq \sup\{\, \beta\delta \mid \delta < \gamma \,\}, & \\ \alpha\gamma \leq \beta\gamma & \text{nach Definition.} \end{array}$$

7. Sei $0 < \alpha$ und β Limeszahl. Zum Beweis von $\alpha\beta$ Limeszahl verwenden wir wieder die in Lemma 5.3.27(1) gegebene Charakterisierung von Limeszahlen. $\alpha\beta \neq 0$: Wegen $1 \leq \alpha$ und $\omega \leq \beta$ ist $0 < \omega = 1\omega \leq \alpha\beta$ nach (6). $\gamma < \alpha\beta \to \gamma + 1 < \alpha\beta$: Sei $\gamma < \alpha\beta = \sup\{\, \alpha\delta \mid \delta < \beta \,\}$, also $\gamma < \alpha\delta$ für ein $\delta < \beta$, also $\gamma + 1 < \alpha\delta + 1 \leq \alpha\delta + \alpha = \alpha(\delta + 1)$ mit $\delta + 1 < \beta$, also $\gamma + 1 < \sup\{\, \alpha\delta \mid \delta < \beta \,\}$.

8. Zu zeigen ist $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$. OBdA sei $0 < \alpha$. Wir verwenden Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma + 1$. Dann ist

$$\begin{aligned} \alpha[\beta + (\gamma + 1)] &= \alpha[(\beta + \gamma) + 1] \\ &= \alpha(\beta + \gamma) + \alpha \\ &= (\alpha\beta + \alpha\gamma) + \alpha \quad \text{nach IH} \\ &= \alpha\beta + (\alpha\gamma + \alpha) \\ &= \alpha\beta + \alpha(\gamma + 1). \end{aligned}$$

Fall γ Limeszahl. Nach (7) ist dann auch $\alpha\gamma$ Limeszahl. Man erhält

$$\begin{aligned} \alpha(\beta + \gamma) &= \sup\{\, \alpha\delta \mid \delta < \beta + \gamma \,\} \\ &= \sup\{\, \alpha(\beta + \varepsilon) \mid \varepsilon < \gamma \,\} \\ &= \sup\{\, \alpha\beta + \alpha\varepsilon \mid \varepsilon < \gamma \,\} \quad \text{nach IH} \\ &= \sup\{\, \alpha\beta + \delta \mid \delta < \alpha\gamma \,\} \\ &= \alpha\beta + \alpha\gamma. \end{aligned}$$

9. $(1+1)\omega = 2\omega = \omega$, aber $1\omega + 1\omega = \omega + \omega$.
10. Ist $0 < \alpha, \beta$, also $1 \leq \alpha, \beta$, so folgt $0 < 1 \cdot 1 \leq \alpha\beta$.
11. Induktion über γ. OBdA sei $\beta \neq 0$. *Fall* 0. Klar. *Fall* $\gamma + 1$. Dann ist

$$\begin{aligned}(\alpha\beta)(\gamma+1) &= (\alpha\beta)\gamma + \alpha\beta \\ &= \alpha(\beta\gamma) + \alpha\beta \quad \text{nach IH} \\ &= \alpha(\beta\gamma + \beta) \quad \text{nach (8)} \\ &= \alpha[\beta(\gamma+1)]\end{aligned}$$

Fall γ Limeszahl. Nach (7) ist dann auch $\beta\gamma$ Limeszahl. Man erhält

$$\begin{aligned}(\alpha\beta)\gamma &= \sup\{\, (\alpha\beta)\delta \mid \delta < \gamma \,\} \\ &= \sup\{\, \alpha(\beta\delta) \mid \delta < \gamma \,\} \quad \text{nach IH} \\ &= \sup\{\, \alpha\varepsilon \mid \varepsilon < \beta\gamma \,\} \\ &= \alpha(\beta\gamma).\end{aligned}$$

12. Existenz: Sei $0 < \beta$, also $1 \leq \beta$ und deshalb $\alpha = 1\alpha \leq \beta\alpha$. Sei γ die kleinste Ordinalzahl mit $\alpha \leq \beta\gamma$. *Fall* $\alpha = \beta\gamma$. Setze $\rho = 0$. *Fall* $\alpha < \beta\gamma$. Ist dann $\gamma = \gamma' + 1$, so ist $\beta\gamma' < \alpha$. Es gibt also ein ρ mit $\beta\gamma' + \rho = \alpha$. Ferner ist $\rho < \beta$, denn aus $\rho \geq \beta$ folgt $\alpha = \beta\gamma' + \rho \geq \beta\gamma' + \beta = \beta(\gamma'+1) = \beta\gamma$ im Widerspruch zur Fallunterscheidungsannahme. Ist γ Limeszahl, so gilt $\alpha < \beta\gamma = \sup\{\, \beta\delta \mid \delta < \gamma \,\}$, also $\alpha < \beta\delta$ für ein $\delta < \gamma$, was nicht sein kann.

Eindeutigkeit: Gelte $\beta\gamma_1 + \rho_1 = \beta\gamma_2 + \rho_2$ mit $\rho_1, \rho_2 < \beta$. Ist etwa $\gamma_1 < \gamma_2$, so folgt

$$\begin{aligned}\beta\gamma_1 + \rho_1 &< \beta\gamma_1 + \beta \\ &= \beta(\gamma_1 + 1) \\ &\leq \beta\gamma_2 \\ &\leq \beta\gamma_2 + \rho_2\end{aligned}$$

und damit ein Widerspruch. Also ist $\gamma_1 = \gamma_2$, und deshalb auch $\rho_1 = \rho_2$. □

Korollar 5.6.3. *Jede Ordinalzahl α läßt sich eindeutig in der Form $\alpha = \omega\gamma + n$ darstellen. Hierbei ist $n = 0$ genau dann, wenn $\alpha = 0$ oder α Limeszahl ist.*

Beweis. Es ist nur noch zu zeigen, daß für jedes γ gilt $\omega\gamma = 0$ oder $\omega\gamma$ ist Limeszahl. Im Fall $\gamma = 0$ ist dies klar. Im Fall $\gamma + 1$ ist $\omega(\gamma+1) = \omega\gamma + \omega$ Limeszahl nach Lemma 5.6.1(8). Im Fall γ Limeszahl ist nach Lemma 5.6.2(7) auch $\omega\gamma$ Limeszahl. □

Die Ordinalzahlexponentiation definieren wir durch

$$\begin{aligned}&\alpha^0 := \begin{cases} 0, & \text{falls } \alpha = 0; \\ 1, & \text{sonst,} \end{cases} \\ &\alpha^{\beta+1} := \alpha^\beta \alpha, \\ &\alpha^\beta := \sup\{\, \alpha^\gamma \mid \gamma < \beta \,\} \quad \text{falls } \beta \text{ Limeszahl.}\end{aligned}$$

Lemma 5.6.4. *(Eigenschaften der Ordinalzahlexponentiation).*

1. $\alpha^\beta \in \mathsf{On}$.
2. $0^\beta = 0$, $1^\beta = \beta$.
3. $1 < \alpha \wedge \beta < \gamma \to \alpha^\beta < \alpha^\gamma$.
4. *Es gibt α, β, γ mit $1 < \gamma$ und $1 < \alpha < \beta$, aber $\alpha^\gamma \not< \beta^\gamma$.*
5. $\alpha \leq \beta \to \alpha^\gamma \leq \beta^\gamma$.
6. *Ist $1 < \alpha$ und β Limeszahl, so ist auch α^β Limeszahl.*
7. $\alpha^{\beta+\gamma} = \alpha^\beta \alpha^\gamma$.
8. $\alpha^{\beta\gamma} = (\alpha^\beta)^\gamma$.

9. $1 < \alpha \to \beta \leq \alpha^\beta$.

Beweis. 1. Induktion über β. *Fall* 0. Klar. *Fall* $\beta+1$. Dann ist $\alpha^{\beta+1} = (\alpha^\beta)\alpha \in \mathsf{On}$, da nach IH $\alpha^\beta \in \mathsf{On}$. *Fall* β Limeszahl. Dann ist $\alpha^\beta = \sup\{\, \alpha^\gamma \mid \gamma < \beta \,\} \in \mathsf{On}$, da nach IH $\alpha^\gamma \in \mathsf{On}$ für alle $\gamma < \beta$.

2. $0^\beta = 0$: Induktion über β. *Fall* 0. $0^0 = 0$ gilt nach Definition. *Fall* $\beta+1$. Dann ist $0^{\beta+1} = (0^\beta)0 = 0$. *Fall* β Limeszahl. $0^\beta = \sup\{\, 0^\gamma \mid \gamma < \beta \,\} = 0$ nach IH. – $1^\beta = 1$: Induktion über β. *Fall* 0. Klar. *Fall* $\beta+1$. Dann ist $1^{\beta+1} = (1^\beta)1 = 1$ nach IH. *Fall* β Limeszahl. $1^\beta = \sup\{\, 1^\gamma \mid \gamma < \beta \,\} = \sup\{\, 1 \mid \gamma < \beta \,\} = 1$ nach IH.

3. Sei $1 < \alpha$. Wir zeigen $\beta < \gamma \to \alpha^\beta < \alpha^\gamma$ durch Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma+1$. Dann ist

$$\begin{aligned}
&\beta < \gamma + 1, \\
&\beta < \gamma \lor \beta = \gamma, \\
&\alpha^\beta < \alpha^\gamma \lor \alpha^\beta = \alpha^\gamma && \text{nach IH,} \\
&\alpha^\beta \leq \alpha^\gamma < \alpha^\gamma + \alpha^\gamma \leq \alpha^{\gamma+1}.
\end{aligned}$$

Fall γ Limeszahl. Sei $\beta < \gamma$, also $\beta < \delta$ für ein $\delta < \gamma$. Dann gilt $\alpha^\beta < \alpha^\delta$ nach IH, also $\alpha^\beta < \sup\{\, \alpha^\delta \mid \delta < \gamma \,\} = \alpha^\gamma$.

4. Für $1 < n$ gilt $n^\omega = \sup\{\, n^m \mid m < \omega \,\} = \omega$ und damit $2^\omega = \omega = 3^\omega$.

5. Wir zeigen die Behauptung $\alpha \leq \beta \to \alpha^\gamma \leq \beta^\gamma$ durch Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma+1$. Sei also $\alpha \leq \beta$. Dann ist

$$\begin{aligned}
\alpha^\gamma &\leq \beta^\gamma && \text{nach IH,} \\
\alpha^{\gamma+1} &= \alpha^\gamma \alpha \\
&\leq \beta^\gamma \alpha \\
&\leq \beta^\gamma \beta \\
&= \beta^{\gamma+1}.
\end{aligned}$$

Fall γ Limeszahl. Sei wieder $\alpha \leq \beta$. Dann ist

$$\begin{aligned}
&\alpha^\delta \leq \beta^\delta && \text{für alle } \delta < \gamma \text{, nach IH,} \\
&\alpha^\delta \leq \sup\{\, \beta^\delta \mid \delta < \gamma \,\} \\
&\sup\{\, \alpha^\delta \mid \delta < \gamma \,\} \leq \sup\{\, \beta^\delta \mid \delta < \gamma \,\} \\
&\alpha^\gamma \leq \beta^\gamma && \text{nach Definition.}
\end{aligned}$$

6. Sei $1 < \alpha$ und β Limeszahl. Zum Beweis von α^β Limeszahl verwenden wir wieder die in Lemma 5.3.27(1) gegebene Charakterisierung von Limeszahlen. $\alpha^\beta \neq 0$: Wegen $1 \leq \alpha$ ist $1 = 1^\beta \leq \alpha^\beta$. $\gamma < \alpha^\beta \to \gamma + 1 < \alpha^\beta$: Sei $\gamma < \alpha^\beta = \sup\{\alpha^\delta \mid \delta < \beta\}$, also $\gamma < \alpha^\delta$ für ein $\delta < \beta$, also $\gamma + 1 < \alpha^\delta + 1 \leq \alpha^\delta 2 \leq \alpha^{\delta+1}$ mit $\delta + 1 < \beta$, also $\gamma + 1 < \sup\{\, \alpha^\delta \mid \delta < \beta \,\}$.

7. Zu zeigen ist $\alpha^{\beta+\gamma} = \alpha^\beta \alpha^\gamma$. OBdA sei $\alpha \neq 0, 1$. Wir verwenden Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma + 1$. Dann ist

$$\begin{aligned}
\alpha^{\beta+\gamma+1} &= \alpha^\beta \alpha^\gamma \alpha && \text{nach IH} \\
&= \alpha^\beta \alpha^{\gamma+1}.
\end{aligned}$$

Fall γ Limeszahl. Man erhält

$$\begin{aligned}
\alpha^{\beta+\gamma} &= \sup\{\, \alpha^\delta \mid \delta < \beta + \gamma \,\} \\
&= \sup\{\, \alpha^{\beta+\varepsilon} \mid \varepsilon < \gamma \,\} \\
&= \sup\{\, \alpha^\beta \alpha^\varepsilon \mid \varepsilon < \gamma \,\} && \text{nach IH} \\
&= \sup\{\, \alpha^\beta \delta \mid \delta < \alpha^\gamma \,\} \\
&= \alpha^\beta \alpha^\gamma.
\end{aligned}$$

8. Zu zeigen ist $\alpha^{\beta\gamma} = (\alpha^\beta)^\gamma$. OBdA sei $\alpha \neq 0, 1$ und $\beta \neq 0$. Wir verwenden Induktion über γ. *Fall* 0. Klar. *Fall* $\gamma + 1$. Dann ist

$$\begin{aligned} \alpha^{\beta(\gamma+1)} &= \alpha^{\beta\gamma}\alpha^\beta \\ &= (\alpha^\beta)^\gamma \alpha^\beta \quad \text{nach IH} \\ &= (\alpha^\beta)^{\gamma+1}. \end{aligned}$$

Fall γ Limeszahl. Wegen $\alpha \neq 0, 1$ und $\beta \neq 0$ sind dann $\alpha^{\beta\gamma}$ und $(\alpha^\beta)^\gamma$ Limeszahlen und man erhält

$$\begin{aligned} \alpha^{\beta\gamma} &= \sup\{\, \alpha^\delta \mid \delta < \beta\gamma \,\} \\ &= \sup\{\, \alpha^{\beta\varepsilon} \mid \varepsilon < \gamma \,\} \\ &= \sup\{\, (\alpha^\beta)^\varepsilon \mid \varepsilon < \gamma \,\} \quad \text{nach IH} \\ &= (\alpha^\beta)^\gamma. \end{aligned}$$

9. Sei $1 < \alpha$. Wir zeigen $\beta \leq \alpha^\beta$ durch Induktion über β. *Fall* 0. Klar. *Fall* $\beta + 1$. Dann ist $\beta \leq \alpha^\beta$ nach IH, also

$$\begin{aligned} \beta + 1 &\leq \alpha^\beta + 1 \\ &\leq \alpha^\beta + \alpha^\beta \\ &\leq \alpha^{\beta+1}. \end{aligned}$$

Fall β Limeszahl. Man erhält

$$\begin{aligned} &= \sup\{\, \gamma \mid \gamma < \beta \,\} \\ &\leq \sup\{\, \alpha^\gamma \mid \gamma < \beta \,\} \quad \text{nach IH} \\ &= \alpha^\beta. \end{aligned}$$

□

Satz 5.6.5. *(*Cantor*-Normalform). Sei $\gamma \geq 2$. Jedes α läßt sich eindeutig darstellen in der Form*

$$\alpha = \gamma^{\alpha_1}\beta_1 + \cdots + \gamma^{\alpha_n}\beta_n \quad \textit{mit } \alpha \geq \alpha_1 > \cdots > \alpha_n \textit{ und } 0 < \beta_i < \gamma.$$

Beweis. Existenz. Induktion über α. Sei δ minimal mit $\alpha < \gamma^\delta$; ein solches δ existiert, da $\alpha \leq \gamma^\alpha$. Dann kann δ keine Limeszahl sein, da sonst $\alpha < \gamma^\varepsilon$ für ein $\varepsilon < \delta$ wäre. Ist $\delta = 0$, so folgt $\alpha = 0$ und die Behauptung ist trivial. Sei also $\delta = \alpha_1 + 1$, also

$$\gamma^{\alpha_1} \leq \alpha < \gamma^{\alpha_1+1}.$$

Division mit Rest ergibt

$$\alpha = \gamma^{\alpha_1}\beta_1 + \rho \quad \text{mit } \rho < \gamma^{\alpha_1}.$$

Offenbar ist $0 < \beta_1 < \gamma$. Ist jetzt $\rho = 0$, so sind wir fertig. Andernfalls haben wir

$$\rho = \gamma^{\alpha_2}\beta_2 + \cdots + \gamma^{\alpha_n}\beta_n \quad \text{nach IH.}$$

Zu zeigen bleibt $\alpha_1 > \alpha_2$. Dies gilt, da aus $\alpha_2 \geq \alpha_1$ folgt $\rho \geq \gamma^{\alpha_2} \geq \gamma^{\alpha_1}$, was nicht sein kann.

Eindeutigkeit. Sei

$$\gamma^{\alpha_1}\beta_1 + \cdots + \gamma^{\alpha_n}\beta_n = \gamma^{\alpha'_1}\beta'_1 + \cdots + \gamma^{\alpha'_m}\beta'_m.$$

und nehmen wir an, daß beide Darstellungen verschieden sind. Da keine Darstellung Verlängerung der anderen sein kann, gibt es dann ein $i \leq n, m$ mit $(\alpha_i, \beta_i) \neq (\alpha'_i, \beta'_i)$. Nach Lemma 5.6.1(4) können wir oBdA $i = 1$ annehmen. Zunächst gilt

$$\begin{aligned} \gamma^{\alpha_1}\beta_1 + \cdots + \gamma^{\alpha_{n-1}}\beta_{n-1} + \gamma^{\alpha_n}\beta_n &< \gamma^{\alpha_1}\beta_1 + \cdots + \gamma^{\alpha_{n-1}}\beta_{n-1} + \gamma^{\alpha_n+1} && \text{da } \beta_n < \gamma \\ &\leq \gamma^{\alpha_1}\beta_1 + \cdots + \gamma^{\alpha_{n-1}}(\beta_{n-1} + 1) && \text{da } \alpha_n < \alpha_{n-1} \\ &\leq \gamma^{\alpha_1}\beta_1 + \cdots + \gamma^{\alpha_{n-1}+1} \\ &\dots \\ &\leq \gamma^{\alpha_1}(\beta_1 + 1). \end{aligned}$$

Wäre etwa $\alpha_1 < \alpha_1'$, so hätte man $\gamma^{\alpha_1}\beta_1 + \cdots + \gamma^{\alpha_n}\beta_n < \gamma^{\alpha_1}(\beta_1 + 1) \leq \gamma^{\alpha_1+1} \leq \gamma^{\alpha_1'}$, was nicht sein kann. Also ist $\alpha_1 = \alpha_1'$. Wäre nun etwa $\beta_1 < \beta_1'$, so hätte man $\gamma^{\alpha_1}\beta_1 + \cdots + \gamma^{\alpha_n}\beta_n < \gamma^{\alpha_1}(\beta_1 + 1) \leq \gamma^{\alpha_1}\beta_1'$, was wieder nicht sein kann. Also ist $\beta_1 = \beta_1'$. □

Korollar 5.6.6. *(*CANTOR*-Normalform zur Basis ω). Jedes α läßt sich eindeutig darstellen in der Form*

$$\alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \quad \textit{mit } \alpha \geq \alpha_1 \geq \cdots \geq \alpha_n. \qquad \square$$

Eine Ordinalzahl α heißt *additive Hauptzahl*, wenn $\alpha \neq 0$ ist und aus $\beta, \gamma < \alpha$ stets folgt $\beta + \gamma < \alpha$.

Korollar 5.6.7. *Additive Hauptzahlen sind genau die Ordinalzahlen der Form ω^ξ.*

Beweis. Dies ergibt sich leicht aus der CANTOR-Normalform zur Basis ω. □

Korollar 5.6.8. *(*CANTOR*-Normalform zur Basis 2). Jedes α läßt sich eindeutig darstellen in der Form*

$$\alpha = 2^{\alpha_1} + \cdots + 2^{\alpha_n} \quad \textit{mit } \alpha \geq \alpha_1 > \cdots > \alpha_n. \qquad \square$$

Wir setzen noch $\omega_0 := 1$, $\omega_{k+1} := \omega^{\omega_k}$ und $\varepsilon_0 := \sup_{k<\omega} \omega_k$.

5.7 Normalfunktionen

In [45] hat VEBLEN den Begriff einer stetigen monotonen Funktion auf einem Abschnitt der Ordinalzahlen untersucht und eine gewisse Hierarchie von Normalfunktionen eingeführt. Seine Absicht war dabei, die CANTORsche Theorie der ε-Zahlen aus [5] zu erweitern.

Sei Ω eine eine reguläre Kardinalzahl $> \omega$ oder $\Omega = \mathsf{On}$. Ein wichtiges Beispiel ist $\Omega = \aleph_1$, also der Fall, daß Ω die Menge der abzählbaren Ordinalzahlen ist. Mit $\alpha, \beta, \gamma, \delta, \varepsilon, \xi, \eta, \zeta$ bezeichnen wir Elemente von Ω. Eine Funktion $\varphi\colon \Omega \to \Omega$ heißt *monoton*, wenn aus $\alpha < \beta$ stets folgt $\varphi\alpha < \varphi\beta$. φ heißt *stetig*, wenn gilt $\varphi\alpha = \sup_{\xi<\alpha} \varphi\xi$ für jede Limeszahl α. φ heißt *normal*, wenn φ monoton und stetig ist.

Lemma 5.7.1. *Für jede monotone Funktion φ gilt $\alpha \leq \varphi\alpha$.*

Beweis. Induktion über α. *Fall* 0. $0 \leq \varphi 0$. *Fall* $\alpha + 1$. $\alpha \leq \varphi\alpha < \varphi(\alpha + 1)$. *Fall* α Limeszahl. $\alpha = \sup_{\xi<\alpha} \xi \leq \sup_{\xi<\alpha} \varphi\xi \leq \varphi\alpha$. □

Eine Klasse $\mathcal{B} \subseteq \Omega$ heißt *beschränkt*, wenn $\sup(\mathcal{B}) \in \Omega$. Eine Klasse $\mathcal{A} \subseteq \Omega$ heißt *abgeschlossen*, wenn für jede beschränkte Teilklasse $\mathcal{B} \subseteq \mathcal{A}$ gilt $\sup(\mathcal{B}) \in \mathcal{A}$. Abgeschlossene unbeschränkte Klassen $\mathcal{A} \subseteq \Omega$ heißen *normal* in Ω oder auch *club* in Ω (club steht für "closed unbounded").

Ist zum Beispiel $\Omega = \Omega_1$, so ist jedes $\mathcal{B} \subseteq \Omega$ eine Menge, und $\mathcal{B}$ ist genau dann beschränkt, wenn $\mathcal{B}$ abzählbar ist. Ist $\Omega = \mathsf{On}$, so ist $\mathcal{B}$ genau dann beschränkt, wenn $\mathcal{B}$ eine Menge ist.

Nach Korollar 5.3.19 (zum Isomorphiesatz von MOSTOWSKI) gibt es zu jedem $\mathcal{A} \subseteq \mathsf{On}$ einen eindeutig bestimmten Isomorphismus einer ordinalen Klasse auf $\mathcal{A}$, also ein $f\colon \mathsf{On} \to \mathcal{A}$ bzw. $f\colon \alpha \to \mathcal{A}$. Diesen Isomorphismus nennt man die *Ordnungsfunktion* von $\mathcal{A}$. f heißt auch die *monotone Aufzählung* von $\mathcal{A}$.

Lemma 5.7.2. *Der Wertebereich einer Normalfunktion ist eine normale Klasse. Umgekehrt ist die Ordnungsfunktion einer normalen Klasse eine Normalfunktion.*

Beweis. Sei φ eine Normalfunktion. $\varphi[\Omega]$ ist unbeschränkt, da für jedes α gilt $\alpha \leq \varphi\alpha$. Wir zeigen jetzt, daß $\varphi[\Omega]$ abgeschlossen ist. Sei also $\mathcal{B} = \{\, \varphi\xi \mid \xi \in \mathcal{A} \,\}$ beschränkt, d.h. $\sup(\mathcal{B}) \in \Omega$. Wegen $\xi \leq \varphi\xi$ ist dann auch $\mathcal{A}$ beschränkt. Zu zeigen ist $\sup(\mathcal{B}) = \varphi\alpha$ für ein α. Hat $\mathcal{A}$ ein maximales Element, so sind wir fertig. Andernfalls sei $\alpha := \sup(\mathcal{A})$. Offenbar ist α eine Limeszahl. Dann gilt $\varphi\alpha = \sup_{\xi<\alpha} \varphi\xi = \sup_{\xi\in\mathcal{A}} \varphi\xi = \sup(\mathcal{B})$.

Sei umgekehrt $\mathcal{A}$ abgeschlossen und unbeschränkt. Wir definieren eine Funktion $\varphi\colon \Omega \to \mathcal{A}$ durch transfinite Rekursion, wie folgt.

$$\varphi\alpha := \min\{\, \gamma \in \mathcal{A} \mid \forall\xi(\xi < \alpha \to \varphi\xi < \gamma) \,\}.$$

φ ist wohldefiniert, da $\mathcal{A}$ unbeschränkt ist. Offenbar ist φ die Ordnungsfunktion von $\mathcal{A}$ und deshalb monoton. Es bleibt zu zeigen, daß φ stetig ist. Sei also α eine Limeszahl. Da $\varphi[\alpha]$ beschränkt ist (dies folgt aus $\varphi\xi < \varphi\alpha$ für $\xi < \alpha$) und $\mathcal{A}$ abgeschlossen ist, gilt $\sup_{\xi\in\alpha} \varphi\xi \in \mathcal{A}$, also nach Definition $\varphi\alpha = \sup_{\xi\in\alpha} \varphi\xi$. □

Lemma 5.7.3. *Die Fixpunkte einer Normalfunktion bilden eine normale Klasse.*

Beweis. (vgl. CANTOR [5, p. 242]). Sei φ eine Normalfunktion. Für jede Ordinalzahl α erhalten wir einen Fixpunkt $\beta \geq \alpha$ von φ durch

$$\beta := \sup\{\, \varphi^n \alpha \mid n \in \mathbb{N} \,\}.$$

Also ist die Klasse der Fixpunkte von φ unbeschränkt. Sie ist auch abgeschlossen, da für jede Klasse $\mathcal{B}$ von Fixpunkten von φ gilt $\varphi(\sup(\mathcal{B})) = \sup\{\, \varphi\alpha \mid \alpha \in \mathcal{B} \,\} = \sup\{\, \alpha \mid \alpha \in \mathcal{B} \,\} = \sup(\mathcal{B})$, d.h. $\sup(\mathcal{B})$ ist ein Fixpunkt von φ. $\square$

Die Ordnungsfunktion der Klasse der Fixpunkte einer Normalfunktion φ heißt nach VEBLEN die *erste Ableitung* φ' von φ. Zum Beispiel ist die erste Ableitung der Funktion ω^ξ die Funktion ε_ξ.

Lemma 5.7.4. *(*VEBLEN *[45, p. 284]). Sei $(\mathcal{A}_\gamma)_{\gamma<\beta}$ mit β Limeszahl eine fallende Folge normaler Klassen. Dann ist auch der Durchschnitt $\bigcap_{\gamma<\beta} \mathcal{A}_\gamma$ normal.*

Beweis. Unbeschränktheit. Sei α gegeben und $\delta_\gamma := \min\{\, \xi \in \mathcal{A}_\gamma \mid \xi > \alpha \,\}$. Dann ist $(\delta_\gamma)_{\gamma<\beta}$ schwach monoton. Sei $\delta := \sup_{\gamma<\beta} \delta_\gamma$. Dann ist $\delta \in \mathcal{A}_\gamma$ für jedes $\gamma < \beta$, da die $\mathcal{A}_\gamma$ fallen. Also ist $\alpha < \delta \in \bigcap_{\gamma<\beta} \mathcal{A}_\gamma$.

Abgeschlossenheit. Sei $\mathcal{B} \subseteq \bigcap_{\gamma<\beta} \mathcal{A}_\gamma$, $\mathcal{B}$ beschränkt. Dann gilt $\mathcal{B} \subseteq \mathcal{A}_\gamma$ für jedes $\gamma < \beta$ und deshalb $\sup(\mathcal{B}) \in \mathcal{A}_\gamma$. Also ist $\sup(\mathcal{B}) \in \bigcap_{\gamma<\beta} \mathcal{A}_\gamma$. $\square$

Wir definieren jetzt die VEBLEN-*Hierarchie von Normalfunktionen*. Sie geht aus von einer beliebigen gegebenen Normalfunktion $\varphi\colon \Omega \to \Omega$. Wir verwenden transfinite Rekursion, um für jedes $\beta \in \Omega$ eine Normalfunktion $\varphi_\beta\colon \Omega \to \Omega$ zu definieren:

$$\varphi_0 := \varphi,$$
$$\varphi_{\beta+1} := (\varphi_\beta)',$$
für Limeszahlen β sei φ_β die Ordnungsfunktion für $\bigcap_{\gamma<\beta} \varphi_\gamma[\Omega]$.

Etwa für $\varphi\alpha := 1 + \alpha$ erhält man $\varphi_\beta\alpha = \omega^\beta + \alpha$. Beginnt man mit $\varphi\alpha := \omega^\alpha$, so ist $\varphi_1\alpha = \varepsilon_\alpha$ und φ_2 zählt die kritischen ε-Zahlen auf, d.h. die Ordinalzahlen α mit $\varepsilon_\alpha = \alpha$.

Lemma 5.7.5. *Sei $\beta > 0$. Dann ist φ_β die Ordnungsfunktion der Klasse aller gemeinsamen Fixpunkte aller φ_γ für $\gamma < \beta$.*

Beweis. Wir müssen zeigen $\varphi_\beta[\Omega] = \{\, \xi \mid \forall\gamma(\gamma < \beta \to \varphi_\gamma\xi = \xi) \,\}$.

$\subseteq$. Dies beweist man durch transfinite Induktion nach β. Im Fall $\beta + 1$ ist jedes $\varphi_{\beta+1}\alpha$ ein Fixpunkt von φ_β und deshalb nach IH auch ein Fixpunkt aller φ_γ für $\gamma < \beta$. Ist β eine Limeszahl, so folgt die Behauptung aus $\varphi_\beta[\Omega] = \bigcap_{\gamma<\beta} \varphi_\gamma[\Omega]$.

$\supseteq$. Sei ξ mit $\forall\gamma(\gamma < \beta \to \varphi_\gamma\xi = \xi)$ gegeben. Ist β ein Nachfolger, so gilt $\xi \in \varphi_\beta[\Omega]$ nach Definition von φ_β. Ist β eine Limeszahl, so gilt $\xi \in \bigcap_{\gamma<\beta} \varphi_\gamma[\Omega] = \varphi_\beta[\Omega]$. $\square$

Hieraus folgt, daß gilt $\varphi_\gamma(\varphi_\beta\xi) = \varphi_\beta\xi$ für jedes $\gamma < \beta$.

Eine weitere Normalfunktion erhält man wie folgt. Aus jeder der normalen Klassen $\varphi_\beta[\Omega]$ nehme man den kleinsten Fixpunkt heraus. Die auf diese Weise gebildete Klasse ist wieder normal, läßt sich also durch eine Normalfunktion aufzählen. Diese Normalfunktion ordnet also jedem β die Ordinalzahl $\varphi_\beta 0$ zu.

Lemma 5.7.6. *Ist φ eine Normalfunktion mit $0 < \varphi 0$, so ist auch $\lambda\beta\, \varphi_\beta 0$ eine Normalfunktion.*

Beweis. Wir zeigen zunächst

$$\beta < \gamma \to \varphi_\beta 0 < \varphi_\gamma 0,$$

und zwar durch Induktion über γ. Sei also $\beta < \gamma$. Man beachte, daß $0 < \varphi_\beta 0$ nach IH oder im Fall $\beta = 0$ nach Annahme. Also ist 0 kein Fixpunkt von φ_β und deshalb $0 < \varphi_\gamma 0$. Hieraus folgt aber $\varphi_\beta 0 < \varphi_\beta(\varphi_\gamma 0) = \varphi_\gamma 0$.

Wir zeigen jetzt, daß $\lambda\beta\varphi_\beta 0$ stetig ist. Sei $\delta := \sup_{\beta<\gamma} \varphi_\beta 0$ mit γ Limeszahl. Wir müssen zeigen $\delta = \varphi_\gamma 0$. Da $\varphi_\beta 0 \in \varphi_\alpha[\Omega]$ für alle $\alpha \leq \beta < \gamma$ und da $\varphi_\alpha[\Omega]$ abgeschlossen ist, gilt $\delta \in \varphi_\alpha[\Omega]$, also $\delta \in \bigcap_{\alpha<\gamma} \varphi_\alpha[\Omega] = \varphi_\gamma[\Omega]$ und deshalb $\delta \geq \varphi_\gamma 0$. Andererseits ist $\varphi_\beta 0 < \varphi_\beta(\varphi_\gamma 0) = \varphi_\gamma 0$, also $\delta \leq \varphi_\gamma 0$. $\square$

Die Fixpunkte dieser Funktion, d.h. die Ordinalzahlen α mit $\varphi_\alpha 0 = \alpha$, heißen *stark kritische* Ordinalzahlen. Man beachte, daß sie von der oben als gegeben vorausgesetzten Normalfunktion $\varphi = \varphi_0$ abhängen. Ihre Ordnungsfunktion wird meist mit Γ bezeichnet. Nach Definition ist also $\Gamma_0 := \Gamma 0$ die kleinste Ordinalzahl β mit $\varphi_\beta 0 = \beta$.

Wir wollen jetzt eine Verallgemeinerung der CANTORschen Normalform herleiten, die auf der VEBLEN-Hierarchie aufbaut (anstelle von ω^ξ).

Lemma 5.7.7.

$$\varphi_{\beta_0}\alpha_0 < \varphi_{\beta_1}\alpha_1 \iff \begin{cases} \alpha_0 < \varphi_{\beta_1}\alpha_1, & \textit{falls } \beta_0 < \beta_1; \\ \alpha_0 < \alpha_1, & \textit{falls } \beta_0 = \beta_1; \\ \varphi_{\beta_0}\alpha_0 < \alpha_1, & \textit{falls } \beta_0 > \beta_1, \end{cases} \tag{5.2}$$

$$\varphi_{\beta_0}\alpha_0 = \varphi_{\beta_1}\alpha_1 \iff \begin{cases} \alpha_0 = \varphi_{\beta_1}\alpha_1, & \textit{falls } \beta_0 < \beta_1; \\ \alpha_0 = \alpha_1, & \textit{falls } \beta_0 = \beta_1; \\ \varphi_{\beta_0}\alpha_0 = \alpha_1, & \textit{falls } \beta_0 > \beta_1. \end{cases} \tag{5.3}$$

Beweis. $\Longleftarrow$. (5.2). Ist $\beta_0 < \beta_1$ und $\alpha_0 < \varphi_{\beta_1}\alpha_1$, so gilt $\varphi_{\beta_0}\alpha_0 < \varphi_{\beta_0}\varphi_{\beta_1}\alpha_1 = \varphi_{\beta_1}\alpha_1$. Ist $\beta_0 = \beta_1$ und $\alpha_0 < \alpha_1$, so folgt $\varphi_{\beta_0}\alpha_0 < \varphi_{\beta_1}\alpha_1$. Ist $\beta_0 > \beta_1$ und $\varphi_{\beta_0}\alpha_0 < \alpha_1$, so gilt $\varphi_{\beta_0}\alpha_0 = \varphi_{\beta_1}\varphi_{\beta_0}\alpha_0 < \varphi_{\beta_1}\alpha_1$. Für (5.3) schließt man analog.

$\Longrightarrow$. Ist die rechte Seite von (5.2) falsch, so gilt

$$\begin{cases} \alpha_1 \le \varphi_{\beta_0}\alpha_0, & \text{falls } \beta_1 < \beta_0; \\ \alpha_1 \le \alpha_0, & \text{falls } \beta_1 = \beta_0; \\ \varphi_{\beta_1}\alpha_1 \le \alpha_0, & \text{falls } \beta_1 > \beta_0, \end{cases}$$

also nach $\Longleftarrow$ (mit 0 und 1 vertauscht) $\varphi_{\beta_1}\alpha_1 < \varphi_{\beta_0}\alpha_0$ bzw. $\varphi_{\beta_1}\alpha_1 = \varphi_{\beta_0}\alpha_0$, also $\neg(\varphi_{\beta_0}\alpha_0 < \varphi_{\beta_1}\alpha_1)$. Ist die rechte Seite von (5.3) falsch, so gilt

$$\begin{cases} \alpha_0 \ne \varphi_{\beta_1}\alpha_1, & \text{falls } \beta_0 < \beta_1; \\ \alpha_0 \ne \alpha_1, & \text{falls } \beta_0 = \beta_1; \\ \varphi_{\beta_0}\alpha_0 \ne \alpha_1, & \text{falls } \beta_0 > \beta_1, \end{cases}$$

und damit nach $\Longleftarrow$ in (5.2) entweder $\varphi_{\beta_0}\alpha_0 < \varphi_{\beta_1}\alpha_1$ oder $\varphi_{\beta_1}\alpha_1 < \varphi_{\beta_0}\alpha_0$, also $\varphi_{\beta_0}\alpha_0 \ne \varphi_{\beta_1}\alpha_1$. □

Korollar 5.7.8. *Gilt $\beta_0 \le \beta_1$, so ist $\varphi_{\beta_0}\alpha \le \varphi_{\beta_1}\alpha$.*

Beweis. Gelte $\beta_0 < \beta_1$. Nach Lemma 5.7.7 (für $\le$) genügt es zu zeigen $\alpha \le \varphi_{\beta_1}\alpha$. Dies folgt aber aus Lemma 5.7.1. □

Korollar 5.7.9. *Gilt $\varphi_{\beta_0}\alpha_0 = \varphi_{\beta_1}\alpha_1$, so ist $\alpha_0 = \alpha_1$ und $\beta_0 = \beta_1$, falls $\alpha_0 < \varphi_{\beta_0}\alpha_0$ und $\alpha_1 < \varphi_{\beta_1}\alpha_1$.*

Beweis. Fall $\beta_0 = \beta_1$. Dann folgt $\alpha_0 = \alpha_1$ aus Lemma 5.7.7. *Fall* $\beta_0 < \beta_1$. Nach Lemma 5.7.7 gilt dann $\alpha_0 = \varphi_{\beta_1}\alpha_1 = \varphi_{\beta_0}\alpha_0$ im Widerspruch zu unserer Annahme. *Fall* $\beta_1 < \beta_0$. Ähnlich. □

Korollar 5.7.10. *Ist φ eine Normalfunktion mit $0 < \varphi 0$, so kann jeder Fixpunkt α von $\varphi = \varphi_0$ eindeutig in der Form $\alpha = \varphi_\beta\alpha'$ mit $\alpha' < \alpha$ geschrieben werden.*

Beweis. Es gilt $\alpha + 1 \le \varphi_{\alpha+1}0$ nach Lemma 5.7.6 und deshalb $\alpha < \varphi_{\alpha+1}\alpha$. Sei nun β minimal mit $\alpha < \varphi_\beta\alpha$. Nach Annahme ist $0 < \beta$. Da α Fixpunkt aller φ_γ mit $\gamma < \beta$ ist, gilt $\alpha = \varphi_\beta\alpha'$ für ein α'. Mit $\alpha < \varphi_\beta\alpha$ folgt $\alpha' < \alpha$.

Eindeutigkeit. Sei noch $\alpha = \varphi_{\beta_1}\alpha_1$ mit $\alpha_1 < \alpha$. Dann ist $\alpha_1 < \varphi_{\beta_1}\alpha_1$, also $\beta \le \beta_1$ nach Wahl von β. Wäre nun $\beta < \beta_1$, so folgte $\varphi_\beta\alpha = \varphi_\beta\varphi_{\beta_1}\alpha_1 = \varphi_{\beta_1}\alpha_1 = \alpha$ im Widerspruch zur Wahl von β. Also ist $\beta = \beta_1$ und damit $\alpha_1 = \alpha$. □

Wir zeigen jetzt, daß sich jede Ordinalzahl eindeutig in einer gewissen φ-Normalform schreiben läßt. Hierbei nehmen wir an, daß unsere Ausgangs-Normalfunktion $\varphi_0 = \varphi$ die Exponentiation zur Basis ω ist.

Satz 5.7.11 (φ-Normalform). *Sei $\varphi_0\xi := \omega^\xi$. Dann läßt sich jede Ordinalzahl α eindeutig schreiben in der Form*

$$\alpha = \varphi_{\beta_1}\alpha_1 + \cdots + \varphi_{\beta_n}\alpha_n$$

mit $\varphi_{\beta_1}\alpha_1 \geq \cdots \geq \varphi_{\beta_n}\alpha_n$ und $\alpha_i < \varphi_{\beta_i}\alpha_i$ für $i = 1, \ldots, k$. Ist $\alpha < \Gamma_0$, so gilt zusätzlich $\beta_i < \varphi_{\beta_i}\alpha_i$ für $i = 1, \ldots, n$.

Beweis. Existenz. Zunächst schreibe man α in CANTOR-Normalform $\alpha = \varphi_0\delta_1 + \cdots + \varphi_0\delta_n$ mit $\delta_1 \geq \cdots \geq \delta_n$. Jeden Summanden mit $\delta_i < \varphi_0\delta_i$ lasse man unverändert. Jeder andere Summand erfüllt $\delta_i = \varphi_0\delta_i$ und kann deshalb nach Korollar 5.7.10 durch $\varphi_\beta\alpha'$ mit $\alpha' < \varphi_\beta\alpha'$ ersetzt werden.

Eindeutigkeit. Sei

$$\alpha = \varphi_{\beta_1}\alpha_1 + \cdots + \varphi_{\beta_n}\alpha_n = \varphi_{\beta'_1}\alpha'_1 + \cdots + \varphi_{\beta'_m}\alpha'_m$$

und nehmen wir an, daß beide Darstellungen verschieden sind. Da keine Darstellung Verlängerung der anderen sein kann, gibt es dann ein $i \leq n, m$ mit $(\beta_i, \alpha_i) \neq (\beta'_i, \alpha'_i)$. Nach Lemma ordadd.g können wir oBdA $i = 1$ annehmen. Wäre etwa $\varphi_{\beta_1}\alpha_1 < \varphi_{\beta'_1}\alpha'_1$, so hätte man (da $\varphi_{\beta'_1}\alpha'_1$ additive Hauptzahl ist und $\varphi_{\beta_1}\alpha_1 \geq \cdots \geq \varphi_{\beta_n}\alpha_n$ gilt)

$$\varphi_{\beta_1}\alpha_1 + \cdots + \varphi_{\beta_n}\alpha_n < \varphi_{\beta'_1}\alpha'_1 \leq \varphi_{\beta'_1}\alpha'_1 + \cdots + \varphi_{\beta'_m}\alpha'_m,$$

was nicht sein kann.

Zu zeigen bleibt, daß im Fall $\alpha < \Gamma_0$ gilt $\beta_i < \varphi_{\beta_i}\alpha_i$ für $i = 1, \ldots, n$. Nehmen wir also an, daß $\varphi_{\beta_i}\alpha_i \leq \beta_i$ für ein i. Dann erhält man

$$\varphi_{\beta_i}0 \leq \varphi_{\beta_i}\alpha_i \leq \beta_i \leq \varphi_{\beta_i}0,$$

also $\varphi_{\beta_i}0 = \beta_i$ und deshalb

$$\Gamma_0 \leq \beta_i = \varphi_{\beta_i}0 \leq \varphi_{\beta_i}\alpha_i \leq \alpha. \qquad \square$$

Aus den $\varphi_\beta(\alpha)$ erhält man also ein eindeutiges Bezeichnungssystem für Ordinalzahlen unterhalb von $\Gamma_0 := \Gamma 0$. Man beachte jedoch, daß nach Definition von Γ_0 gilt $\Gamma_0 = \varphi_{\Gamma_0}0$.

5.8 Anmerkungen

Die hier vorgestellte Mengenlehre wird in der Literatur mit ZFC bezeichnet (ZERMELO-FRAENKEL-Mengenlehre mit Auswahlaxiom C für Choice). ZERMELO hat 1908 die Axiome angegeben mit Ausnahme des Regularitätsaxioms (VON NEUMANN, 1925) und des Ersetzungsschemas (FRAENKEL, 1922). Auch SKOLEM betrachtete Prinzipien, die zu den beiden nachträglichen Axiomen verwandt sind. In ZFC gibt es nur Mengen als Objekte, Klassen sind nur eine Sprechweise und können immer mit Formeln identifiziert werden.

Die in Abschnitt 5.7 definierte Hierarchie von Normalfunktionen wurde von VEBLEN [45] auf Funktionen mit mehr als einem Argument erweitert. SCHÜTTE hat in [30] diese Funktionen genauer untersucht und gezeigt, daß und wie sie für eine konstruktive Darstellung eines weit über Γ_0 hinausreichenden Abschnitts der Ordinalzahlen verwendet werden können. Dazu hat er sogenannte "Klammersymbole" zur Bezeichnung der mehrstelligen VEBLEN-Funktionen eingeführt.

BACHMANN hat die VEBLEN-Hierarchie unter Verwendung der ersten überabzählbaren Ordinalzahl Ω erweitert. Sein Ansatz wurde später durch Hinzunahme von Symbolen für höhere Zahlenklassen erweitert, zuerst von PFEIFFER für endliche Zahlenklassen und dann von ISLES für transfinite Zahlenklassen. Die entstehende Theorie war jedoch sehr kompliziert und es war schwierig, mit ihr zu arbeiten. Eine Idee von Feferman hat dann den Gegenstand sehr vereinfacht: Er führte Funktionen $\theta_\alpha\colon \mathsf{On} \to \mathsf{On}$ mit $\alpha \in \mathsf{On}$ ein, die wieder eine Hierarchie von Normalfunktionen bildeten und die VEBLEN-Hierarchie erweiterten; man schreibt meist $\theta\alpha\beta$ anstelle von $\theta_\alpha(\beta)$ und sieht θ als eine zweistellige Funktion an. Die Ordinalzahlen $\theta\alpha\beta$ werden wie folgt durch transfinite Rekursion über α definiert. Nehmen wir an, daß θ_ξ für jedes $\xi < \alpha$ schon definiert ist. Sei $C(\alpha, \beta)$ die Menge aller Ordinalzahlen, die sich aus Ordinalzahlen $< \beta$ und etwa den Konstanten $0, \aleph_1, \ldots, \aleph_\omega$ mit Hilfe der Funktionen $+$ und $\theta{\restriction}\{\, \xi \mid \xi < \alpha \,\} \times \mathsf{On}$ erzeugen lassen. Eine Ordinalzahl β heiße α-kritisch, wenn $\beta \notin C(\alpha, \beta)$. $\theta_\alpha \colon \mathsf{On} \to \mathsf{On}$ wird dann definiert als die Ordnungsfunktion der Klasse aller α-kritischen Ordinalzahlen.

Buchholz hat in [3] bemerkt, daß das zweite Argument β in $\theta\alpha\beta$ nicht wesentlich benutzt wird, und daß die Funktionen $\alpha \mapsto \theta\alpha\aleph_v$ mit $v = 0, 1, \ldots, \omega$ ein Bezeichnungssystem für Ordinalzahlen von derselben Stärke erzeugen wie das System mit der zweistelligen θ-Funktion. Er hat deshalb direkt Funktionen ψ_v mit $v \leq \omega$ definiert, die $\alpha \mapsto \theta\alpha\aleph_v$ entsprechen. Genauer definiert er $\psi_v\alpha$ für $\alpha \in \mathsf{On}$ und $v \leq \omega$ durch transfinite Rekursion über α, und zwar simultan für alle v, wie folgt.

$$\psi_v\alpha := \min\{\, \gamma \mid \gamma \notin C_v(\alpha) \,\},$$

wobei $C_v(\alpha)$ die Menge aller Ordinalzahlen ist, die sich aus den Ordinalzahlen $< \aleph_v$ durch die Funktionen $+$ und alle $\psi_u \restriction \{\, \xi \mid \xi < \alpha \,\}$ mit $u \leq \omega$ erzeugen lassen.

6. Beweistheorie der Arithmetik

In diesem Kapitel nehmen wir die Behandlung der Beweistheorie wieder auf. Allerdings verlassen wir den Bereich der reinen Logik und betrachten induktiv erzeugte Datenstrukturen; das Standardbeispiel ist die Menge der natürlichen Zahlen, die wir uns aus aus der Null mit der Nachfolgerfunktion erzeugt denken. Wir behandeln die Frage nach Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der Arithmetik. Es wird sich zeigen, daß wir hier eine scharfe Schranke erhalten, nämlich die Ordinalzahl ε_0. Als Folgerung hieraus werden wir in Abschnitt 6.4 zeigen, daß ein Normalisierungssatz, der die Teilformeleigenschaft impliziert, in der Arithmetik nicht gelten kann.

6.1 Gödelisierung von Ordinalzahlen

Um in arithmetischen Theorien über Ordinalzahlen sprechen zu können, verwenden wir eine GÖDELisierung von Ordinalzahlen. Dies ist sicher nur für abzählbare Mengen von Ordinalzahlen möglich. Wir beschränken uns hier auf den durch ε_0 bestimmten Abschnitt der Ordinalzahlen. Unsere GÖDELisierung können wir dann aus der CANTOR-Normalform erhalten. Mit α, β, γ bezeichnen wir in diesem Abschnitt stets Ordinalzahlen $< \varepsilon_0$.

Wir müssen ferner wissen, daß gewissen Relationen und Funktionen über den Ordinalzahlen $< \varepsilon_0$ berechenbare Relationen und Funktionen entsprechen. Dazu stellen wir folgendes fest.

Lemma 6.1.1. *Seien $\omega^{\alpha_m} + \cdots + \omega^{\alpha_0}$ und $\omega^{\beta_n} + \cdots + \omega^{\beta_0}$* CANTOR-*Normalformen (mit $m, n \geq -1$). Dann gilt*

$$\omega^{\alpha_m} + \cdots + \omega^{\alpha_0} < \omega^{\beta_n} + \cdots + \omega^{\beta_0}$$

genau dann, wenn es ein $i \geq 0$ gibt so daß $\alpha_{m-i} < \beta_{n-i}$, $\alpha_{m-i+1} = \beta_{n-i+1}, \ldots, \alpha_m = \beta_n$, oder $m < n$ und $\alpha_m = \beta_n, \ldots, \alpha_0 = \beta_{n-m}$.

Beweis. Übung. □

Wir verwenden die Bezeichnungen 1 für ω^0, a für $\omega^0 + \cdots + \omega^0$ mit a Exemplaren von ω^0 und $\omega^\alpha a$ für $\omega^\alpha + \cdots + \omega^\alpha$ wieder mit a Exemplaren von ω^α. □

Lemma 6.1.2. *Seien $\omega^{\alpha_m} + \cdots + \omega^{\alpha_0}$ und $\omega^{\beta_n} + \cdots + \omega^{\beta_0}$* CANTOR-*Normalformen. Dann gilt*

$$\omega^{\alpha_m} + \cdots + \omega^{\alpha_0} + \omega^{\beta_n} + \cdots + \omega^{\beta_0} = \omega^{\alpha_m} + \cdots + \omega^{\alpha_i} + \omega^{\beta_n} + \cdots + \omega^{\beta_0},$$

wobei i minimal ist so daß $\alpha_i \geq \beta_n$; falls es kein solches i gibt, sei $i = m + 1$ (also $\omega^{\beta_n} + \cdots + \omega^{\beta_0}$).

Beweis. Übung. □

Man kann auch eine kommutative Version der Addition definieren. Dies ist die sogenannte *natürliche Summe* oder HESSENBERG-*Summe* zweier Ordinalzahlen. Für CANTOR-Normalformen $\omega^{\alpha_m} + \cdots + \omega^{\alpha_0}$ und $\omega^{\beta_n} + \cdots + \omega^{\beta_0}$ wird sie definiert durch

$$(\omega^{\alpha_m} + \cdots + \omega^{\alpha_0}) \# (\omega^{\beta_n} + \cdots + \omega^{\beta_0}) := \omega^{\gamma_{m+n+1}} + \cdots + \omega^{\gamma_0},$$

wobei $\gamma_{m+n+1}, \ldots, \gamma_0$ eine fallende Permutation von $\alpha_m, \ldots, \alpha_0, \beta_n, \ldots, \beta_0$ ist.

Lemma 6.1.3. *$\#$ ist assoziativ, kommutativ und wachsend in beiden Argumenten.*

Beweis. Übung. □

Wir definieren jetzt eine Bijektion zwischen Ordinalzahlen $< \varepsilon_0$ und natürlichen Zahlen. Für diese Definition ist es nützlich, Ordinalzahlen darzustellen in der Form

$$\omega^{\alpha_m} a_m + \cdots + \omega^{\alpha_0} a_0 \quad \text{mit } \alpha_m > \cdots > \alpha_0 \text{ und } a_i \neq 0 \ (m \geq -1).$$

Für jede Ordinalzahl α definieren wir ihre GÖDELnummer $\ulcorner \alpha \urcorner$ induktiv durch

$$\ulcorner \omega^{\alpha_m} a_m + \cdots + \omega^{\alpha_0} a_0 \urcorner := \Big(\prod_{i \leq m} p_{\ulcorner \alpha_i \urcorner}^{a_i} \Big) - 1,$$

wobei p_n die n-te Primzahl ist, beginnend mit $p_0 := 2$. Für jede natürliche Zahl x definieren wir ihre entsprechende Ordinalzahl $\mathsf{o}(x)$ induktiv durch

$$\mathsf{o}\Big(\big(\prod_{i \leq m} p_i^{a_i} \big) - 1 \Big) := \sum_{i \leq m} \omega^{\mathsf{o}(i)} a_i,$$

wobei die Summe als natürliche Summe zu verstehen ist.

Lemma 6.1.4. *1.* $\mathsf{o}(\ulcorner \alpha \urcorner) = \alpha$.
2. $\ulcorner \mathsf{o}(x) \urcorner = x$.

Beweis. Übung. □

Wir haben also eine einfache Bijektion zwischen den Ordinalzahlen $< \varepsilon_0$ und den natürlichen Zahlen. Es ist nützlich, für einige Relationen und Funktionen auf den Ordinalzahlen besondere Bezeichnungen für die entsprechenden Relationen und Funktionen auf den natürlichen Zahlen einzuführen. Wir schreiben

$$\begin{array}{ll} x \prec y & \text{für } \mathsf{o}(x) < \mathsf{o}(y), \\ \omega^x & \text{für } \ulcorner \omega^{\mathsf{o}(x)} \urcorner, \\ x \oplus y & \text{für } \ulcorner \mathsf{o}(x) + \mathsf{o}(y) \urcorner, \\ xa & \text{für } \ulcorner \mathsf{o}(x) a \urcorner, \\ \omega_k & \text{für } \ulcorner \omega_k \urcorner, \end{array}$$

wobei $\omega_0 := 1$, $\omega_{k+1} := \omega^{\omega_k}$.

6.2 Beweisbarkeit von Anfangsfällen der transfiniten Induktion

Wir wollen jetzt Anfangsfälle der transfiniten Induktion in der PEANO-Arithmetik herleiten, also

$$(\forall x. \forall y {\prec} x\, A(y) \to A(x)) \to \forall x {\prec} a\, A(x)$$

für alle natürlichen Zahlen a und beliebige Formeln $A(x)$. Später werden wir sehen, daß unsere Resultate hier optimal sind in dem Sinn, daß für das volle System von Bezeichnungen für Ordinalzahlen $< \varepsilon_0$ das entsprechende Axiom der transfiniten Induktion bis ε_0, also

$$(\forall x. \forall y {\prec} x\, Py \to Px) \to \forall x\, Px$$

mit einem Prädikatensymbol P in der PEANO-Arithmetik unbeweisbar ist. Alle diese Resultate stammen von GENTZEN [10].

Unter einem *arithmetischen System* Z verstehen wir eine auf der Minimallogik aufgebaute Theorie (einschließlich der Gleichheitsaxiome) mit den folgenden Eigenschaften. Die Sprache von Z besteht aus einem festen (möglicherweise abzählbar unendlichen) Vorrat von Relations- und Funktionssymbolen, von denen wir annehmen, daß sie feste Relationen und Funktionen über den natürlichen Zahlen bezeichnen, für die ein Berechnungsverfahren bekannt ist. Unter den Funktionssymbolen müssen Symbole S für die

Nachfolgerfunktion und 0 für (die 0-stellige Funktion) Null vorkommen. Unter den Relationssymbolen müssen Symbole $=$ für die Gleichheit und $\prec$ für die eben eingeführte Ordnung vom Typ ε_0 der natürlichen Zahlen vorkommen. Um das allgemeine Prinzip der transfiniten Induktion formulieren zu können, nehmen wir noch an, daß auch ein einstelliges Relationssymbol P vorhanden ist, das wie eine freie Mengenvariable verwendet wird.

Terme werden aus Objektvariablen x, y, z mittels $f(t_1, \dots, t_m)$ aufgebaut, wobei f ein Funktionssymbol ist. Wir identifizieren einen geschlossenen Term mit seinem Wert; dies ist ein bequemer Weg, in unserem formalen System die Annahme auszudrücken, daß für jedes Funktionssymbol ein Berechnungsverfahren bekannt ist. Terme der Form $\mathsf{S}(\mathsf{S}(\dots\mathsf{S}(0)\dots))$ heißen *Numerale*. Wir verwenden die Bezeichnung $\mathsf{S}^n 0$ oder auch nur n für sie; gelegentlich schreiben wir auch $\underline{a}$, definiert durch $\underline{0} := 0$ und $\underline{n+1} := \mathsf{S}(\underline{n})$. *Formeln* werden aus atomaren Formeln $R(t_1, \dots, t_m)$ mit einem Relationssymbol R und aus $\perp$ mittels $A \to B$ und $\forall x A$ aufgebaut. Wie üblich kürzen wir $A \to \perp$ durch $\neg A$ ab.

Die *Axiome* von Z sollen immer die Peano-*Axiome* enthalten, also die Allabschlüsse von

$$\mathsf{S}(x) = \mathsf{S}(y) \to x = y, \tag{6.1}$$

$$\mathsf{S}(x) = 0 \to A, \tag{6.2}$$

$$A(0) \to (\forall x. A(x) \to A(\mathsf{S}(x))) \to \forall x A(x), \tag{6.3}$$

wobei $A(x)$ eine beliebige Formel ist. Um in unserem formalen System die Annahme auszudrücken, daß für jedes Relationssymbol ein Berechnungsverfahren bekannt ist, verwenden wir als Axiome $R\boldsymbol{n}$ falls $R\boldsymbol{n}$ wahr ist, und $\neg R\boldsymbol{n}$ falls $R\boldsymbol{n}$ falsch ist. Für $\prec$ fordern wir Irreflexivität und Transitivität als Axiome, und ferner – wie bei Schütte [32] – die Allabschlüsse von

$$x \prec 0 \to A, \tag{6.4}$$

$$z \prec y \oplus \omega^0 \to (z \prec y \to A) \to (z = y \to A) \to A, \tag{6.5}$$

$$x \oplus 0 = x, \tag{6.6}$$

$$x \oplus (y \oplus z) = (x \oplus y) \oplus z, \tag{6.7}$$

$$0 \oplus x = x, \tag{6.8}$$

$$\omega^x 0 = 0, \tag{6.9}$$

$$\omega^x \mathsf{S}(y) = \omega^x y \oplus \omega^x, \tag{6.10}$$

$$z \prec y \oplus \omega^{\mathsf{S}(x)} \to z \prec y \oplus \omega^{e(x,y,z)} m(x, y, z), \tag{6.11}$$

$$z \prec y \oplus \omega^{\mathsf{S}(x)} \to e(x, y, z) \prec \mathsf{S}(x), \tag{6.12}$$

wobei $\oplus$, $\omega^x y$, e und m die entsprechenden Funktionssymbole bezeichnen und A eine beliebige Formel ist. Wir erlauben auch eine beliebige Menge von wahren Formeln $\forall \boldsymbol{x}\, A$ mit A quantorenfrei und ohne P als Axiome. Solche Formeln heißen *Π_1-Formeln.*

Wir können ferner ein *ex-falso-quodlibet Axiom* oder sogar ein *Stabilitätsaxiom* für P hinzunehmen:

$$\forall x. \perp \to Px,$$
$$\forall x. \neg\neg Px \to Px.$$

Dies führt auf ein intuitionistisches arithmetisches System (wie die Heyting-Arithmetik HA) oder auf ein klassisches arithmetisches System (wie die Peano-Arithmetik PA). Man beachte, daß in Anwesenheit der Stabilitätsaxiome für alle Relationssymbole und für P sich (6.2), (6.4) und (6.5) durch ihre üblicheren klassischen Versionen ersetzen lassen, nämlich

$$\mathsf{S}(x) \neq 0, \tag{6.13}$$

$$x \not\prec 0, \tag{6.14}$$

$$z \prec y \oplus \omega^0 \to z \neq y \to z \prec y. \tag{6.15}$$

Wir werden auch *eingeschränkte* arithmetische Systeme Z_k betrachten. Sie sind wie Z definiert, aber mit dem Induktionsschema (6.3) eingeschränkt auf Formeln A der Stufe $\mathsf{lev}(A) \leq k$. Die *Stufe* einer Formel A ist definiert durch

$$\begin{aligned}\mathsf{lev}(R\boldsymbol{t}) &:= \mathsf{lev}(\bot) := 0,\\ \mathsf{lev}(A \to B) &:= \max(\mathsf{lev}(A)+1, \mathsf{lev}(B)),\\ \mathsf{lev}(\forall x A) &:= \max(1, \mathsf{lev}(A)).\end{aligned}$$

Der triviale Spezialfall $A(0) \to \forall x A(\mathsf{S}(x)) \to \forall x A(x)$, welcher der Fallunterscheidung entspricht, ist jedoch für beliebige Formeln $A(x)$ erlaubt. Dies werden wir in Satz 6.2.2 verwenden.

Satz 6.2.1. *(*Gentzen*). Das Schema der transfiniten Induktion bis ω_n, also jede Formel*

$$(\forall x.\forall y{\prec}x\, A(y) \to A(x)) \to \forall x{\prec}\omega_n\, A(x),$$

ist beweisbar in Z.

Beweis. Jeder Formel $A(x)$ ordnen wir eine Formel $A^+(x)$ zu (mit Bezug auf eine feste Variable x), und zwar durch

$$A^+(x) := \forall y.\forall z{\prec}y\, A(z) \to \forall z{\prec}y \oplus \omega^x\, A(z).$$

Wir zeigen zunächst

$$A(x) \text{ ist progressiv } \implies A^+(x) \text{ ist progressiv},$$

wobei "$B(x)$ ist *progressiv*" bedeutet $\forall x.\forall y{\prec}x\, B(y) \to B(x)$. Sei also $A(x)$ progressiv und

$$\forall y{\prec}x\, A^+(y). \tag{6.16}$$

Wir haben zu zeigen, daß $A^+(x)$ richtig ist. Gelte also weiter

$$\forall z{\prec}y\, A(z) \tag{6.17}$$

und $z \prec y \oplus \omega^x$. Zu zeigen ist dann $A(z)$.

Fall $x = 0$. Dann $z \prec y \oplus \omega^0$. Nach (6.5) genügt es, $A(z)$ aus der Annahme $z \prec y$ und auch aus der Annahme $z = y$ herzuleiten. Gilt $z \prec y$, so folgt $A(z)$ aus (6.17), und im Fall $z = y$ folgt $A(z)$ aus (6.17) und der Progressivität of $A(x)$.

Fall $\mathsf{S}(x)$. Aus $z \prec y \oplus \omega^{\mathsf{S}(x)}$ erhalten wir $z \prec y \oplus \omega^{e(x,y,z)} m(x,y,z)$ nach (6.11) und $e(x,y,z) \prec \mathsf{S}(x)$ nach (6.12). Aus (6.16) ergibt sich $A^+(e(x,y,z))$. Nach Definition von A^+ erhalten wir

$$\forall u{\prec}y \oplus \omega^{e(x,y,z)} v\, A(u) \to \forall u{\prec}(y \oplus \omega^{e(x,y,z)} v) \oplus \omega^{e(x,y,z)}\, A(u)$$

und also, unter Verwendung von (6.7) und (6.10)

$$\forall u{\prec}y \oplus \omega^{e(x,y,z)} v\, A(u) \to \forall u{\prec}y \oplus \omega^{e(x,y,z)} \mathsf{S}(v)\, A(u).$$

Ferner ergibt sich aus (6.17) und (6.9), (6.6)

$$\forall u{\prec}y \oplus \omega^{e(x,y,z)} 0\, A(u).$$

Mit Hilfe einer geeigneten Instanz des Induktionsschemas erhalten wir

$$\forall u{\prec}y \oplus \omega^{e(x,y,z)} m(x,y,z)\, A(u)$$

und deshalb $A(z)$.

Wir zeigen jetzt, durch Induktion über n, wie man für eine beliebige Formel $A(x)$ eine Herleitung von

$$\big(\forall x.\forall y{\prec}x\, A(y) \to A(x)\big) \to \forall x{\prec}\omega_n\, A(x)$$

erhalten kann. Gelte also die linke Seite, d.h. $A(x)$ sei progressiv.

Fall 0. Dann gilt $x \prec \omega^0$ und also $x \prec 0 \oplus \omega^0$ nach (6.8). Nach (6.5) genügt es, $A(x)$ aus $x \prec 0$ und auch aus $x = 0$ herzuleiten. Nun gilt $x \prec 0 \to A(x)$ nach (6.4), und $A(0)$ folgt aus der Progressivität von $A(x)$.

Fall $n + 1$. Da $A(x)$ progressiv ist, ist nach der obigen Überlegung auch $A^+(x)$ progressiv. Eine Anwendung der IH auf $A^+(x)$ liefert $\forall x{\prec}\omega_n\, A^+(x)$, und also $A^+(\omega_n)$ aufgrund der Progressivität von $A^+(x)$. Nun ergibt sich aus der Definition von $A^+(x)$ (mit (6.4) und (6.8)) sofort $\forall z{\prec}\omega^{\omega_n}\, A(z)$. □

Man beachte, daß wir im Induktionsschritt dieses Beweises die transfinite Induktion bis ω^{n+1} für $A(x)$ aus der transfiniten Induktion bis ω^n für die kompliziertere Formel $A^+(x)$ hergeleitet haben.

Wir wollen jetzt dieses Resultat verschärfen zu einem entsprechenden Satz für die Teilsysteme Z_k von Z (s. PARSONS [24]). Für $i \geq 1$ sei $\omega_i[m]$ definiert durch $\omega_1[m] := m$, $\omega_{i+1}[m] := \omega^{\omega_i[m]}$.

Satz 6.2.2. *Sei $1 \leq \ell \leq k$. Dann läßt sich in Z_k für jede Formel $A(x)$ einer Stufe $\leq \ell$ die transfinite Induktion bis $\omega_{k-\ell+2}[m]$ für beliebiges m herleiten, also*

$$(\forall x \forall y {\prec} x\, A(y) \to A(x)) \to \forall x {\prec} \omega_{k-\ell+2}[m]\, A(x).$$

Beweis. Man beachte zunächst, daß für jede Formel $A(x)$ einer Stufe $\ell \geq 1$ die im Beweis von Satz 6.2.1 konstruierte Formel $A^+(x)$ die Stufe $\ell + 1$ hat, und daß wir für den Beweis von

$$A(x) \text{ ist progressiv } \implies A^+(x) \text{ ist progressiv}$$

Induktion mit einer Induktionsformel der Stufe ℓ benutzt haben.

Sei jetzt (oBdA) $A(x)$ eine Formel der Stufe $\ell \geq 1$, und nehmen wir an, daß $A(x)$ progressiv ist. Sei $A^0 := A$, $A^{i+1} := (A^i)^+$. Dann ist $\mathsf{lev}(A^i) = \ell + i$, und folglich können wir in Z_k beweisen, daß $A^1, A^2, \ldots A^{k-\ell+1}$ sämtlich progressiv sind. Aus der Progressivität von $A^{k-\ell+1}(x)$ ergibt sich $A^{k-\ell+1}(0)$, $A^{k-\ell+1}(1)$, $A^{k-\ell+1}(2)$ und allgemein $A^{k-\ell+1}(m)$ für jedes m, also $A^{k-\ell+1}(\omega_1[m])$. Wegen

$$A^{k-\ell+1}(x) \iff (A^{k-\ell})^+(x) \iff \forall y.\forall z {\prec} y\, A^{k-\ell}(z) \to \forall z {\prec} y \oplus \omega^x\, A^{k-\ell}(z)$$

erhalten wir zunächst (mit $y = 0$) $\forall z {\prec} \omega_2[m]\, A^{k-\ell}(z)$ und dann $A^{k-\ell}(\omega_2[m])$ aus der Progressivität von $A^{k-\ell}$. Durch Wiederholung dieses Arguments erhalten wir schließlich $\forall z {\prec} \omega_{k-\ell+2}[m]\, A^0(z)$. $\square$

Unser nächstes Ziel ist ein Beweis, daß diese Schranken scharf sind. Genauer werden wir zeigen, daß man in Z (unabhängig davon, wieviele wahre Π_1-Formeln als Axiome verwendet werden) nicht die transfinite Induktion bis ε_0 beweisen kann, also die Formel

$$(\forall x.\forall y {\prec} x\, Py \to Px) \to \forall x Px$$

mit einem Relationssymbol P, und daß man in Z_k nicht die transfinite Induktion bis ω_{k+1} beweisen kann, also die Formel

$$(\forall x.\forall y {\prec} x\, Py \to Px) \to \forall x {\prec} \omega_{k+1}\, Px.$$

Dies wird sich durch Anwendung der Methode der Normalisierung auf arithmetische Systeme ergeben.

6.3 Normalisierung für die Arithmetik mit der Omega-Regel

Wir werden in Abschnitt 6.4 zeigen, daß für arithmetische Systeme Z kein Normalisierungssatz gilt, der für jede Herleitung einer Formel A in Z eine andere Herleitung derselben Formel A in Z liefert, die nur Formeln einer Stufe kleiner oder gleich der Stufe von A verwendet. Der Grund dafür sind die Induktionsaxiome, die von beliebiger Stufe sein können.

Hier umgehen wir diese Schwierigkeit in einer etwas drastischen Art und Weise: Wir geben unsere bisherige Auffassung auf, daß alle Herleitungen endlich sein müssen, und ersetzen die Induktionsaxiome durch eine Regel mit unendlich vielen Prämissen, die sogenannte ω-Regel. Sie wurde zuerst von HILBERT vorgeschlagen und später von LORENZEN, NOVIKOV und SCHÜTTE genauer studiert. Die ω-Regel erlaubt es, $\forall x A(x)$ aus $A(0), A(1), A(2), \cdots$ zu erschließen, also

$$\frac{\begin{matrix} \mathcal{D}_0 & \mathcal{D}_1 & & \mathcal{D}_i & \\ A(0) & A(1) & \ldots & A(i) & \ldots \end{matrix}}{\forall x A(x)}\;\omega$$

Herleitungen kann man dann als beschriftete unendliche (abzählbar verzweigte) Bäume auffassen, wobei die Beschriftung aus der hergeleiteten Formel und dem Namen der angewandten Regel besteht. Da wir unendliche Herleitungen induktiv definieren, muß jede solche Herleitung fundiert sein, d.h. sie kann keinen unendlichen absteigenden Pfad enthalten.

Offenbar macht die ω-Regel auch die Regel $\forall^+$ der Alleinführung überflüssig. Wir können deshalb auf freie Objektvariablen verzichten.

Ferner ist klar, daß jede Herleitung in einem arithmetischen System Z in eine unendliche Herleitung mit der ω-Regel übersetzt werden kann; dies werden wir in Lemma 6.3.3 ausführen. Die entstehende unendliche Herleitung hat eine beachtenswerte Eigenschaft: in jeder Anwendung der ω-Regel sind die Schnittränge der unendlich vielen Teilherleitungen beschränkt, und auch die Mengen ihrer freien Annahmen sind durch eine endliche Menge beschränkt. Hierbei ist der *Schnittrang* einer Herleitung die kleinste Zahl $\geq$ der Stufe jeder Teilherleitung, die durch $\to^+$ als Hauptprämisse von $\to^-$ oder die durch die ω-Regel als Hauptprämisse von $\forall^-$ gebildet wurde. Unter der *Stufe* einer Herleitung verstehen wir die Stufe ihres Typs, d.h. der von ihr hergeleiteten Formel. Eine unendliche Herleitung nennen wir normal, wenn ihr Schnittrang 0 ist, und wir werden unten zeigen, daß jede (möglicherweise unendliche) Herleitung von endlichem Schnittrang in eine Herleitung mit Schnittrang 0 umgeformt werden kann. Die entstehende normale Herleitung wird immer noch i.a. unendlich sein; man könnte also meinen, daß dieses Resultat nutzlos ist. Es wird sich jedoch zeigen, daß wir die Tiefe der entstehenden Herleitung in einer informativen Art und Weise beschränken können, und dies wird uns in Abschnitt 6.4 in die Lage versetzen, das gewünschte Resultat über die Unbeweisbarkeit der transfiniten Induktion zu erhalten. Dieses Programm wollen wir jetzt durchführen.

Wir führen die Systeme Z^∞ der ω-Arithmetik wie folgt ein. Z^∞ hat dieselbe Sprache und - abgesehen von den Induktionsaxiomen - dieselben Axiome wie Z. Herleitungen in Z^∞ sind unendliche Objekte. Es ist nützlich, für sie eine Termbezeichnung zu verwenden; d, e, f bezeichnen solche (unendlichen) Herleitungsterme. Für unsere Zwecke genügt es, nur Herleitungen zu betrachten, deren Tiefe unterhalb von ε_0 beschränkt ist.

Wir definieren den Begriff "d ist eine Herleitung der Tiefe $\leq \alpha$" (geschrieben $|d| \leq \alpha$) induktiv wie folgt.

(A). Jede Annahmenvariable u^A mit A geschlossen und jedes Axiom Ax^A ist eine Herleitung der Tiefe $\leq \alpha$, für jedes α.

($\to^+$). Ist d^B eine Herleitung der Tiefe $\leq \alpha_0 < \alpha$, so ist $(\lambda u^A d^B)^{A\to B}$ eine Herleitung der Tiefe $\leq \alpha$.

($\to^-$). Sind $d^{A\to B}$ und e^A Herleitungen der Tiefen $\leq \alpha_i < \alpha$ (i=1,2), so ist $(d^{A\to B}e^A)^B$ eine Herleitung der Tiefe $\leq \alpha$.

(ω). Sind $d_i^{A(i)}$ Herleitungen der Tiefen $\leq \alpha_i < \alpha$ $(i < \omega)$, so ist $\langle d_i^{A(i)} \rangle_{i<\omega}^{\forall x A(x)}$ eine Herleitung der Tiefe $\leq \alpha$.

($\forall^-$). Ist $d^{\forall x A(x)}$ eine Herleitung der Tiefe $\leq \alpha_0 < \alpha$, so ist $(d^{\forall x A(x)} i)^{A(i)}$ eine Herleitung der Tiefe $\leq \alpha$.

Man beachte, daß es in $\forall^-$ genügt, Zahlen statt Terme in den Nebenprämissen zu verwenden. Der Grund dafür ist, daß wir nur geschlossene Terme betrachten müssen, und solche Terme können wir hier mit Zahlen identifizieren. Der Schnittrang $\mathsf{cr}(d)$ einer Herleitung d ist wie folgt definiert.

$$
\begin{aligned}
&\mathsf{cr}(u^A) := \mathsf{cr}(\mathsf{Ax}^A) := 0,\\
&\mathsf{cr}(\lambda u d) := \mathsf{cr}(d),\\
&\mathsf{cr}(d^{A\to B}e^A) := \begin{cases} \max(\mathsf{lev}(A\to B), \mathsf{cr}(d), \mathsf{cr}(e)), & \text{falls } d = \lambda u d',\\ \max(\mathsf{cr}(d), \mathsf{cr}(e)), & \text{sonst,}\end{cases}\\
&\mathsf{cr}(\langle d_i\rangle_{i<\omega}) := \sup_{i<\omega} \mathsf{cr}(d_i),\\
&\mathsf{cr}(d^{\forall x A(x)} j) := \begin{cases} \max(\mathsf{lev}(\forall x A(x)), \mathsf{cr}(d)), & \text{falls } d = \langle d_i\rangle_{i<\omega},\\ \mathsf{cr}(d), & \text{sonst.}\end{cases}
\end{aligned}
$$

Offenbar ist $\mathsf{cr}(d) \in \mathbb{N} \cup \{\omega\}$ für alle d. Für unsere Zwecke genügt es, nur Herleitungen mit endlichem Schnittrang zu betrachten, d.h. mit $\mathsf{cr}(d) \in \mathbb{N}$. Es wird auch genügen, nur Herleitungen mit endlich vielen freien Variablen zu betrachten. Wir verwenden die Bezeichnung $d[u_1^{A_1}, \ldots, u_n^{A_n}]$ für Herleitungen mit freien Annahmevariablen unter $u_1^{A_1}, \ldots, u_n^{A_n}$. Mit $|d|$ bezeichnen wir das kleinste α so daß $|d| \leq \alpha$.

Lemma 6.3.1. *Ist d eine Herleitung der Tiefe $\leq \alpha$ mit freien Annahmevariablen unter $u, \boldsymbol{u}$ und vom Schnittrang $\mathsf{cr}(d) = k$, und ist e eine Herleitung der Tiefe $\leq \beta$, mit freien Annahmevariablen unter $\boldsymbol{u}$*

und vom Schnittrang $\mathsf{cr}(e) = \ell$, *so ist* $d[u := e]$ *eine Herleitung mit freien Annahmevariablen unter* $\boldsymbol{u}e$, *von der Tiefe* $|d[u := e]| \leq \beta + \alpha$ *und von einem Schnittrang* $\mathsf{cr}(d[u := e]) \leq \max(\mathsf{lev}(e), k, \ell)$.

Beweis. Einfache Induktion über die Tiefe von d. □

Unter Verwendung dieses Lemmas können wir jetzt unsere Systeme Z_k (d.h. Arithmetik mit Induktionsaxiomen beschränkt auf Formeln der Stufe $\leq k$) und Z in Z^∞ einbetten. In dieser Einbettung beziehen wir uns auf die Anzahl $n_I(d)$ von geschachtelten Anwendungen des Induktionsschemas innerhalb einer Z_k-Herleitung d. $n_I(d)$ ist durch Induktion über d wie folgt definiert.

$$\begin{aligned}
&n_I(u) := n_I(\mathsf{Ax}) := 0,\\
&n_I(\mathsf{Ind}) := 1,\\
&n_I(\mathsf{Ind}\,\boldsymbol{t}de) := \max(n_I(d), n_I(e)+1),\\
&n_I(de) := \max(n_I(d), n_I(e)), \quad \text{falls } d \text{ nicht von der Form } \mathsf{Ind}\,\boldsymbol{t}d_0 \text{ ist},\\
&n_I(\lambda ud) := n_I(\lambda xd) := n_I(dt) := n_I(d).
\end{aligned}$$

Ferner benötigen wir im nächsten Lemma den Begriff der *langen Normalform* einer Herleitung. Der Einfachheit halber beschränken wir uns wieder auf das $\to$-Fragment der minimalen Aussagenlogik; alle Überlegungen gelten aber genauso für die volle Sprache.

In Abschnitt 2.2 hatten wir die Gestalt normaler Herleitungen im $\to$-Fragment der minimalen Aussagenlogik studiert. Insbesondere hatten wir die Form von Ästen in einer normalen Herleitung genau analysiert und festgestellt, daß in jedem Ast alle Beseitigungsregeln vor allen Einführungsregeln kommen, und daß an einem eindeutig bestimmten Minimalknoten eine Minimalformel steht, die Teilformel aller Formeln im Einführungsteil und im Beseitigungsteil des Astes ist. Im Begriff der langen Normalform wird zusätzlich verlangt, daß jede Minimalformel atomar ist.

Für Terme des λ-Kalküls definiert man die *η-Expansion einer Variablen* durch

$$\eta_V(x^{\boldsymbol{\tau}\to\iota}) := \lambda \boldsymbol{z}^{\boldsymbol{\tau}}.x\eta_V(\boldsymbol{z}),$$

also durch Induktion über den Typ der Variablen. Die *η-Expansion eines Terms* läßt sich dann durch Induktion über Terme definieren durch

$$\eta\big(\lambda \boldsymbol{y}.(x\boldsymbol{M})^{\boldsymbol{\tau}\to\iota}\big) := \lambda \boldsymbol{y}, \boldsymbol{z}^{\boldsymbol{\tau}}.x\eta(\boldsymbol{M})\eta_V(\boldsymbol{z}).$$

Man beachte. daß stets $\eta(x) = \eta_V(x)$ gilt. – Damit ist klar:

Lemma 6.3.2. *Jeder Term läßt sich durch Normalisierung und anschließende η-Expansion in lange Normalform bringen.* □

Lemma 6.3.3. *Sei eine Z_k-Herleitung in langer Normalform (siehe Lemma 6.3.2, oder auch Abschnitt 6.5) gegeben mit $\leq m$ geschachtelten Anwendungen des Induktionsschemas* (6.3), *d.h. von*

$$A(0) \to (\forall x.A(x) \to A(\mathsf{S}(x))) \to \forall x A(x),$$

alle mit $\mathsf{lev}(A) \leq k$. *Wir betrachten Teilherleitungen d^B nicht von der Form $\mathsf{Ind}\,\boldsymbol{t}$ oder $\mathsf{Ind}\,\boldsymbol{t}d_0$. Für jede solche Teilherleitung und geschlossene Substitutionsinstanz $B\sigma$ von B konstruieren wir $(d^\infty_\sigma)^{B\sigma}$ in Z^∞ mit freien Annahmevariablen $u^{C\sigma}$ für u^C freie Annahmevariable von d, so daß $|d^\infty_\sigma| < \omega^{m+1}$ und $\mathsf{cr}(d^\infty_\sigma) \leq k$, und ferner so daß d durch $\to$-Einführung erzeugt ist gdw d^∞_σ es ist, sowie d durch $\forall$-Einführung erzeugt oder von der Form $\mathsf{Ind}\,\boldsymbol{t}d_0e$ ist gdw d^∞_σ durch die ω-Regel erzeugt ist.*

Beweis. Rekursion über solche Teilherleitungen d.

Fall u^C oder Ax. Dann können wir $u^{C\sigma}$ oder Ax nehmen.

Fall $\mathsf{Ind}\,\boldsymbol{t}de'$. Da die Herleitung in langer Normalform ist, gilt $e' = \lambda x\lambda v\, e$. Nach IH haben wir d^∞_σ und e^∞_σ. (Man beachte, daß weder d noch e eine der verbotenen Formen $\mathsf{Ind}\,\boldsymbol{t}$ oder $\mathsf{Ind}\,\boldsymbol{t}d_0$ haben kann, da beide in langer Normalform sind). Wir schreiben $e^\infty_\sigma(t, f)$ für $e^\infty_\sigma[x, v := t, f]$, und setzen

$$(\mathsf{Ind}\,\boldsymbol{t}d(\lambda x\lambda v\, e))^\infty_\sigma := \langle d^\infty_\sigma, e^\infty_\sigma(0, d^\infty_\sigma), e^\infty_\sigma(1, e^\infty_\sigma(0, d^\infty_\sigma)), \dots \rangle.$$

Nach IH ist $|e_\sigma^\infty| \leq \omega^{m-1}\cdot p$ und $|d_\sigma^\infty| \leq \omega^m\cdot q$ für gewisse $p, q < \omega$. Nach Lemma 6.3.1 erhalten wir

$$|e_\sigma^\infty(0, d_\sigma^\infty)| \leq \omega^m\cdot q + \omega^{m-1}\cdot p,$$
$$|e_\sigma^\infty(1, e_\sigma^\infty(0, d_\sigma^\infty))| \leq \omega^m\cdot q + \omega^{m-1}\cdot 2p$$

und so weiter, schließlich also

$$|(\mathsf{Ind}\,\boldsymbol{t}d(\lambda x\lambda v\, e))_\sigma^\infty| \leq \omega^m\cdot(q+1).$$

Für den Schnittrang ergibt sich nach IH $\mathsf{cr}(d_\sigma^\infty), \mathsf{cr}(e_\sigma^\infty) \leq k$ und deshalb

$$\mathsf{cr}(e_\sigma^\infty(0, d_\sigma^\infty)) \leq \max(\mathsf{lev}(A(0)), \mathsf{cr}(d_\sigma^\infty), \mathsf{cr}(e_\sigma^\infty)) \leq k,$$
$$\mathsf{cr}(e_\sigma^\infty(1, e_\sigma^\infty(0, d_\sigma^\infty))) \leq \max(\mathsf{lev}(A(1)), k, \mathsf{cr}(e_\sigma^\infty)) = k,$$

und so weiter, schließlich also

$$\mathsf{cr}((\mathsf{Ind}\,\boldsymbol{t}d(\lambda x\lambda v\, e))_\sigma^\infty) \leq k.$$

Fall $\lambda u^C\, d^B$. Nach IH haben wir $(d_\sigma^\infty)^{B\sigma}$, möglicherweise mit freien Annahmen $u^{C\sigma}$. Setze $(\lambda u\, d)_\sigma^\infty := \lambda u^{C\sigma}\, d_\sigma^\infty$.

Fall de mit d nicht von der Form $\mathsf{Ind}\,\boldsymbol{t}$ oder $\mathsf{Ind}\,\boldsymbol{t}d_0$. Nach IH haben wir d_σ^∞ und e_σ^∞. Da de Teilherleitung einer normalen Herleitung ist, kann d und also auch d_σ^∞ nicht durch $\to$-Einführung erzeugt sein. Also ist $(de)_\sigma^\infty := d_\sigma^\infty e_\sigma^\infty$ normal und $\mathsf{cr}(d_\sigma^\infty e_\sigma^\infty) = \max(\mathsf{cr}(d_\sigma^\infty), \mathsf{cr}(e_\sigma^\infty)) \leq k$. Ferner haben wir offenbar $|d_\sigma^\infty e_\sigma^\infty| < \omega^{m+1}$.

Fall $(\lambda x\, d)^{\forall x B(x)}$. Für jedes i und jede Substitutionsinstanz $B(i)\sigma$ haben wir nach IH $d_{\sigma,i}^\infty$. Setze $(\lambda x\, d)_\sigma^\infty := \langle d_{\sigma,i}^\infty\rangle_{i<\omega}$.

Fall $(\mathsf{Ind}\,\boldsymbol{t}det)^{B[x:=t]}$. Nach IH haben wir $((\mathsf{Ind}\,\boldsymbol{t}de)_\sigma^\infty)^{(\forall x B)\sigma}$, und $(\mathsf{Ind}\,\boldsymbol{t}de)_\sigma^\infty = \langle d_i\rangle_{i<\omega}$, wie im zweiten Fall des Beweises definiert. Setze $(\mathsf{Ind}\,\boldsymbol{t}det)_\sigma^\infty := d_i$, wobei i das Numeral mit demselben Wert wie $t\sigma$ ist.

Fall $(dt)^{B[x:=t]}$ mit d nicht von der Form $\mathsf{Ind}\,\boldsymbol{t}d_0e$. Nach IH haben wir $(d_\sigma^\infty)^{(\forall x B)\sigma}$. Da dt eine Teilherleitung einer normalen Herleitung ist, kann d nicht durch $\forall$-Einführung erzeugt sein, also d_σ^ω auch nicht durch die ω-Regel. Wir können deshalb setzen $(dt)_\sigma^\infty := d_\sigma^\infty i$, wobei i das Numeral mit demselben Wert wie $t\sigma$ ist. □

Eine Herleitung heißt *konvertierbar*, wenn sie von der Form $(\lambda u\, d)e$ oder $\langle d_i\rangle_{i<\omega}j$ ist. Sie kann dann konvertiert werden in $d[u := e]$ bzw. d_j. Hierbei entsteht $d[u := e]$ aus d durch Substituieren von e für alle freien Vorkommen von u in d. Eine Herleitung heißt *normal*, wenn sie keine konvertible Teilherleitung enthält. Man beachte, daß eine Herleitung genau dann normal ist, wenn sie den Schnittrang 0 hat.

Eine Herleitung heißt *einfache Anwendung*, wenn sie von der Form $d_0 d_1 \dots d_m$ mit d_0 eine Annahmevariable oder ein Axiom ist.

Wir wollen jetzt eine Operation definieren, die durch wiederholte Konversionen eine gegebene Herleitung in eine normale Herleitung umformt, wobei die Endformel erhalten bleibt und keine zusätzlichen freien Annahmevariablen eingeführt werden. Die übliche Methode zum Erreichen dieses Ziels muß hier an unsere spezielle Situation unendlicher Herleitungen angepaßt werden. Wir verwenden ein besonders einfaches Argument, das auf Tait [38] zurückgeht.

Lemma 6.3.4. *Zu jeder Herleitung d^A einer Tiefe $\leq \alpha$ und vom Schnittrang $k+1$ findet man eine Herleitung $(d^k)^A$ mit freien Annahmevariablen unter denen von d, mit einer Tiefe $\leq 2^\alpha$ und vom Schnittrang $\leq k$.*

Beweis. Induktion über α. Wir beschränken uns auf den Fall einer Herleitung der Form de mit $|d| \leq \alpha_1 < \alpha$ und $|e| \leq \alpha_2 < \alpha$, die keine einfache Anwendung ist. Zunächst betrachten wir den Unterfall, in dem $d^k = \lambda u\, d_1$ und $\mathsf{lev}(d) = k+1$ ist. Dann gilt $\mathsf{lev}(e) \leq k$ nach Definition der Stufen (die Stufe einer Herleitung war definiert als die Stufe der hergeleiteten Formel). Folglich hat $d_1[u := e^k]$ einen Schnittrang $\leq k$ nach Lemma 6.3.1. Ferner hat ebenfalls nach Lemma 6.3.1 $d_1[u := e^k]$ eine Tiefe $\leq 2^{\alpha_2} + 2^{\alpha_1} \leq 2^{\max(\alpha_2,\alpha_1)+1} \leq 2^\alpha$. Also können wir $(de)^k$ als $d_1[u := e^k]$ definieren.

Im Unterfall $d^k = \langle d_i \rangle_{i<\omega}$, $\mathsf{lev}(d) = k + 1$ und $e^k = j$ können wir $(de)^k$ als d_j wählen, da d_j einen Schnittrang $\leq k$ und eine Tiefe $\leq 2^\alpha$ hat.

Sind wir nicht in den obigen Unterfällen, so können wir einfach $(de)^k := d^k e^k$ setzen. Diese Herleitung hat offenbar eine Tiefe $\leq 2^\alpha$. Sie hat auch einen Schnittrang $\leq k$, was man wie folgt einsieht. Im Fall $\mathsf{lev}(d) \leq k+1$ sind wir fertig. Aber $\mathsf{lev}(d) \geq k+2$ ist unmöglich, da wir angenommen haben, daß de keine einfache Anwendung ist.

Um dies einzusehen, beachte man folgendes. Ist de keine einfache Anwendung, so muß es von der Form $d_0 d_1 \ldots d_n e$ sein mit d_0 weder Annahmevariable noch Axiom und d_0 auch nicht selbst von der Form $d'd''$; dann muß d_0 durch $\to$-Einführung oder durch die ω-Regel erzeugt sein, und es gäbe einen Schnitt mit einem Schnittrang $\geq k + 2$, was nach Annahme ausgeschlossen ist. □

Als unmittelbare Folgerung erhalten wir

Satz 6.3.5. *(Normalisierung für Z^∞). Für jede Herleitung d^A einer Tiefe $\leq \alpha$ und vom Schnittrang $\leq k$ findet man eine normale Herleitung $(d^*)^A$ mit freine Annahmen unter denen von d^A und einer Tiefe $\leq 2^\alpha_k$, wobei $2^\alpha_0 := \alpha, 2^\alpha_{m+1} = 2^{2^\alpha_m}$.* □

Wie in Satz 2.3.2 können wir jetzt die Struktur normaler Herleitungen in Z^∞ untersuchen. Insbesondere erhalten wir

Satz 6.3.6. *(Teilformeleigenschaft für Z^∞). Sei $d^B[u_1^{A_1}, \ldots, u_n^{A_n}]$ eine normale Herleitung in Z^∞ und e^C eine Teilherleitung von d^B. Dann ist C Teilformel von B oder von einem A_i.*

Beweis. Induktion über d. □

6.4 Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion

Wir verwenden jetzt die Technik der Normalisierung für die Arithmetik mit der ω-Regel zu einem Beweis, daß das Axiom der transfiniten Induktion bis ε_0 in Z unbeweisbar ist, also

$$\mathsf{Z} \nvdash (\forall x. \forall y {\prec} x\, Py \to Px) \to \forall x\, Px$$

mit einem Relationssymbol P, und daß die transfinite Induktion bis ω_{k+1} unbeweisbar ist in Z_k, also

$$\mathsf{Z}_k \nvdash (\forall x. \forall y {\prec} x\, Py \to Px) \to \forall x {\prec} \omega_{k+1}\, Px.$$

Offenbar genügt es, die Unbeweisbarkeit für auf der klassischen Logik basierende arithmetische Systeme zu zeigen. Wir können also annehmen, daß wir die klassischen Versionen (6.13), (6.14) und (6.15) der Axiome aus Abschnitt 6.2 verwendet haben.

Unser Beweis basiert auf einer Idee von Schütte, nämlich eine sogenannte *Progressionsregel* zu dem unendlichen System hinzuzunehmen. Diese Regel erlaubt es, Pj (wobei j eine beliebige Zahl ist) aus allen Pi für $i \prec j$ zu erschließen.

Eine Herleitung in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ einer Tiefe $\leq \alpha$ und vom Schnittrang $\leq k$ wird erzeugt durch die induktiven Klauseln von Abschnitt 6.3 und die zusätzliche Klausel

(Prog). Sind für alle $i \prec j$ Herleitungen d_i^{Pi} der Tiefe $\leq \alpha_i < \alpha$ bereits konstruiert, so ist $\langle d_i^{Pi} \rangle_{i \prec j}^{Pj}$ eine Herleitung der Tiefe $\leq \alpha$.

Den Schnittrang dieser Herleitung definieren wir durch $\mathsf{cr}(\langle d_i \rangle_{i \prec j}) := \sup_{i \prec j} \mathsf{cr}(d_i)$.

Da die Progressionsregel nur Herleitungen atomarer Formeln betrifft, verändert sie nicht die Schnittränge von Herleitungen. Folglich überträgt sich der Normalisierungsbeweis für Z^∞ unverändert auf $\mathsf{Z}^\infty + \mathsf{Prog}(P)$. Insbesondere haben wir

Lemma 6.4.1. *Für jede Herleitung d^A in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ einer Tiefe $\leq \alpha$ und vom Schnittrang $\leq k+1$ findet man eine Herleitung $(d^k)^A$ in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ mit freien Annahmevariablen unter denen von d, einer Tiefe $\leq 2^\alpha$ und vom Schnittrang $\leq k$.* □

Wir zeigen jetzt, daß man aus der Progressionsregel für P leicht die Progressivität von P herleiten kann.

Lemma 6.4.2. *Es gibt eine normale Herleitung von $\forall x.\forall y{\prec}x\, Py \to Px$ in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ der Tiefe 5.*

Beweis.

$$\dfrac{\dots \qquad \dfrac{\dfrac{\dots \qquad \dfrac{\dfrac{\forall y{\prec}j\, Py}{i \prec j \to Pi}\,\forall^- \qquad i \prec j}{Pi}\,\to^- \qquad \dots \quad (\text{alle } i \prec j)}{Pj}\,\mathsf{Prog}}{\forall y{\prec}j\, Py \to Pj}\,\to^+ \qquad \dots \quad (\text{alle } j)}{\forall x.\forall y{\prec}x\, Py \to Px}\,\omega$$

□

Die wesentliche Beobachtung ist jetzt, daß eine normale Herleitung von $P\ulcorner\beta\urcorner$ im wesentlichen eine Tiefe mindestens β erfordert. Um jedoch eine scharfe Abschätzung für die Teilsysteme Z_k zu erhalten, können wir Lemma 6.4.1 nicht bis herunter zum Schnittrang 0 (also bis zur Normalform) anwenden, sondern wir müssen bereits beim Schnittrang 1 aufhören. Derartige Herleitungen, also solche vom Schnittrang ≤ 1, nennen wir *quasinormal*; sie lassen sich ebenfalls leicht analysieren.

Wir beginnen mit einem Beweis, daß jede quasinormale Herleitung einer quantorenfreien Formel sich stets ohne Erhöhung ihres Schnittrangs oder ihrer Tiefe in eine quasinormale Herleitung derselben Formel umformen läßt, die

1. die ω-Regel nicht benutzt, und
2. die Regel $\forall^-$ höchstens in Anfangsteilen von mit einem Axiom beginnenden Ästen enthält.

Man beachte hierbei, daß alle unsere Axiome von der Form $\forall \boldsymbol{x} A$ sind mit A quantorenfrei.

Ferner benötigen wir den Begriff einer *Quasiteilformel*, der induktiv durch die folgenden Klauseln definiert ist.

- A, B sind Quasiteilformeln von $A \to B$;
- $A(i)$ ist eine Quasiteilformel von $\forall x A(x)$, für jede Zahl i;
- Ist A eine Quasiteilformel von B und C eine atomare Formel, so sind $C \to A$ und $\forall x A$ Quasiteilformeln von B;
- Die Relation "… ist eine Quasiteilformel von …" ist reflexiv und transitiv.

Zum Beispiel ist $Q \to \forall x.Px \to A$ eine Quasiteilformel von $A \to B$.

Wir übertragen jetzt die Teilformeleigenschaft für normale Herleitungen und beweisen entsprechend eine Quasiteilformeleigenschaft für quasinormale Herleitungen.

Satz 6.4.3. *(Quasiteilformeleigenschaft). Ist $d^B[u_1^{A_1}, \dots, u_n^{A_n}]$ eine quasinormale Herleitung in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ und ist e^C eine Teilherleitung von d^B, so ist C eine Quasiteilformel von B oder von einem A_i.*

Beweis. Induktion über die Länge der Endposition eines Astes in d. □

Korollar 6.4.4. *Sei d eine quasinormale Herleitung in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ einer Formel $\forall \boldsymbol{x} A$ aus quantorenfreien Annahmen, mit A quantorenfrei. Dann endet in d jeder Ast einer Ordnung > 0 mit einer quantorenfreien Formel.*

Beweis. Andernfalls würde die Hauptprämisse der $\to^-$-Regel, deren Nebenprämisse die Endformel dieses Astes ist, einen Quantor auf der linken Seite von $\to$ enthalten. Dies widerspricht dem Satz. □

Unser nächstes Ziel ist es, die ω-Regel zu eliminieren. Hierfür benötigen wir den Begriff einer *Instanz* einer Formel. Er ist induktiv wie folgt definiert.

- Ist B' Instanz von B und A quantorenfrei, so ist $A \to B'$ Instanz von $A \to B$.
- $A(i)$ ist Instanz von $\forall x A(x)$, für jede Zahl i.
- Die Relation "… ist eine Instanz von …" ist reflexiv und transitiv.

Lemma 6.4.5. *Sei d eine quasinormale Herleitung in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ einer Formel A aus quantorenfreien Annahmen, wobei A kein $\forall$ auf der linken Seite eines $\to$ enthält. Dann findet man für jede quantorenfreie Instanz A' von A eine quasinormale Herleitung d' von A' aus denselben Annahmen und quantorenfreien Instanzen der Axiome so daß*

- *d' die ω-Regel nicht verwendet,*
- *d' die Regel $\forall^-$ höchstens in Anfangsteilen von mit einem Axiom beginnenden Ästen enthält, und*
- *$|d'| \leq |d|$.*

Beweis. Induktion über die Tiefe d. Wir unterscheiden Fälle entsprechend der letzten Regel in d.

Fall $\to^-$.

$$\frac{A \to B \qquad A}{B} \to^-$$

Nach Korollar 6.4.4 muß A quantorenfrei sein. Sei B' eine quantorenfreie Instanz von B. Dann ist $A \to B'$ nach Definition eine quantorenfreie Instanz von $A \to B$. Die Behauptung folgt jetzt aus der IH.

Fall $\to^+$.

$$\frac{B}{A \to B} \to^+$$

Jede Instanz von $A \to B$ hat die Form $A \to B'$ mit einer quantorenfreien Instanz B' von B. Also folgt die Behauptung aus der IH.

Fall $\forall^-$.

$$\frac{\forall x A(x) \qquad i}{A(i)} \forall^-$$

Wieder ist jede quantorenfreie Instanz von $A(i)$ auch eine quantorenfreie Instanz von $\forall x A(x)$, und die Behauptung folgt aus der IH.

Fall ω.

$$\frac{\dots \quad A(i) \quad \dots \quad (\text{für alle } i < \omega)}{\forall x A(x)} \omega$$

Jede quantorenfreie Instanz von $\forall x A(x)$ hat die Form $A(i)'$ mit $A(i)'$ quantorenfreie Instanz von $A(i)$. Also folgt die Behauptung wieder aus der IH. $\square$

Eine Herleitung d in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ heißt *$P\boldsymbol{\alpha}, \neg P\boldsymbol{\beta}$-Widerlegung*, wenn $\boldsymbol{\alpha}$ und $\boldsymbol{\beta}$ disjunkt sind und d eine Formel $\boldsymbol{A} \to B := A_1 \to \dots \to A_k \to B$ herleitet, wobei $\boldsymbol{A}$ und die freien Annahmen in d wahre quantorenfreie Formeln ohne P sind oder unter $P\ulcorner\alpha_1\urcorner, \dots, P\ulcorner\alpha_m\urcorner, \neg P\ulcorner\beta_1\urcorner, \dots, \neg P\ulcorner\beta_n\urcorner$ vorkommen, und B eine falsche quantorenfreie Formel ohne P ist oder unter $P\ulcorner\beta_1\urcorner, \dots, P\ulcorner\beta_n\urcorner$ vorkommt.

Lemma 6.4.6. *Sei d eine quasinormale $P\boldsymbol{\alpha}, \neg P\boldsymbol{\beta}$-Widerlegung. Dann ist*

$$\min(\boldsymbol{\beta}) \leq |d| + \mathsf{lh}(\boldsymbol{\alpha}),$$

wobei $\mathsf{lh}(\boldsymbol{\alpha})$ die Länge der Liste $\boldsymbol{\alpha}$ bezeichnet.

Beweis. Induktion über $|d|$. Nach Lemma 6.4.5 können wir annehmen, daß d die ω-Regel nicht enthält, und die Regel $\forall^-$ höchstens in Anfangsteilen von mit einem Axiom beginnenden Ästen enthält. Wir unterscheiden Fälle entsprechend der letzten Regel in d.

Fall $\to^+$. Nach unserer Definition von Widerlegungen folgt die Behauptung unmittelbar aus der IH.

Fall $\to^-$. Dann ist $d = f^{C \to \boldsymbol{A} \to B} e^C$. Ist C eine wahre quantorenfreie Formeln ohne P oder von der Form $P\ulcorner\gamma\urcorner$ mit $\gamma < \min(\boldsymbol{\beta})$, so folgt die Behauptung aud der IH für f:

$$\min(\boldsymbol{\beta}) \leq |f| + \mathsf{lh}(\boldsymbol{\alpha}') + 1 \leq |d| + \mathsf{lh}(\boldsymbol{\alpha}').$$

Ist C eine falsche quantorenfreie Formel ohne P oder von der Form $P\ulcorner\gamma\urcorner$ mit $\min(\boldsymbol{\beta}) \leq \gamma$, so folgt die Behauptung aus der IH für e:

$$\min(\boldsymbol{\beta}) \leq |e| + \mathsf{lh}(\boldsymbol{\alpha}') + 1 \leq |d| + \mathsf{lh}(\boldsymbol{\alpha}').$$

Zu behandeln bleibt der Fall, daß C eine quantorenfreie Implikation ist, die P enthält. Dann ist $\mathsf{lev}(C) \geq 1$, also $\mathsf{lev}(C \to \boldsymbol{A} \to B) \geq 2$. Wegen $\mathsf{cr}(d) \leq 1$ muß dann d eine mit einem Axiom beginnende einfache Anwendung sein. Unsere einzigen Axiome, die P enthalten können, sind $\mathsf{Eq}_P \colon \forall x, y.x = y \to Px \to Py$ und $\mathsf{Stab}_P \colon \forall x.\neg\neg Px \to Px$, und von diesen hat nur Stab_P die richtige Form. Also $f = \mathsf{Stab}_P\ulcorner\gamma\urcorner$ und deshalb $e\colon \neg\neg P\ulcorner\gamma\urcorner$. Aus $\mathsf{lev}(\neg\neg P\ulcorner\gamma\urcorner) = 2$, der Annahme $\mathsf{cr}(e) \leq 1$ und wieder der Form unserer P enthaltenden Axiome folgt, daß e mit $\to^+$ endet, also $e = \lambda u^{\neg P\ulcorner\gamma\urcorner} e_0^{\perp}$ und damit insgesamt

$$\dfrac{\dfrac{\mid f}{\neg\neg P\ulcorner\gamma\urcorner \to P\ulcorner\gamma\urcorner} \qquad \dfrac{\dfrac{\begin{matrix}[u\colon \neg P\ulcorner\gamma\urcorner] \\ \mid e_0\end{matrix}}{\bot}}{\neg\neg P\ulcorner\gamma\urcorner}}{P\ulcorner\gamma\urcorner}$$

Die Behauptung folgt jetzt aus der IH für e_0.

Fall $\forall^-$. Nach Annahme befinden wir uns dann im Anfangsteil eines mit einem Axiom beginnenden Astes. Da d eine $P\boldsymbol{\alpha}, \neg P\boldsymbol{\beta}$-Widerlegung ist, muß das Axiom P enthalten. Das Gleichheitsaxiom $\mathsf{Eq}_P\colon \forall x,y.x = y \to Px \to Py$ kann es nicht sein, da $\ulcorner\gamma\urcorner = \ulcorner\delta\urcorner \to P\ulcorner\gamma\urcorner \to P\ulcorner\delta\urcorner$ in keinem Fall ($\gamma = \delta$ oder $\gamma \neq \delta$) Endformel einer $P\boldsymbol{\alpha}, \neg P\boldsymbol{\beta}$-Widerlegung ist. Aus demselben Grund kommt auch das Stabilitätsaxiom $\mathsf{Stab}_P\colon \forall x.\neg\neg Px \to Px$ nicht in Frage. Der Fall $\forall^-$ kann also nicht eintreten.

Fall $\mathsf{Prog}(P)$. Dann ist $d = \langle d_\delta^{P\ulcorner\delta\urcorner}\rangle_{\delta<\gamma}^{P\ulcorner\gamma\urcorner}$. Nach Annahme über d ist γ in $\boldsymbol{\beta}$. Ohne Beschränkung der Allgemeinheit sei $\gamma = \beta_i := \min(\boldsymbol{\beta})$, denn andernfalls wäre die Prämissenherleitung $d_{\beta_i}\colon P\ulcorner\beta_i\urcorner$ eine quasinormale $P\boldsymbol{\alpha}, \neg P\boldsymbol{\beta}$-Widerlegung, auf die man die IH anwenden könnte.

Gibt es keine $\alpha_j < \gamma$, so ist das Argument einfach: jedes d_δ ist eine $P\boldsymbol{\alpha}, \neg P\boldsymbol{\beta}, \neg P\delta$-Widerlegung, also gilt nach IH (da auch kein $\alpha_j < \delta$ ist)

$$\min(\boldsymbol{\beta}, \delta) = \delta \leq |d_\delta|,$$

also $\gamma = \min(\boldsymbol{\beta}) \leq |d|$.

Wir befassen uns jetzt mit dem Fall, daß einige α_j kleiner als γ sind. Man beachte zunächst, daß es höchstens endlich viele α_j geben kann, die unmittelbar vor γ kommen; sei also ε die kleinste Ordinalzahl mit

$$\forall\delta.\varepsilon \leq \delta < \gamma \to \delta \in \boldsymbol{\alpha}.$$

Dann ist $\varepsilon, \varepsilon+1, \dots, \varepsilon+k-1 \in \boldsymbol{\alpha}$ und $\varepsilon + k = \gamma$. ε ist entweder ein Nachfolger oder eine Limeszahl. Ist $\varepsilon = \varepsilon' + 1$, so folgt aus der IH (da $d_{\varepsilon'}$ eine $P\boldsymbol{\alpha}, \neg P\boldsymbol{\beta}, \neg P(\varepsilon-1)$-Widerlegung ist), daß

$$\varepsilon - 1 \leq |d_{\varepsilon-1}| + \mathsf{lh}(\boldsymbol{\alpha}') - k,$$

wobei $\boldsymbol{\alpha}'$ eine Folge von $\alpha_j < \gamma$ ist. Also ist $\varepsilon \leq |d| + \mathsf{lh}(\boldsymbol{\alpha}') - k$, und deshalb

$$\gamma \leq |d| = \mathsf{lh}(\boldsymbol{\alpha}').$$

Ist ε eine Limeszahl, so gibt es eine Folge $\langle\delta_{f(n)}\rangle_n$ mit Limes ε, und so daß alle $\alpha_j < \varepsilon$ kleiner als $\delta_{f(0)}$ sind. Nach IH folgt

$$\delta_{f(n)} \leq |d_{f(n)}| + \mathsf{lh}(\boldsymbol{\alpha}') - k,$$

und deshalb $\varepsilon \leq |d_{f(n)}| + \mathsf{lh}(\boldsymbol{\alpha}') - k$, also $\gamma \leq |d| + \mathsf{lh}(\boldsymbol{\alpha}')$. □

Jetzt können wir das folgende Resultat beweisen (s. Mints [21] und Parsons [24]).

Satz 6.4.7. *Das Axiom der transfiniten Induktion bis ε_0 ist unbeweisbar in Z, also*

$$\mathsf{Z} \nvdash (\forall x.\forall y{\prec}x\, Py \to Px) \to \forall x\, Px$$

mit einem Prädikatensymbol P, und die transfinite Induktion bis ω_{k+1} ist unbeweisbar in Z_k, also

$$\mathsf{Z}_k \nvdash (\forall x.\forall y{\prec}x\, Py \to Px) \to \forall x{\prec}\omega_{k+1}\, Px.$$

Beweis. Wir beschränken uns auf den zweiten Teil. Nehmen wir also an, die transfinite Induktion bis ω_{k+1} ist herleitbar in Z_k. Aufgrund der Einbettung von Z_k in Z^∞ (Lemma 6.3.3) und der normalen Herleitbarkeit der Progressivität von P in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ mit endlicher Tiefe (Lemma 6.4.2) können wir schließen, daß $\forall x{\prec}\omega_{k+1}\, Px$ herleitbar ist in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ mit einer Tiefe $< \omega^{m+1}$ und Schnittrang $\leq k$. Nun liefern $k-1$ Anwendungen von Lemma 6.4.1 eine Herleitung derselben Formel $\forall x{\prec}\omega_{k+1}\, Px$ in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ mit einer Tiefe $\gamma < 2_{k-1}^{\omega^{m+1}} < \omega_{k+1}$ und Schnittrang ≤ 1, also auch eine Herleitung von $P\ulcorner\gamma + 2\urcorner$ in $\mathsf{Z}^\infty + \mathsf{Prog}(P)$ mit der Tiefe $\gamma + 1$ und Schnittrang ≤ 1. Dies widerspricht Lemma 6.4.6. □

Wir wollen uns schließlich noch überlegen, daß man als Folgerung hieraus die Unmöglichkeit der Normalisierung für die Arithmetik erhalten kann. Der Normalisierungssatz für die Logik erster Stufe ist nicht besonders nützlich, wenn man ihn auf eines unserer arithmetischen Systeme Z anwendet. Der Grund liegt darin, daß in einer Herleitung Induktionsaxiome beliebiger Komplexität vorkommen können. Es liegt deshalb nahe zu versuchen, die Induktionsaxiome durch eine Induktionsregel zu ersetzen, welche den Schluß auf $\forall x A(x)$ aus einer Herleitung von $A(0)$ und einer Herleitung von $A(\mathsf{S}(x))$ mit der zusätzlichen Annahme $A(x)$ gestattet, die an dieser Stelle zu streichen ist; man beachte, daß diese Induktionsregel äquivalent zum Induktionsschema ist. Dann kann man versuchen, die entstehende Herleitung in dem neuen System Z mit der Induktionsregel zu normalisieren. Wir zeigen jetzt, daß sogar eine sehr schwache Form eines solchen Normalisierungssatzes in Z mit der Induktionsregel nicht gelten kann.

Satz 6.4.8. *Die folgende Form eines Normalisierungssatzes in Z mit der Induktionsregel ist falsch. Für jede Herleitung $d[\boldsymbol{u}^{\boldsymbol{A}}]\colon B$ mit $\boldsymbol{A}, B$ Formeln der Stufe $\leq \ell$ gibt es eine Herleitung $d^*[\boldsymbol{u}^{\boldsymbol{A}}]\colon B$, die nur Formeln einer Stufe $\leq k$ enthält, wobei k nur von ℓ abhängt.*

Beweis. Nehmen wir an, ein solcher Normalisierungssatz würde gelten. Man betrachte die Formel

$$(\forall x.\forall y{\prec}x\, Py \to Px) \to \forall x{\prec}\omega_{n+1}\, Px,$$

die die transfinite Induktion bis ω_{n+1} ausdrückt und von der Stufe 3 ist. Nach Satz 6.2.1 ist sie herleitbar in Z. Aus unserer Annahme folgt nun, daß es eine Herleitung dieser Formel gibt, die nur Formeln einer Stufe $\leq k$ enthält, für ein von n unabhängiges k. Also beweist Z_k die transfinite Induktion bis ω_{n+1} für jedes n. Dies widerspricht aber Satz 6.4.7. □

6.5 Nachtrag: η-Expansion

Für manche Untersuchungen ist es zweckmäßig, anstelle der η-Konversion ihre Umkehrung zu verwenden, die sogenannte *η-Expansion*. Ein Problem besteht jedoch darin, daß die η-Expansion zusammen mit der β-Konversion zu Schleifen führen kann:

$$MN \to_{\eta\uparrow} (\lambda x.Mx)N \to_\beta MN \quad \text{und} \quad \lambda xM \to_{\eta\uparrow} \lambda x(\lambda xM)x \to_\beta \lambda xM.$$

Wir wollen deshalb die η-Expansion von Termen in Anwendungspositionen und auch von Abstraktionen ausschließen. Dies läßt sich wie folgt erreichen.

Definition 6.5.1. (η-Expansion). $\to_{\eta\uparrow}$ sei der Termabschluß der Konversionsregel

$$M^{\rho\to\sigma} \mapsto_{\eta\uparrow} \lambda x^\rho.Mx \quad \text{falls } x \notin \mathsf{FV}(M) \text{ und } M \text{ neutral ist.}$$

Ferner sei $\to_{\beta\eta\uparrow} := \to_\beta \cup \to_{\eta\uparrow}$. Für $\gamma \in \{\eta\uparrow, \beta\eta\uparrow\}$ können wir wie in Definition 2.1 die Begriffe *γ-normal* (oder in *γ-Normalform*), *γ-Reduktionsfolge* und *stark γ-normalisierbar* definieren. Man beachte, daß man von γ-Redexen nicht sinnvoll sprechen kann.

Offenbar lassen sich die Terme in $\eta\uparrow$-Normalform charakterisieren durch

$$M ::= (x\boldsymbol{M})^\iota \mid \lambda xM \mid ((\lambda xM)N\boldsymbol{L})^\iota;$$

die Terme in $\beta\eta\uparrow$-Normalform erhält man durch Weglassen der letzten Regel.

In diesem Abschnitt zeigen wir die Terminierung und Konfluenz von $\to_{\beta\eta\uparrow}$. Eine Schwierigkeit liegt in der Möglichkeit der Interaktion von η-Expansion mit β-Konversion. Dies führt dazu, daß die Substitutionseigenschaften aus Lemma 2.2.1(1), (3) und (4) alle für $\to \beta\eta\uparrow$ nicht mehr gelten.

Definition 6.5.2. (Äußere η-Expansion). Wir definieren $\eta_\rho(M^\rho) \in \Lambda_\rho$ zusammen mit seiner *Expansionshöhe* $\mu_\rho \in \mathbb{N}$ durch Rekursion über ρ.

$$\begin{array}{lcllcl}
\eta_\iota(M) & := & M, & \mu_\iota & := & 0,\\
\eta_{\rho\to\sigma}(M) & := & \lambda x^\rho \eta_\sigma(M\eta_\rho(x)). & \mu_{\rho\to\sigma} & := & \mu_\rho + \mu_\sigma + 1.
\end{array}$$

Beispiele: $\eta_{\iota\to\iota}(y) = \lambda x.yx$ und für $\rho = (\iota \to \iota) \to (\iota \to \iota)$ ist $\eta_\rho(z) = \lambda y\lambda x.z(\lambda u.yu)x$.

Lemma 6.5.3. *(Eigenschaften der äußeren η-Expansion).*

$$M^\rho \to_{\eta\uparrow}^{\mu_\rho} \eta(M) \quad \textit{falls } M \textit{ neutral ist.} \tag{6.18}$$

Sind $\boldsymbol{M}, M, N, \boldsymbol{L}$ in $\eta\uparrow$-Normalform, so auch $\eta_\rho(x\boldsymbol{M})$ und $\eta_\rho((\lambda xM)N\boldsymbol{L})$. (6.19)

Wenn $M \to_\beta M'$, so ist $\eta(M) \to_\beta \eta(M')$. (6.20)

$$\eta(M)[x := N] = \eta(M[x := N]). \tag{6.21}$$

$$\eta(M)\boldsymbol{N} \xrightarrow{*}_\beta \eta(M\eta(\boldsymbol{N})). \tag{6.22}$$

$$\eta(\eta(M)) \xrightarrow{*}_\beta \eta(M). \tag{6.23}$$

Beweis. (6.18). Induktion über den Typ ρ von M. Der Anfangsfall $\rho = \iota$ ist klar, und im Fall $\rho \to \sigma$ haben wir

$$\begin{array}{lll} M^{\rho\to\sigma} & \to_{\eta\uparrow} & \lambda x.Mx \quad \text{falls } M \text{ neutral} \\ & \to_{\eta\uparrow}^{\mu_\rho} & \lambda x.M\eta_\rho(x) \quad \text{nach IH}(\rho) \\ & \to_{\eta\uparrow}^{\mu_\sigma} & \lambda x.\eta_\sigma(M\eta_\rho(x)) \quad \text{nach IH}(\sigma). \end{array}$$

(6.19). Induktion über ρ. Der Fall ι ist klar, und im Fall $\rho \to \sigma$ haben wir

$$\begin{array}{l} \eta_{\rho\to\sigma}(x\boldsymbol{M}) = \lambda y^\rho \eta_\sigma(x\boldsymbol{M}\eta_\rho(y)), \\ \eta_{\rho\to\sigma}((\lambda xM)N\boldsymbol{L}) = \lambda y^\rho \eta_\sigma((\lambda xM)N\boldsymbol{L}\eta_\rho(y)). \end{array}$$

Die Behauptung folgt in beiden Fällen aus den IHn für ρ und σ.

(6.20). Induktion über den Typ von M.

(6.21). Induktion über den Typ von M. Sei oBdA M vom Typ $\rho \to \sigma$. Dann gilt

$$\begin{array}{lll} \eta(M)[x := N] & = & \lambda y\eta(M\eta(y))[x := N] \\ & = & \lambda y\eta(M[x := N]\eta(y)) \quad \text{nach IH} \\ & = & \eta(M[x := N]). \end{array}$$

(6.22). Induktion über den Typ von M. Der Fall ι ist klar. Im Fall $\rho \to \sigma$ können wir oBdA annehmen, daß $\boldsymbol{N}$ nicht leer ist. Man erhält

$$\begin{array}{lll} \eta_{\rho\to\sigma}(M)N\boldsymbol{L} & = & \lambda x\eta(M\eta(x))N\boldsymbol{L} \\ & \to_\beta & \eta(M\eta(N))\boldsymbol{L} \\ & \xrightarrow{*}_\beta & \eta(M\eta(N)\eta(\boldsymbol{L})) \quad \text{nach IH.} \end{array}$$

(6.23). Induktion über den Typ von M. Der Fall ι ist klar. Im Fall $\rho \to \sigma$ haben wir

$$\begin{array}{lll} \eta(\eta(M)) & = & \lambda x\eta(\eta(M)\eta(x)) \\ & \xrightarrow{*}_\beta & \lambda x\eta(\eta(M\eta(\eta(x)))) \quad \text{nach (6.22)} \\ & \xrightarrow{*}_\beta & \lambda x\eta(M\eta(x)) \quad \text{nach IH} \\ & = & \eta(M). \end{array}$$

□

Definition 6.5.4. (η-Expansion). Wir definieren $\exp(M^\rho) \in \Lambda_\rho$ zusammen mit seiner *Expansionshöhe* $\#_\eta(M^\rho) \in \mathbb{N}$ durch Rekursion über ρ.

$$\begin{array}{lll@{\qquad}lll} \exp(x\boldsymbol{M}) & := & \eta(x\exp(\boldsymbol{M})), & \#_\eta(x\boldsymbol{M}) & := & \mu_\rho + \#_\eta(\boldsymbol{M}), \\ \exp(\lambda xM) & := & \lambda x\exp(M), & \#_\eta(\lambda xM) & := & \#_\eta(M), \\ \exp((\lambda xM)N\boldsymbol{L}) & := & \eta((\lambda x\exp(M))\exp(N)\exp(\boldsymbol{L})). & \#_\eta((\lambda xM)N\boldsymbol{L}) & := & \mu_\rho + \#_\eta(M, N, \boldsymbol{L}). \end{array}$$

Hierbei steht $\#_\eta(\boldsymbol{M})$ für $\sum_i \#_\eta(M_i)$.

Lemma 6.5.5. *(Eigenschaften der η-Expansion).*

$$M \to_{\eta\uparrow}^{\#_\eta(M)} \exp(M), \quad \textit{und } \exp(M) \textit{ ist in } \eta\uparrow\textit{-Normalform.} \tag{6.24}$$

$$\eta(\exp(M)) \xrightarrow{*}_\beta \exp(M). \tag{6.25}$$

$$\exp(M)[x := \eta(x)] \xrightarrow{*}_\beta \exp(M). \tag{6.26}$$

$$\eta(\exp(M)\exp(\boldsymbol{N})) \xrightarrow{*}_\beta \exp(M\boldsymbol{N}). \tag{6.27}$$

$$\exp(M)[x := \exp(N)] \xrightarrow{*}_\beta \exp(M[x := N]). \tag{6.28}$$

$$\textit{Wenn } M \to_{\eta\uparrow} M', \textit{ so ist } \exp(M) = \exp(M') \textit{ und } \#_\eta(M) = \#_\eta(M') + 1. \tag{6.29}$$

Beweis. (6.24) beweist man leicht durch Induktion über M, unter Verwendung von (6.18) bzw. (6.19).

(6.25) und (6.26) werden simultan durch Induktion über ρ bewiesen. (6.25). Für neutrales M ist $\exp(M)$ ein η-Bild und die Behauptung folgt aus (6.23). Im Fall einer Abstraktion haben wir

$$\begin{aligned}
\eta(\exp(\lambda x M)) &= \eta(\lambda x \exp(M)) \\
&= \lambda x \eta((\lambda x \exp(M))\eta(x)) \\
&\xrightarrow{*}_\beta \lambda x \eta(\exp(M)[x := \eta(x)]) \quad \text{nach IH(6.26) mit (6.20)} \\
&\xrightarrow{*}_\beta \lambda x \exp(M) \quad \text{nach IH(6.25)} \\
&= \exp(\lambda x M).
\end{aligned}$$

(6.26). *Fall $x\boldsymbol{M}$.*

$$\begin{aligned}
\exp(x\boldsymbol{M})[x := \eta(x)] &= \eta(x \exp(\boldsymbol{M})[x := \eta(x)] \\
&= \eta(\eta(x) \exp(\boldsymbol{M})[x := \eta(x)]) \quad \text{nach (6.21)} \\
&\xrightarrow{*}_\beta \eta(\eta(x) \exp(\boldsymbol{M})) \quad \text{nach IH(6.26) mit (6.20)} \\
&\xrightarrow{*}_\beta \eta(\eta(x\eta(\exp(\boldsymbol{M})))) \quad \text{nach (6.22) mit (6.20)} \\
&\xrightarrow{*}_\beta \eta(x\eta(\exp(\boldsymbol{M}))) \quad \text{nach (6.23)} \\
&\xrightarrow{*}_\beta \eta(x \exp(\boldsymbol{M})) \quad \text{nach (6.25) mit (6.20)} \\
&= \exp(x\boldsymbol{M}).
\end{aligned}$$

Die restlichen Fälle erhält man leicht aus der IH.

(6.27). Induktion über M. Wir können oBdA annehmen, daß $\boldsymbol{N}$ nicht leer ist.

$$\begin{aligned}
\eta(\exp(x\boldsymbol{M})\exp(\boldsymbol{N})) &= \eta(\eta(x \exp(\boldsymbol{M}))\exp(\boldsymbol{N})) \\
&\xrightarrow{*}_\beta \eta(\eta(x \exp(\boldsymbol{M})\eta(\exp(\boldsymbol{N})))) \quad \text{nach (6.22) mit (6.20)} \\
&\xrightarrow{*}_\beta \eta(x \exp(\boldsymbol{M})\eta(\exp(\boldsymbol{N}))) \quad \text{nach (6.23)} \\
&\xrightarrow{*}_\beta \eta(x \exp(\boldsymbol{M})\exp(\boldsymbol{N})) \quad \text{nach (6.25)} \\
&= \exp(M\boldsymbol{M}\boldsymbol{N}) \\
\eta(\exp(\lambda x M)\exp(N)\exp(\boldsymbol{L})) &= \eta((\lambda x \exp(M))\exp(N)\exp(\boldsymbol{L})) \\
&= \exp((\lambda x M)N\boldsymbol{L}) \\
\eta(\exp((\lambda x M)N\boldsymbol{L})\exp(\boldsymbol{K})) &= \eta(\eta((\lambda x \exp(M))\exp(N)\exp(\boldsymbol{L}))\exp(\boldsymbol{K})) \\
&\xrightarrow{*}_\beta \eta(\eta((\lambda x \exp(M))\exp(N)\exp(\boldsymbol{L}))\eta(\exp(\boldsymbol{K}))) \quad \text{nach (6.22), (6.20)} \\
&\xrightarrow{*}_\beta \eta((\lambda x \exp(M))\exp(N)\exp(\boldsymbol{L})\exp(\boldsymbol{K})) \quad \text{nach (6.23), (6.25), (6.20)} \\
&= \exp((\lambda x M)N\boldsymbol{L}).
\end{aligned}$$

(6.28). Induktion über M.

$$\begin{aligned}
\exp(x\boldsymbol{M})[x := \exp(N)] &= \eta(x \exp(\boldsymbol{M}))[x := \exp(N)] \\
&= \eta(\exp(N)\boldsymbol{M}[x := \exp(N)]) \quad \text{nach (6.21)} \\
&\xrightarrow{*}_\beta \eta(\exp(N) \exp(\boldsymbol{M}[x := \exp(N)])) \quad \text{nach IH mit (6.20)} \\
&\xrightarrow{*}_\beta \exp(N\boldsymbol{M}[x := N]) \quad \text{nach (6.27).}
\end{aligned}$$

Die restlichen Fälle ergeben sich leicht aus der IH.

(6.29). Induktion über $M \to_{\eta\uparrow} M'$. Im Fall einer $\eta\uparrow$-Kopfkonversion betrachten wir zunächst den Unterfall $x\boldsymbol{M} \mapsto_{\eta\uparrow} \lambda y.x\boldsymbol{M}y$. Man erhält

$$\begin{aligned}
\exp(x\boldsymbol{M}) &= \eta(x \exp(\boldsymbol{M})) \\
&= \lambda y \eta(x \exp(\boldsymbol{M})\eta(y)) \\
&= \exp(\lambda y.x\boldsymbol{M}y)
\end{aligned}$$

und

$$\begin{aligned}
\#_\eta(x\boldsymbol{M}^{\rho\to\sigma}) &= \mu_{\rho\to\sigma} + \#_\eta(\boldsymbol{M}) \\
&= \mu_\sigma + \#_\eta(\boldsymbol{M}) + \mu_\rho + 1 \\
&= \mu_\sigma + \#_\eta(\boldsymbol{M}, y^\rho) + 1 \\
&= \#_\eta(\lambda y.x\boldsymbol{M}y) + 1.
\end{aligned}$$

Im Fall einer inneren Konversion folgt die Behauptung wieder sofort aus der IH. □

Proposition 6.5.6. *Wenn* $M \to_\beta M'$*, so ist* $\exp(M) \to_\beta^+ \exp(M')$.

Beweis. Induktion über $M \to_\beta M'$. Im Fall einer β-Kopfkonversion erhalten wir

$$\begin{aligned}
\exp((\lambda x M)N\boldsymbol{L}) &= \eta((\lambda x \exp(M)) \exp(N) \exp(\boldsymbol{L})) \\
&\to_\beta \eta(\exp(M)[x := \exp(N)] \exp(\boldsymbol{L})) \quad \text{nach (6.20)} \\
&\xrightarrow{*}_\beta \eta(\exp(M[x := N]) \exp(\boldsymbol{L})) \quad \text{nach (6.28) mit (6.20)} \\
&\xrightarrow{*}_\beta \exp(M[x := N]\boldsymbol{L}) \quad \text{nach (6.27).}
\end{aligned}$$

Die Fälle $x\boldsymbol{M}M\boldsymbol{N} \to x\boldsymbol{M}M'\boldsymbol{N}$ und $(\lambda x M)\boldsymbol{N} \to (\lambda x M')\boldsymbol{N}$ sind klar nach der IH. Auch die restlichen Fälle ergeben sich sofort aus der IH. □

Korollar 6.5.7. $\to_{\beta\eta\uparrow}$ *ist terminierend.*

Beweis. Nach der Proposition kann man jede β-Reduktion auf dem Term M durch eine positive Anzahl von β-Reduktionen auf $\exp(M)$ simulieren, während nach (6.29) η-Expansionen den Term $\exp(M)$ unverändert lassen. Die Terminierung von $\to_{\beta\eta\uparrow}$ folgt jetzt durch Induktion über $\exp(M)$ bzgl. der terminierenden Relation $\to_\beta$ und Nebeninduktion über $\#_\eta(M)$. □

Proposition 6.5.8. $\to_{\beta\eta\uparrow}$ *ist konfluent.*

Beweis. Gelte $N_0 \xleftarrow{*}_{\beta\eta\uparrow} M \xrightarrow{*}_{\beta\eta\uparrow} N_1$. Nach (6.29) und Proposition 6.5.6 folgt $\exp(N_0) \xleftarrow{*}_\beta \exp(M) \xrightarrow{*}_\beta \exp(N_1)$. Die Konfluenz von $\to_\beta$ liefert einen Term N mit $\exp(N_0) \xrightarrow{*}_\beta N \xleftarrow{*}_\beta \exp(N_1)$. Nach (6.24) gilt $N_i \to_{\eta\uparrow} \exp(N_i)$; damit folgt die Behauptung.

$$\begin{tikzcd}
 & M \arrow[dl, "\ast"', "\beta\eta\uparrow"] \arrow[dr, "\ast", "\beta\eta\uparrow"'] & \\
N_0 \arrow[dd, "\ast"', "\eta\uparrow"] & & N_1 \arrow[dd, "\ast", "\eta\uparrow"'] \\
 & \exp(M) \arrow[dl, "\ast"', "\beta"] \arrow[dr, "\ast", "\beta"'] & \\
\exp(N_0) \arrow[dr, "\ast", "\beta"'] & & \exp(N_1) \arrow[dl, "\ast"', "\beta"] \\
 & N &
\end{tikzcd}$$

□

Man kann leicht zeigen, daß sich die Terme in $\eta\uparrow$-Normalform charakterisieren lassen durch

$$M ::= (x\boldsymbol{M})^\iota \mid \lambda xM \mid ((\lambda xM)N\boldsymbol{L})^\iota,$$

und daß man bei Weglassen der letzten Regel die Terme in $\beta\eta\uparrow$-Normalform erhält.

6.6 Anmerkungen

Die Beweisbarkeit von Anfangsfällen der transfiniten Induktion (s. Abschnitt 6.2) wurde von Gentzen [10] gezeigt; in dieser Arbeit zeigte Gentzen auch die Unbeweisbarkeit der transfiniten Induktion bis ε_0 für ein Relationssymbol P (s. Abschnitt 6.4).

Unsere Darstellung basiert auf Gentzens [10], verwendet aber wie in Schütte [29] ein unendliches Beweissystem mit ω-Regel (s. Abschnitt 6.3), in das sich die übliche Arithmetik einbetten läßt. Dies ist hier für ein System des natürlichen Schließens durchgeführt; eine ähnliche Darstellung für ein Gentzen-Tait-System findet sich in [33].

Gentzens Resultat über die Unbeweisbarkeit der transfiniten Induktion bis ε_0 war das erste Beispiel eines mathematisch leicht zu verstehenden wahren Satzes, der nicht in der erststufigen Arithmetik beweisbar ist. Dies steht im Gegensatz zu Gödels Unvollständigkeitssatz: dort ist der unbeweisbare Satz ausschließlich durch metamathematische Betrachtungen motiviert. Einen mehr kombinatorische Aussage dieser Art, in der Ordinalzahlen nicht explizit vorkommen, wurde zuerst von Paris [23] gefunden.

Der einfache Beweis in Abschnitt 6.5 der Terminierung und Konfluenz von $\to_{\beta\eta\uparrow}$ stammt in dieser Form von Felix Joachimski; man kann ihn als eine verbesserte Variante der η-Expansor-Methode von di Cosmo und Kesner [7] ansehen.

Weiterführende Bücher über Beweistheorie sind Schütte [31, 32], Buchholz et al. [4], Takeuti [40], Girard [11], Pohlers [26] sowie [43].

Literatur

1. E.W. Beth. Semantic construction of intuitionistic logic. *Medelingen de KNAW N.S.*, 19(11), 1956.
2. E.W. Beth. *The foundations of mathematics.* North–Holland, Amsterdam, 1959.
3. Wilfried Buchholz. A new system of proof–theoretic ordinal functions. *Annals of Pure and Applied Logic*, 32(3):195–207, 1986.
4. Wilfried Buchholz, Solomon Feferman, Wolfram Pohlers, and Wilfried Sieg. *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof–Theoretical Studies*, volume 897 of *Lecture Notes in Mathematics.* Springer, Berlin, 1981.
5. Georg Cantor. Beträge zur Begründung der transfiniten Mengenlehre. *Mathematische Annalen*, 49, 1897.
6. C.C. Chang and H.J. Keisler. *Model Theory*, volume 73 of *Studies in Logic.* North–Holland, Amsterdam, 3rd edition, 1990.
7. Roberto Di Cosmo and Delia Kesner. A confluent reduction for the extensional typed λ–calculus with pairs, sums, recursion and terminal object. In *ICALP 93*, number 700 in Lecture Notes in Computer Science, pages 645–656. Springer Verlag, Berlin, Heidelberg, New York, 1993.
8. Harvey Friedman. Equality between functionals. In R. Parikh, editor, *Logic Colloquium, Lecture Notes in Mathematics 453*, pages 22–37. Springer, 1975.
9. Gerhard Gentzen. Untersuchungen über das logische Schließen. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1934.
10. Gerhard Gentzen. Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie. *Mathematische Annalen*, 119:140–161, 1943.
11. Jean-Yves Girard. *Proof Theory and Logical Complexity.* Bibliopolis, Napoli, 1987.
12. Kurt Gödel. Die Vollständigkeit der Axiome des logischen Funktionenkalküls. *Monatshefte für Mathematik und Physik*, 37:349–360, 1930.
13. Kurt Gödel. Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931.
14. Carl A. Gunter. *Semantics of Programming Languages: Structures and Techniques.* Foundations of Computing. The MIT Press, Cambridge, Massachusetts, 1992.
15. Ingebrigt Johansson. Der Minimalkalkül, ein reduzierter intuitionistischer Formalismus. *Compositio Mathematica*, 4:119–136, 1937.
16. Stephen C. Kleene. *Introduction to Metamathematics.* D. van Nostrand Comp., New York, 1952.
17. Saul A. Kripke. Semantical analysis of intuitionistic logic I. In J. Crossley and M. Dummett, editors, *Formal Systems and Recursive Functions*, pages 93–130. North–Holland, Amsterdam, 1965.
18. Jerzy Łoś. Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. In *Mathematical Interpretation of Formal Systems*, pages 98–113. North–Holland, Amsterdam, 1955.
19. L. Löwenheim. Über Möglichkeiten im Relativkalkül. *Mathematische Annalen*, 76:447–470, 1915.
20. A. Malzew. Untersuchungen aus dem Gebiete der mathematischen Logik. *Rec. Math. N. S.*, 1:323–336, 1936.
21. Grigori E. Mints. Exact estimates of the provability of transfinite induction in the initial segments of arithmetic. *Journal of Soviet Math*, 1:85–91, 1973. Translated from Zapiski Nauch. Sem. Leningrad *20*, 134–144 (1971).
22. V.P. Orevkov. Lower bounds for increasing complexity of derivations after cut elimination. *Zapiski Nauchnykh Seminarov Leningradskogo*, 88:137–161, 1979.
23. J.B. Paris. Some independence results for Peano arithmetic. *The Journal of Symbolic Logic*, 43(4):725–731, 1978.
24. Charles Parsons. Transfinite induction in subsystems of number theory (abstract). *The Journal of Symbolic Logic*, 38(3):544–545, 1973.
25. Rósza Péter. *Rekursive Funktionen.* Akademie–Verlag, Berlin, 1957.
26. Wolfram Pohlers. *Proof Theory*, volume 1407 of *Lecture Notes in Mathematics.* Springer Verlag, Berlin, Heidelberg, New York, 1989.
27. Harvey E. Rose. *Subrecursion: Functions and hierarchies*, volume 9 of *Oxford Logic Guides.* Clarendon Press, Oxford, 1984.
28. Rósza Péter. *Recursive Functions.* Academic Press, New York, 1967.
29. Kurt Schütte. Beweistheoretische Erfassung der unendlichen Induktion in der Zahlentheorie. *Mathematische Annalen*, 122:369–389, 1951.

30. Kurt Schütte. Kennzeichnung von Ordinalzahlen durch rekursiv definierte Funktionen. *Mathematische Annalen*, 127:16–32, 1954.
31. Kurt Schütte. *Beweistheorie.* Springer Verlag, Berlin, Heidelberg, New York, 1960.
32. Kurt Schütte. *Proof Theory.* Springer Verlag, Berlin, Heidelberg, New York, 1977.
33. Helmut Schwichtenberg. Proof theory: Some applications of cut-elimination. In J. Barwise, editor, *Handbook of Mathematical Logic*, volume 90 of *Studies in Logic and the Foundations of Mathematics*, chapter Proof Theory and Constructive Mathematics, pages 867–895. North–Holland, Amsterdam, 1977.
34. Joseph R. Shoenfield. *Mathematical Logic.* Addison–Wesley Publ. Comp., Reading, Massachusetts, 1967.
35. T. Skolem. Logisch–kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theorem über dichte Mengen. *Skrifter utgitt av Videnkapsselskapet i Kristiania, I, Mat. Naturv. Kl.*, 4:36 pp., 1920.
36. T. Skolem. über die Nicht–Charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschließlich Zahlenvariablen. *Fund. Math.*, 23:150–161, 1934.
37. Richard Statman. Bounds for proof-search and speed-up in the predicate calculus. *Annals of Mathematical Logic*, 15:225–287, 1978.
38. William W. Tait. Infinitely long terms of transfinite type I. In J. Crossley and M. Dummett, editors, *Formal Systems and Recursive Functions*, pages 176–185. North–Holland, Amsterdam, 1965.
39. William W. Tait. A realizability interpretation of the theory of species. In R. Parikh, editor, *Logic Colloquium Boston 1971/72*, volume 453 of *Lecture Notes in Mathematics*, pages 240–251. Springer Verlag, Berlin, Heidelberg, New York, 1975.
40. Gaisi Takeuti. *Proof Theory.* North–Holland, Amsterdam, second edition, 1987.
41. Anne S. Troelstra. *Choice Sequences.* Oxford Logic Guides. Clarendon Press, Oxford, 1977.
42. Anne S. Troelstra and Helmut Schwichtenberg. *Basic Proof Theory.* Cambridge University Press, 1996.
43. Anne S. Troelstra and Helmut Schwichtenberg. *Basic Proof Theory.* Cambridge University Press, 2nd edition, 2000.
44. Anne S. Troelstra and Dirk van Dalen. *Constructivism in Mathematics. An Introduction*, volume 121, 123 of *Studies in Logic and the Foundations of Mathematics.* North–Holland, Amsterdam, 1988.
45. Oswald Veblen. Continuous increasing functions of finite and transfinite ordinals. *Transactions AMS*, 9:280–292, 1908.

Index