



DIPLOMARBEIT

Titel der Diplomarbeit

“ZFC vs. NFU – Eine philosophische Untersuchung zur
Mengenlehre”

Verfasser

Günther Eder

angestrebter akademischer Grad

Magister der Philosophie (Mag. phil.)

Wien, im November 2008

Studienkennzahl lt. Studienblatt: A296

Studienrichtung lt. Studienblatt: Philosophie

Betreuerin: MMag. DDr. Esther Ramharter

Inhaltsverzeichnis

1	Vorwort	4
2	Das System NFU	8
2.1	Einfache Folgerungen	10
2.2	Relationen und Funktionen	10
2.3	Ordnungsrelationen	11
2.4	Äquivalenzrelationen	12
2.5	Funktionen	13
3	Das System ZFC	14
3.1	Einfache Folgerungen	15
3.2	Funktionen und Relationen	15
4	Ein erster Vergleich	17
4.1	Extensionalität	17
4.2	Geordnete Paare	19
4.3	Aussonderung vs. S-Komprehension	19
4.4	Stratifizierung und heterogene Relationen	22
4.5	Fundierungsaxiom contra $V \in V$	24
4.6	Mengen und Klassen	28
5	Kardinal- und Ordinalzahlen in NFU	30
5.1	Kardinalzahlen: Grundlegende Begriffe	30
5.2	Ordinalzahlen: Grundlegende Begriffe	35
5.3	Mehr ordinale Arithmetik	38
5.4	Mehr kardinale Arithmetik	41
6	Kardinal- und Ordinalzahlen in ZFC	46
6.1	Ordinalzahlen: Grundlegende Begriffe	46
6.2	Exkurs: Die natürlichen Zahlen	50
6.3	Kardinalzahlen: Grundlegende Begriffe	52
7	Mengenlehre und Arithmetik: Ein Vergleich	54
7.1	Was sind und was sollen die Zahlen?	54
7.2	Unendlichkeit	57
7.3	Das Paradox von Burali-Forti	59
7.4	Induktive Definitionen, Indizierungen und lange Wohlordnungen	61
7.5	‘Die Struktur des Unendlichen’	63
8	Schluss	67
8.1	Von Frege über Russell zu TT	67
8.2	Quine’s NF und der modifizierte logische Begriff der Menge	72
8.3	ZFC und der iterative Begriff der Menge	74
8.4	Intuitive Begriffe, Formalisierungen und Vergleichbarkeit von Mengenlehren	78

9 Anhang	83
10 Literatur	85

1 Vorwort

Es gibt eine Geschichte der Mengenlehre - manche meinen es sei *die* Geschichte der Mengenlehre - die sich ungefähr so erzählt: Am Anfang steht die sogenannte 'naive Mengenlehre'. Je nachdem wer die Geschichte erzählt, werden mit der naiven Mengenlehre Namen wie Cantor, Dedekind, Frege oder auch Bolzano verbunden. Themen der naiven Mengenlehre sind in den meisten Fällen Fragen betreffend die Grundlagen der Arithmetik, i.e. Fragen, die um den Begriff der 'Zahl' kreisen. Dedekind entdeckte 1872, dass man mit Mengen die reellen Zahlen aus den rationalen Zahlen 'konstruieren' kann. Cantor zeigte, dass es verschiedene Größen von unendlichen Mengen, dass es 'mehr' reelle als natürliche Zahlen gibt und dass man die Begriffe der Ordinal- und Kardinalzahl in natürlicher Weise auf unendliche Zahlen erweitern kann. Frege versucht (wie auch Dedekind) das 'Heiligtum' der Mathematik - die natürlichen Zahlen - auf 'logische' Grundlagen zu stellen, und zu diesen logischen Grundlagen gehören nach Frege auch - Mengen. Das wesentliche *feature* der naiven Mengenlehre ist, dass jede Eigenschaft eine Menge bestimmt, nämlich die Menge aller Dinge, die diese Eigenschaft haben. Man kann dieses Prinzip aus Cantors berühmter 'Definition' einer 'Menge' aus dem Jahr 1883 herauslesen:

"Unter einer Mannigfaltigkeit oder Menge verstehe ich nämlich allgemein jedes Viele, welches sich als Eines denken läßt, d.h. jeden Inbegriff bestimmter Elemente, welcher durch ein Gesetz zu einem ganzen verbunden werden kann..."¹

Relativ bald tauchen die ersten 'Ungereimtheiten' auf. Cantor selbst stellt in einem Brief an Dedekind fest, dass die Menge aller Ordinalzahlen keine 'konsistente Vielheit' sein kann. Ein weiteres Problem ergibt sich durch den Satz von Cantor, der besagt, dass die Potenzmenge $\mathcal{P}(A)$ einer Menge A immer größer ist als A selbst. Auf die Menge aller Dinge angewandt, würde das ergeben, dass es eine Menge gibt, die größer ist als die Menge aller Dinge.

Doch der *knockout* für die naive Mengenlehre - vor allem aber für Frege - kommt mit Russell. Unbestreitbar ist es eine 'Eigenschaft' einer Menge, nicht Element von sich selbst zu sein. Nach dem naiven Komprehensionsprinzip von oben kann also die Menge gebildet werden, die aus allen Mengen besteht, die sich nicht selbst enthalten. Fragt man nun, ob sich diese Menge selbst enthält oder nicht, gelangt man zu dem bekannten Widerspruch.

Die Widersprüche, die sich aus dem naiven Umgang mit Mengen ergaben, hatten zur Folge, dass man sich klarer zu werden versuchte über die Grundlagen

¹Auch die noch berühmtere 'Definition' des Mengenbegriffs von 1895 - 1897 ist in dieser Hinsicht problematisch. Dort steht: "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 könnte aus dieser 'Definition' ebenfalls ein naives Komprehensionsprinzip herauslesen, das *beliebige Zusammenfassungen* erlaubt.

dieser neuen Disziplin und es wurden verschiedene Vorschläge gemacht, die Paradoxien zu vermeiden. Die beiden wichtigsten sind die Typentheorie von Russell und die Axiomatisierung von Zermelo. In beiden Fällen gibt es ein Bündel von Axiomen, aus denen dann alle weiteren Tatsachen über Mengen abgeleitet werden können.

Der Vorschlag von Russell bedeutet - in der mengentheoretischen Sichtweise - einfach eine Beschränkung der Regeln, wie man Sätze in der Sprache der Mengenlehre bilden kann. Eine Menge, die sich selbst enthält (oder auch nicht enthält), kann es deshalb nicht geben, weil hinter dem $\in$ in einem Satz wie $x \in y$, y den nächsthöheren von x *Typ* haben muss, um überhaupt *sinnvoll* zu sein. ‘Mengentheoretische Sichtweise’ deshalb, weil bei Russell Mengen eigentlich nicht vorkommen (als *Objekte* seiner Theorie). ‘Dinge’ der Typentheorie Russells sind *Funktionen*, *Funktionen von Funktionen* und dgl.

Eine andere Möglichkeit, den Antinomien zu entkommen, ist jene von Zermelo. Der Ausweg besteht im Wesentlichen darin, keine ‘zu großen’ Mengen zuzulassen, d. h. man beschränkt das Komprehensionsaxiom derart, dass es ‘Zusammenfassungen’ nur mehr dann erlaubt, wenn man die Dinge, auf die eine Eigenschaft zutrifft, aus einer schon gegebenen Menge aussondern kann. Zusätzlich werden, neben einem Axiom, das sicherstellen soll, dass Mengen allein durch ihre Elemente bestimmt sind, diesem *Aussonderungsprinzip* Axiome beigelegt, die die Existenz schon gegebener Mengen gewährleisten.

Der weitere Hergang der Geschichte verläuft dann relativ geradlinig. Es werden noch zwei Axiome in den Kanon aufgenommen (Ersetzung und Fundierung von Fraenkel und Von Neumann), alle Axiome werden in eine prädikatenlogische Fassung gebracht (Skolem) und die daraus resultierende First-Order-Theorie heißt dann ZFC (*Zermelo-Fraenkel-Choice*, *Choice* für das Auswahlaxiom).

Man hört heute wenig von Alternativen zu ZFC, und wenn, dann sind es meistens Varianten von ZFC, die im Prinzip zu gleichen oder ähnlichen Begriffen von ‘Menge’ und ‘Enthalten-Sein’ führen (bzw. von im Prinzip gleichen Intuitionen ausgehen). ZFC ist das bei weitem gängigste Axiomensystem der Mengenlehre, auch NBG (Neumann-Bernays-Gödel), eine konservative Erweiterung von ZFC, ist noch relativ gebräuchlich. NBG lässt auch Quantifikation über (echte) Klassen (den Dingen die in ZFC ‘zu groß’ sind, um noch Mengen sein zu können) zu, liefert aber für Mengen dieselben Resultate.

Eine ‘wirkliche’ Alternative soll im Folgenden vorgestellt und im Vergleich zu ZFC besprochen werden. Das System geht zurück auf Quine’s NF (New Foundations), das er in seinem berühmten Aufsatz ‘*New Foundations for mathematical logic*’ von 1937 vorgestellt hat. Das System, das ich hier besprechen möchte, ist im Wesentlichen die Version von Jensen, bzw. von Holmes, und wird mit NFU bezeichnet.

Ein ‘wirkliche’ Alternative ist NFU aus mehreren Gründen: In NFU gibt es Mengen, die in ZFC nicht existieren. Die Menge aller Mengen, die es in ZFC nicht gibt, existiert, ebenso die Menge aller Ordinalzahlen. Die Kardinal- und

Ordinalzahlen in NFU sind völlig verschieden von denen in ZFC. Bestimmte Sätze aus der Theorie der Kardinalzahlen, etwa der Satz von Cantor, gelten nicht (oder zumindest nur eingeschränkt).

Der wesentlichste Unterschied zwischen den Systemen liegt aber nicht so sehr in der ‘Mathematik’ (Dingen wie: Was ist eine Kardinalzahl und was gilt für Ordinalzahlen? Wie repräsentiert man übliche mathematische Objekte wie Gruppen, topologische Räume? usw.), sondern in den *Intuitionen* die dahinter stecken. Man könnte sagen, der Unterschied ist ein genuin *Philosophischer*. Diese Zusammenhänge etwas näher zu beleuchten wird den Schlussteil dieser Arbeit ausmachen.

Bevor es losgeht noch einige Bemerkungen: Der Aufbau von NFU im Allgemeinen und der Arithmetik in NFU im Besonderen folgt im Großen und Ganzen dem von Holmes’ Buch ‘*Set Theory with a Universal Set*’. Bestimmte, mir nicht wichtig erscheinende Dinge wurden weggelassen, andere, die mir unzureichend (oder schlicht falsch) dargestellt schienen, wurden genauer beleuchtet. Vor allem aber wurde das System, das bei Holmes noch eine eher informelle Gestalt hatte in eine präzisere prädikatenlogische Form gebracht. Die Abschnitte, die sich mit dem Aufbau von ZFC beschäftigen, benutzen vor allem die wunderbare ‘*Einführung in die Mengenlehre*’ von Ebbinghaus und das Skriptum ‘*Axiomatic Set Theory 1*’ von Agatha Walczak-Typke.

Beide vorgestellten Systeme sind *First-Order-Theorien*, d.h. beide benutzen als *Hintergrundlogik* irgendeinen prädikatenlogischen Kalkül erster Stufe mit Identität (derer es wie Sand am Meer gibt). Die (wohlgeformten) *Formeln* baut man wie üblich rekursiv über die *atomaren Formeln* auf. Was eine atomare Formel ist, hängt davon ab, welche *nichtlogischen Konstanten*² zur Verfügung stehen. Die mit Abstand wichtigste (und für ZFC *einzige*) nichtlogische Konstante, mit der wir es im Folgenden zu tun haben werden, ist die zweistellige $\in$ -Relation, d.h. die wichtigste atomare Formel ist der Form $x \in y$.

Hier noch einige Konventionen: Oft wird statt $\forall x(x \in A \rightarrow \varphi(x))$ der kürze halber $\forall x \in A(\varphi(x))$ geschrieben (analog $\exists x \in A(\varphi(x))$ statt $\exists x(x \in A \wedge \varphi(x))$) und falls mehrere Variablen auf eine Menge beschränkt sind, schreiben wir $\forall x_1, \dots, x_n \in A(\varphi(x_1, \dots, x_n))$ statt der ausführlichen Schreibweise $\forall x_1 \dots \forall x_n((x_1 \in A \wedge \dots \wedge x_n \in A) \rightarrow \varphi(x_1, \dots, x_n))$. Die Menge aller Variablen, die in einer Formel φ vorkommen, bezeichnen wir mit $\text{var}(\varphi)$ und für $\neg(x \in y)$ schreiben wir $x \notin y$.

Falls wir von einem x Existenz und Eindeutigkeit beweisen können, so schreiben wir für ‘das eindeutig bestimmte x mit $\varphi(x)$ ’ manchmal auch ‘ $\iota x \varphi(x)$ ’³. Der iota-Operator bildet gewissermaßen den umgangssprachlichen bestimmten Artikel nach. Definitionen von Objekten über solche ‘Kennzeichnungen’ (*definite descriptions*) sind in der Mathematik keine Seltenheit. Selten sind eher die Anlässe, wo so etwas explizit zur Sprache gebracht wird. Wir wollen aber den Gebrauch

²Die Menge der nichtlogischen Konstanten, in der eine bestimmte Theorie formuliert ist, nennen wir auch *Sprache*.

³Die iota-Notation stammt von Bertrand Russell.

dieses iota-Operators so unproblematisch wie möglich halten und ein einfaches Beispiel zur Erläuterung bringen: Falls wir von einer Folge $(a_n)_{n \in \mathbb{N}}$ von reellen Zahlen zeigen können, dass $\varphi(a) :\leftrightarrow \forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |a_n - a| < \epsilon$, dann sagen wir, dass diese Folge *konvergiert*. Wir können nun zeigen, dass ein a , das obige Formel erfüllt, eindeutig bestimmt sein muss, falls es überhaupt eines gibt. Für dieses eindeutig bestimmte a schreibt man üblicherweise auch $\lim_{n \rightarrow \infty} a_n$, d.h. mit dem iota-Operator können wir definieren: $\lim_{n \rightarrow \infty} a_n := \iota a \varphi(a)$.

Wir werden im Weiteren immer wieder ungezwungen vom iota-Operator Gebrauch machen.⁴

⁴Falls man mehr über ‘Kennzeichnungen’ und dgl. erfahren möchte, wende man sich z.B. an Ramharter/Rieck [2006] S. 35 ff.

2 Das System NFU

NFU ist eine Theorie erster Stufe in der Sprache $\{\in, Set, (\cdot)\}$, wobei $\in$ eine zweistellige Relation, Set ein einstelliges Prädikat und $(\cdot)$ eine zweistellige Funktion ist. Intuitiv stellt man sich $\in$ als die *Element-Relation* vor und Set als Eigenschaft eine Menge zu sein (im Gegensatz zu *Urelementen* oder *Atomen*). $(\cdot)$ ist die Funktion, die je zwei Objekten x und y (dem geordneten Paar (x, y) im ‘metasprachlichen Sinn’) das geordnete Paar (x, y) zuordnet.⁵

Die Axiome von NFU⁶ sind:

1. $\exists y [Set(y) \wedge \forall x (x \in y \leftrightarrow \varphi(x))]$ (S-Komprehensionsschema)
ist ein Axiom falls $\varphi(x)$ stratifiziert ist und y nicht frei in $\varphi(x)$ vorkommt.⁷
2. $\forall x \forall y [(Set(x) \wedge Set(y) \wedge \forall z (z \in x \leftrightarrow z \in y)) \rightarrow x = y]$ (Extensionalität)
3. $\forall x [\neg Set(x) \rightarrow \neg \exists y (y \in x)]$ (Atome)
4. $\forall x \forall y \forall u \forall v ((x, y) = (u, v) \rightarrow x = u \wedge y = v)$ (Geordnete Paare)
5. $\forall X [\forall x \forall y (x \in X \wedge y \in X \rightarrow x \neq \emptyset \wedge (x = y \vee x \cap y = \emptyset))] \rightarrow$
 $\exists Y \forall x (x \in X \rightarrow \exists z (Y \cap x = \{z\}))$ (Auswahlaxiom)⁸

Definition 2.1. Eine Formel $\varphi(x)$ mit x frei heißt *stratifiziert*, wenn es eine Funktion $\sigma : var(\varphi) \rightarrow N$ gibt, derart, dass für jede atomare Formel in $\varphi(x)$

1. $\sigma(x) = \sigma(y)$ falls $x = y$

⁵Die Begriffe ‘Funktion’ und ‘Relation’ (sowie ‘Prädikat’ für ‘einstellige Relation’) sind dreideutig: Z.B. verwendet man den Begriff ‘Relation’ einmal *syntaktisch* für eine *Relationskonstante* im Sinne der logischen Konstanten einer First-Order-Sprache (hier etwa die $\in$ -Relation); das andere mal verwendet man den Begriff auch für Relationen im *semantischen*, mengentheoretischen Sinne, (wie er in 2.2. definiert wird) als Menge von geordneten Paaren. Weiters versteht man unter einer Relation auch ein zweistelliges, definierbares Prädikat in der Sprache der jeweiligen Mengenlehre. Grob gesprochen besteht der Zusammenhang zwischen den drei Verwendungsweisen darin, dass man 1. aus Relationen im ersten Sinne Formeln bilden kann, die dann Relationen im 3. Sinne sein können, die wiederum (möglicherweise) Relationen im 2. Sinne (Mengen von geordneten Paaren) bestimmen.

⁶Wir werden im Folgenden unter NFU immer die Axiome 1 - 5 verstehen, insbesondere also das Auswahlaxiom dazurechnen. In der Literatur meint man mit NFU meistens nur die Axiome 1 - 4. Weil das Auswahlaxiom aber auch bei ZFC ‘immer schon dabei’ ist, und es auch für Mengenlehren der Art von NFU von tragender Bedeutung ist, wird es in dieser Arbeit als Bestandteil der Axiome von NFU angesehen.

⁷Genauer: Falls die freien Variablen von φ $x, x_1, x_2, \dots, x_n$ sind, und y nicht frei in φ vorkommt, dann ist der universelle Abschluss von $\exists y \forall x (x \in y \leftrightarrow \varphi(x, x_1, x_2, \dots, x_n))$ ein Axiom

⁸Diese Formulierung des Auswahlaxioms benutzt Zeichen, die an dieser Stelle eigentlich noch nicht bekannt sind. Strenggenommen ‘verstehen wir’ nämlich noch nicht, was die Zeichen $\{x\}$ und $\emptyset$ bedeuten. Wir setzen hier auf das Vorverständnis des Lesers; Rechtfertigen lässt sich diese Vorgangsweise, weil im *Aufbau der Theorie* das Auswahlaxiom erst benötigt wird, wenn die erwähnten Zeichen schon definiert worden sind. Ähnliches gilt für die Axiome von ZFC, die im nächsten Abschnitt präsentiert werden.

2. $\sigma(x) + 1 = \sigma(y)$ falls $x \in y$
3. $\sigma(x) = \sigma(y) = \sigma(v)$ falls $v = (x, y)$

Bemerkungen

- Die ersten beiden Bedingungen dieser Definition besagen, grob gesprochen, dass eine Formel stratifiziert ist, wenn sie wohlgeformt wäre gemäß der Typentheorie. Die dritte Bedingung besagt erstens, dass u und v in einem Paar (u, v) vom selben Typ sein müssen und zweitens, dass überdies auch das Paar (u, v) selbst vom selben Typ wie u und v ist. Man sagt auch, dass Paar ist ein *type-level-orderd pair*.⁹
- Axiom (1) garantiert für jede stratifizierte Elementbedingung die Existenz der Menge aller Dinge, die dieser Bedingung genügen. Die dadurch bestimmten Mengen schreiben wir auch $\{x : \varphi(x)\}$. Es ist zu beachten, dass Stratifiziertheit von $\varphi(x)$ eine hinreichende, aber keine notwendige Bedingung für die Existenz von $\{x : \varphi(x)\}$ ist.¹⁰ D.h. manchmal kann es durchaus möglich sein, *über Umwege* die Existenz einer Menge zu beweisen, auch wenn die definierende Bedingung nicht stratifiziert ist.
- Axiom (2) besagt, dass zwei Mengen als gleich gelten, falls sie dieselben Elemente enthalten. Die Einschränkung auf Mengen ist eine notwendige Bedingung für die Existenz von Dingen, die keine Elemente haben, sogenannten Urelementen oder Atomen.
- (3) fordert von allen Nicht-Mengen, also Atomen, dass sie elementlos sind; dass es überhaupt Atome gibt, steht von vornherein noch nicht fest. *Falls* es aber Atome gibt, so besteht das Universum, über das wir quantifizieren, genau aus den Mengen und Atomen und nichts anderem, d.h. jedes ‘Ding’, jeder mögliche Wert einer Variable, ist entweder *Menge* oder *Atom*.
(4) fordert einfach nur die Eigenschaft, die jedes geordnete Paar erfüllen muss.
- Das Auswahlaxiom wird im folgenden noch keine Rolle spielen und wurde nur der Vollständigkeit halber in die Liste mit aufgenommen. Zum ‚Einsatz‘ kommt es erst, wenn es darum geht, grundlegende arithmetische Begriffe zu entwickeln. Intuitiv besagt es, dass man zu jeder Menge von disjunkten Mengen eine sogenannte Auswahlmenge bilden kann, die genau ein Element jeder dieser Mengen enthält.

⁹In Holmes [1998] wird das sichergestellt, indem gefordert wird, dass die linke und rechte Projektion eines geordneten Paares jeweils vom selben Typ wie das geordnete Paar selbst sein muss.

¹⁰Und anders als Rosser in [1953] suggeriert, war sich Quine dieser Tatsache durchaus bewusst. Vgl. Rosser [1953] S. 205 ff.

2.1 Einfache Folgerungen

Eine erste (und eine der wichtigsten) Folgerungen aus den Axiomen ist

Satz 2.1. $\exists y [Set(y) \wedge \forall x (x \in y \leftrightarrow x = x)]$

Wir erhalten diesen Satz durch Einsetzen der stratifizierten Formel $x = x$ für $\varphi(x)$ in das S-Komprehensionsschema. Die durch diesen Satz eindeutig bestimmte Menge, die *alles* umfasst, nennen wir das *Universum* oder die *Allmenge* und bezeichnen sie mit V . Durch Einsetzung von $Set(x)$ für $\varphi(x)$ bekommen wir

Satz 2.2. $\exists y \forall x (x \in y \leftrightarrow Set(x))$

Die durch 1.2. eindeutig bestimmte Menge aller Mengen bezeichnen wir mit V_{Set} .

Satz 2.3. $\exists y \forall x (x \in y \leftrightarrow x \notin A)$

1.3. sichert die Existenz der Komplemente A^C jeder Menge A . Mit 1.1. liefert das die leere Menge $\emptyset := V^C$.

Mit der Definition $A \subseteq B := \forall x (x \in A \rightarrow x \in B)$ erhalten wir

Satz 2.4. $\exists y \forall x (x \in y \leftrightarrow x \subseteq A)$

Die durch 1.4. definierte Menge heißt *Potenzmenge* von A und wir schreiben dafür $\mathcal{P}(A)$.

Korollar 2.1. $V_{Set} = \mathcal{P}(V)$

Weitere Instanzen von (1) ergeben die Existenz der Mengen $\{x\}$, $x \cup y$ (boolesche Vereinigung) und $x \cap y$ (boolescher Durchschnitt).

Die für NFU sehr wichtige und oft benutzte Operation, die jedes x auf seine Singleton $\{x\}$ ‘abbildet’ nennen wir sinnigerweise auch ‘Singleton-Operation’. Für das iterierte Singleton $\{\{x\}\}$ verwenden wir auch die Bezeichnung ‘SSingleton’.

Das Universum V bildet mit $\cup, \cap$ und der Komplement-Operation eine *boolesche Algebra*.

Für eine Menge $\mathcal{F}$ von nichtleeren Mengen existieren die Mengen $\bigcup \mathcal{F} := \{x : \exists A (x \in A \wedge A \in \mathcal{F})\}$ und $\bigcap \mathcal{F} := \{x : \forall A (A \in \mathcal{F} \rightarrow x \in A)\}$.¹¹

2.2 Relationen und Funktionen

Relationen (im mengentheoretischen Sinne) sind Mengen von geordneten Paaren.

Eine Instanz von (1) erlaubt uns die Bildung des kartesischen Produkts $A \times B$ von zwei Mengen: $A \times B := \{x : x = (a, b) \wedge a \in A \wedge b \in B\}$.

Für Relationen R , für die gilt: $R \subseteq A \times B$ können wir definieren:

¹¹Für Mengen von Atomen sind Durchschnitt und Vereinigung nicht definiert.

Definition 2.2. 1. $\text{dom}(R) := \{a : a \in A \wedge \exists b \in B((a, b) \in R)\}$

2. $\text{rng}(R) := \{b : b \in B \wedge \exists a \in A((a, b) \in R)\}$

3. $R|C := \{(x, y) : (x, y) \in R \wedge x \in C\}$ für $C \subseteq A$

4. $R''C := \text{rng}(R|C)$

Die Mengen, die durch 1 - 4 definiert sind nennen wir auch, die *Domain*, den *Wertebereich*, die *Einschränkung* einer Relation auf eine Menge C und das *Bild* einer Menge C .

Wir sagen von einer Relation R , dass sie *auf einer Menge A* ist, wenn $\text{dom}(R) = \text{rng}(R) = A$, i.e. $R \subseteq A \times A$ und schreiben in Zukunft auch xRy falls $(x, y) \in R$.

Falls R eine Relation ist, dann bezeichnet R^{-1} die inverse Relation, und es ist $R^{-1} := \{(x, y) : yRx\}$.

Für eine (zweistellige) Relation R auf einer Menge A bezeichne $SI(R)$ die - nach S-Komprehension existierende - eindeutig bestimmte Menge $\{(\{x\}, \{y\}) : xRy\}$. Allgemein soll $SI^n(R)$ diejenige Relation sein, die man durch n -malige Anwendung der Singleton-Operation auf Elemente von A erhält.

2.3 Ordnungsrelationen

Für die folgenden Definitionen denken wir uns die Quantoren auf die jeweilige Menge, auf der die Ordnungsrelation definiert ist, beschränkt.

Definition 2.3. Eine Relation R auf einer Menge A heißt dann eine *strikte Totalordnung* (oder *Totalordnung im Sinne von $<$*) falls

1. $\forall x \neg(xRx)$
2. $\forall x \forall y \forall z (xRy \wedge yRz \rightarrow xRz)$
3. $\forall x \forall y (xRy \vee yRx \vee x = y)^{12}$

Eine Relation R auf einer Menge A heißt *schwache Totalordnung* (oder *Totalordnung im Sinne von $\leq$*) wenn statt (1)

- $\forall x (xRx)$

gilt und zusätzlich

- $\forall x \forall y (xRy \wedge yRx \rightarrow x = y)$

¹²Formal ist eine 'Ordnungsrelation auf einer Menge A ' (wie jede Struktur, die wir mengentheoretisch auffassen wollen) ein *geordnetes Paar* $z := (R, A)$ (bzw. geordnetes n -Tupel), derart, dass für alle Elemente $x \in \text{pr}_2(z)$ gilt: $\neg(x\text{pr}_1(z)x)$. usw. Wir unterdrücken im Folgenden diese 'Paarnotation' - der besseren Lesbarkeit und Suggestivität halber - und verwenden die Ausdrucksweise, die bisher entwickelt wurde.

Eine lineare Ordnung R auf einer Menge A heie *Wohlordnung*, falls gilt:

- $\forall B [(B \neq \emptyset \wedge B \subseteq A) \rightarrow \exists y \in B \forall x \in B yRx]$ ¹³

Falls R eine Wohlordnung auf einer Menge A ist, schreiben wir dafr abkrzend auch $WO(R_A)$ und falls $a \in A$, dann definieren wir das durch a eindeutig bestimmte *Anfangsstck* bzgl. R als

Definition 2.4. $Pred(a, R, A) := \{x \in A : xRa\}$

Die Einschrnkung einer Wohlordnung auf $Pred(a, R, A)$ schreiben wir $R_{Pred(a, R, A)}$ und definieren $R_{Pred(a, R, A)} := R \cap (Pred(a, R, A) \times Pred(a, R, A))$

Die Einschrnkung einer Wohlordnung auf $Pred(a, R, A)$ ist selbst eine Wohlordnung.

Wenn nicht ausdrcklich anders erwhnt, soll sich im Weiteren $Pred(a, R, A)$ immer auf eine strikte Ordnung R beziehen, d. h. $Pred(a, R, A)$ bezeichnet immer ein *echtes Anfangsstck* der Wohlordnung R .

2.4 quivalenzrelationen

Definition 2.5. Eine Relation $\sim$ auf einer Menge A heit *quivalenzrelation*, wenn gilt:

1. $\forall x, y, z (x \sim y \wedge y \sim z \rightarrow x \sim z)$
2. $\forall x \neg(x \sim x)$
3. $\forall x, y (x \sim y \rightarrow y \sim x)$

Definition 2.6. Eine Partition einer Menge A ist eine Familie F von Teilmengen von A derart, dass

- $\forall x, y \in A (x \cap y = \emptyset)$
- $\bigcup F = A$

Falls $\sim$ eine quivalenzrelation auf einer Menge A ist, definiert $\sim$ eine Partition von A . Die paarweise disjunkten Elemente von F nennen wir *quivalenzklassen* und schreiben dafr $\bar{x}, \bar{y}, \dots$. Eine Element $x \in \bar{x}$ heit *Reprsentant* der quivalenzklasse $\bar{x}$ und eine Menge die aus jeder quivalenzklasse (genau) einen Reprsentanten enthlt, heit *Reprsentantensystem*.¹⁴

¹³Ein Beispiel fr eine strikte Totalordnung ist die bliche Ordnung $<$ auf den natrlichen Zahlen. Definiert man $x \leq y :\leftrightarrow x < y \vee x = y$ erhlt man eine schwache Totalordnung auf den natrlichen Zahlen.

¹⁴Dass jede quivalenzrelation ein Reprsentantensystem besitzt folgt aus dem Auswahlaxiom.

2.5 Funktionen

Eine *Funktion von A nach B* ist eine Relation, die jedem $a \in A$ genau ein $b \in B$ zuordnet. Wir schreiben manchmal auch $Funk(f)_{A,B}$ für 'f ist eine Funktion von A nach B', wobei A der *Definitionsbereich* und B der *Zielbereich* der Funktion ist.

Definition 2.7. $Funk(f)_{A,B} := f \subseteq A \times B \wedge \forall a \in A \exists! b \in B ((a, b) \in f)$

Das eindeutig bestimmte y , derart dass $(a, b) \in f$, bezeichnen wir mit $f(a)$. Wir werden im Weiteren auch die üblichere Schreibweise $f : A \longrightarrow B$ für $Funk(f)_{A,B}$ verwenden.

Eine Funktion $f : A \longrightarrow B$ heißt *surjektiv*, falls $\forall b \in B \exists a \in A (f(a) = b)$ und *injektiv*, falls $\forall a, a' \in A (a \neq a' \rightarrow f(a) \neq f(a'))$. Eine Funktion heißt *bijektiv*, falls sie injektiv und surjektiv ist.

Für eine Menge A bezeichne id_A die *identische Funktion* auf A , für die gilt: $id_A(a) = a$ für alle $a \in A$. Falls f und g Funktionen sind, derart, dass $f : A \longrightarrow B$ und $g : B \longrightarrow C$, dann bezeichnet $g \circ f$ die *Komposition* (Hinter-einanderausführung) der Funktionen f und g , sprich: $(g \circ f)(a) = g(f(a))$.

Eine Operation¹⁵, die für die Arithmetik in NFU eine zentrale Rolle spielen wird, ist diejenige, die jeder Menge A die Menge ihrer Singletons zuordnet. Wir bezeichnen die Menge aller Singletons einer Menge A mit $\mathcal{P}_1(A)$, die Menge aller Singletons von Singletons einer Menge A als $\mathcal{P}_2(A)$ usw.¹⁶

¹⁵Man verwendet den Begriff *Operation* in der Mengenlehre meistens um sich von *Funktionen* abzugrenzen. Funktionen sind dabei Objekte (Mengen von geordneten Paaren) der jeweiligen Mengenlehre. Operationen dagegen müssen das nicht sein. Wir werden später sehen, dass die Singleton-Operation $x \mapsto \{x\}$ in NFU keine Funktion ist (noch weniger natürlich in ZFC).

¹⁶Ist z.B. $A = \{a, b, c\}$, dann ist $\mathcal{P}_1(A) = \{\{a\}, \{b\}, \{c\}\}$

3 Das System ZFC

ZFC ist eine Theorie erster Stufe in der Sprache $\{\in\}$. Die Axiome von ZFC sind:

1. $\forall x \forall y (\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y)$ (Extensionalität)
2. Wenn $\varphi(x)$ eine Formel in der Sprache $\{\in\}$ und y nicht frei in $\varphi(x)$ vorkommt, dann ist $\forall z \exists y \forall x (x \in y \leftrightarrow x \in z \wedge \varphi(x))$ ein Axiom (Aussonderung)
Wir schreiben für solche Mengen auch $\{x \in z : \varphi(x)\}$.
3. $\forall x \forall y \exists z (x \in z \wedge y \in z)$ (Paarmenge)
4. $\forall v \exists z \forall x \forall y (x \in y \wedge y \in v \rightarrow x \in z)$ (Vereinigung)
5. Wenn $\varphi(x, z)$ eine Formel ist mit x und z frei, und y nicht frei in $\varphi(x, z)$ vorkommt, dann ist $\forall x \in A \exists! z \varphi(x, z) \rightarrow \exists y \forall x \in A \exists z \in y \varphi(x, z)$ ein Axiom. (Ersetzung)
6. $\exists x (\emptyset \in x \wedge \forall y \in x (y \cup \{y\} \in x))$ (Unendlichkeit)
7. $\forall x \exists y \forall z (z \subseteq x \rightarrow z \in y)$ (Potenzmenge)
8. $\forall x (x \neq \emptyset \rightarrow \exists y \in x \forall z (z \notin y \vee z \notin x))$ (Fundierung)
9. $\forall X [\forall x \forall y (x \in X \wedge y \in X \rightarrow x \neq \emptyset \wedge (x = y \vee x \cap y = \emptyset))] \rightarrow \exists Y \forall x (x \in X \rightarrow \exists z (Y \cap x = \{z\}))$ (Auswahlaxiom)¹⁷

Bemerkungen

- Wie auch in NFU stellt (1) sicher, dass zwei Mengen gleich sind, wenn sie dieselben Elemente haben. Da die einzigen Objekte in ZFC Mengen sind, brauchen wir uns keine Gedanken über Atome oder andere Nicht-Mengen machen.
- (2) ist die ‘ZFC-Entsprechung’ des S-Komprehensionsschemas in NFU. Das Aussonderungsschema erlaubt die Bildung von Mengen, indem man aus einer bestimmten vorgegebenen Menge diejenigen Elemente aussondert, die der Eigenschaft $\varphi(x)$ genügen.
- Axiom (3) stellt für beliebige x, y die Existenz von $\{x, y\}$ (und damit von $\{x\} = \{x, x\}$) sicher, (4) die Existenz der Vereinigungsmenge $\bigcup v$ für eine Menge v von Mengen und (7) die Existenz der Potenzmenge.
- Wie das Aussonderungsaxiom ist auch das Ersetzungsaxiom ein Axiomenschema. Inhaltlich besagt es, dass, wenn die Formel $\varphi(x, y)$ ‘funktional’ ist, i.e. wenn $\varphi(x, y)$ eine Operation auf V definiert, dann existiert das Bild jeder Menge unter dieser Operation.

¹⁷Manchmal werden wir auch Teilsysteme von ZFC betrachten. Dabei soll ZF für das System aller aufgelisteten Axiome bis auf das Auswahlaxiom stehen. ZFC^- soll das System bezeichnen, dass alle Axiome bis auf das Fundierungsaxiom beinhaltet.

- Das Fundierungsaxiom stellt unter anderem sicher, dass es keine unendlichen absteigenden ‘ $\in$ -Folgen’ gibt, d.h. dass jede Menge ein ‘ $\in$ -minimales’ Element hat. Man umschreibt den Umstand, der in (8) ausgedrückt auch so: *Die $\in$ -Relation ist fundiert.*

3.1 Einfache Folgerungen

Mit dem Aussonderungsaxiom können wir zeigen, dass $\exists y \forall x (x \notin y)$. Diese nach (1) eindeutig bestimmte Menge nennen wir die *leere Menge* und bezeichnen sie wie auch in NFU mit $\emptyset$.

Satz 3.1. $\neg \exists y \forall x (x \in y)$

Satz 2.1. besagt, dass es keine Menge gibt, die alle Mengen umfasst, d.h. die Allmenge (das Universum) V existiert nicht.

Eine direkte Folge aus dem Fundierungsaxiom ist

Satz 3.2. $\neg \exists x_1, \dots, x_n (x_1 \in x_2 \wedge x_2 \in x_3 \dots \wedge x_n \in x_1)$

Insbesondere schließt das Fundierungsaxiom Mengen x aus, für die $x \notin x$.

3.2 Funktionen und Relationen

Mit dem Paarmengenaxiom von oben kann man das sogenannte *Kuratowski-geordnete Paar* definieren:

Definition 3.1. $(x, y) := \{\{x\}, \{x, y\}\}$ ¹⁸

Man kann leicht nachprüfen, dass die für das geordnete Paar einzig wichtige Bedingung $(x, y) = (x', y') \rightarrow x = x' \wedge y = y'$ erfüllt ist.

Das kartesische Produkt $A \times B$ definieren wir durch Benutzung des Ersetzungsaxioms.¹⁹

Die Definitionen bzgl. Relationen und Funktionen können aus dem vorigen Abschnitt übernommen werden. Der einzige Unterschied ist, dass wir statt der primitiven ‘Paarfunktion’ nun eine mengentheoretische ‘Implementation’ des geordneten Paares haben.

Ein Wort noch zum Auswahlaxiom: Wir hatten das Auswahlaxiom in der Form vorgestellt, dass es für jede Menge von Mengen eine weitere Menge gibt,

¹⁸Wir verwenden auch hier die Notation (x, y) für das geordnete Paar. Man sollte sich jedoch im Klaren sein, dass in NFU damit ein *primitives Zeichen* gemeint ist, während wir hier in ZFC ein durch mengentheoretische Operationen *definiertes* Paar verstehen.

¹⁹Für ein beliebiges $b \in B$ gilt: $\forall a \in A \exists! c (c = (a, b))$. Mit dem Ersetzungsaxiom kann man die Menge $prod(A, b) := \{c : \exists a \in A (c = (a, b))\}$ bilden. Da $\forall b \in B \exists! c (c = prod(A, b))$ kann man durch nochmalige Anwendung des Ersetzungsaxioms die Menge $Prod(A, B) := \{prod(A, b) : b \in B\}$. Mit dem Vereinigungsmengenaxiom kann man dann $A \times B := \bigcup Prod(A, B)$ definieren.

die als Durchschnitt mit jeder einzelnen dieser Mengen ein Singleton hat. In der Mathematik werden aber häufig andere (äquivalente) Formulierungen gebraucht. Vor allem in der Theorie der Ordinal- und Kardinalzahlen wird das Auswahlaxiom in der Form des sogenannten ‘Wohlordnungssatzes’ benutzt, der besagt, dass *jede Menge wohlgeordnet werden kann*, i.e. $\forall A \exists R (WO(R_A))$. Wir können, sowohl auf der Basis von ZFC ohne Auswahlaxiom als auch auf der Basis von NFU ohne Auswahlaxiom zeigen, dass beide Formulierungen tatsächlich äquivalent sind und werden von diesem Resultat in beiden Systemen ohne Scheu Gebrauch machen.

4 Ein erster Vergleich

Schon beim Aufbau der beiden Mengenlehren wird klar, dass sich NFU und ZFC in wesentlichen Punkten voneinander unterscheiden. Abgesehen davon, dass der Aufbau von ZFC sozusagen ‘stufenweise’ erfolgt, während man in NFU qua S-Komprehensionsschema ‘alles auf einmal’ bekommt, beginnen die Unterschiede schon bei der Existenz und Eindeutigkeit der leeren Menge $\emptyset$.

Während ZFC nur *eine* Art von Dingen kennt, unterscheidet man in NFU zwischen Mengen und Urelementen. In ZFC gibt es nur ein ‘Ding’, das keine Elemente hat, nämlich die leere Menge; aber nicht jedes ‘Ding’ in NFU, das keine Elemente hat, ist notwendigerweise die leere Menge. Man kann also nicht - wie in ZFC - schliessen: Es gibt ein Ding, das keine Elemente hat, und qua Extensionalität ist dieses Ding eindeutig bestimmt; ‘Ding’ und ‘Menge’ stimmen in ZFC überein, in NFU sind aber nur bestimmte Dinge Mengen. In NFU muss, um Existenz und Eindeutigkeit der leeren Menge sicherzustellen, anders argumentiert werden. Bevor wir uns diese Sache etwas genauer ansehen, noch zwei Bemerkungen:

Die Tatsache, dass der Aufbau von ZFC ‘Schritt für Schritt’ erfolgt im Gegensatz zu NFU, wo man alles *ab initio* bekommt, muss man nicht unbedingt für einen ‘wesentlichen’ Unterschied zwischen den beiden Systemen halten. Holmes liefert z.B. eine endliche Axiomatisierung von NFU, die im Prinzip dem stufenweisen Aufbau von ZFC sehr ähnlich ist. D.h. auch in NFU können wir ‘schrittweise’ - analog zu ZFC - die Existenz bestimmter Mengen postulieren (z.b. die Existenz der Allmenge, der Komplemente, der Singletons usw.) und aus der Existenz dieser bestimmten Mengen dann das S-Komprehensionsschema als *Theorem* ableiten.

Dass NFU Urelemente zulässt und ZFC nicht, ist ‘unwesentlich’ nur *in einem bestimmten Sinne*. Auch in ZFC kann man das Extensionalitätsaxiom derart abschwächen, dass Urelemente zugelassen sind. Im Gegensatz zu NFU gilt aber für diese neue Theorie, dass sie nichts wesentlich Neues beweist. Für NFU im Vergleich zu NF (das ist das System, das nur aus dem starken Extensionalitätsaxiom und S-Komprehension besteht) gilt dieser Zusammenhang *nicht*. Während etwa das Auswahlaxiom integraler Bestandteil von NFU ist, ist es in NF schlicht *falsch*. NFU ist ein gutes Beispiel dafür, dass das Verhältnis zwischen Theorien, die Urelemente zulassen, und ihren atomlosen Gegenstücken nicht unbedingt gleichmäßig ist.

4.1 Extensionalität

Das Extensionalitätsaxiom besagt umgangssprachlich: Mengen mit gleichen Elementen sind gleich. Oder: Mengen sind allein durch ihre Elemente bestimmt. In ZFC bereitet das Extensionalitätsaxiom (kurz *Ex*) keine Probleme (wenn man das Folgende als ‘Problem’ betrachtet). Der Umstand, dass es bei Mengen ausschließlich auf die Elemente ankommt, wird korrekt durch

$$\forall x \forall y [\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y]$$

ausgedrückt. Nicht so in NFU. Dort wollen wir nämlich auch Urelemente zulassen, und wir wollen unter Urelementen solche Dinge verstehen, die keine Elemente haben. Aber auch die leere Menge soll keine Elemente haben und dennoch eindeutig sein.

Wollen wir Ex also wie oben beibehalten, stehen wir vor folgender Situation: Entweder es gibt nur ein Urelement und keine leere Menge oder es gibt eine eindeutig bestimmte leere Menge, aber keine Urelemente.

Um dies zu verhindern, beschränkt man Ex in NFU auf Mengen, i.e.

$$\forall x \forall y [Set(x) \wedge Set(y) \wedge \forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y]$$

Außerdem fordern wir für Urelemente, dass sie elementlos sein sollen, i.e.

$$\forall x [\neg Set(x) \rightarrow \neg \exists y (z \in x)]$$

Mit dem S-Komprehensionsschema ergibt das insgesamt, dass es eine Menge gibt, die keine Elemente hat und überdies mit Ex_{neu} , dass alle anderen Mengen ohne Elemente jener gleich sind, was die Definition *der* leeren Menge rechtfertigt. Außerdem aber stellt Ex_{neu} sicher, dass, falls x und y keine Mengen sind, mit Ex_{neu} keine Aussage bzgl. Gleichheit oder Ungleichheit getroffen werden kann.

Eine Unschönheit, die diese Änderung von Ex mit sich gebracht hat ist, dass wir die ursprüngliche Sprache, die nur aus der $\in$ -Relation und der primitiven Paarfunktion bestanden hat, um das primitive Prädikat ‘Set(x)’ erweitern mussten, und es stellt sich die natürliche Frage, ob das notwendig ist, oder ob sich dieses Prädikat irgendwie definieren lässt. Zumindest für $\in$ allein lässt sich die Frage verneinen, denn würde man $Set(x)$ durch $\exists y (y \in x)$ definieren, wäre auch die leere Menge ein Urelement, was nicht sehr wünschenswert ist. Z.B. hätten wir gerne, dass die leere Menge Teilmenge jeder Menge ist; für andere Urelemente soll das aber nicht gelten. Wir müssten also ein bestimmtes ‘Urelement’ (die leere Menge) ‘auszeichnen’ und es gibt keinen Grund anzunehmen, wir könnten dies mit Recht tun.

Auf andere Weise kann man sich das Prädikat $Set(x)$ dennoch sparen, indem man die leere Menge $\emptyset$ als Konstante in die Sprache aufnimmt. Wir definieren dann $Set(x)$ durch $\exists y (y \in x) \vee x = \emptyset$ und behalten Ex_{neu} bei. Was wir uns in dieser Variante sparen ist ein Axiom, das uns für $\neg Set(x)$ die notwendige Bedingung $\neg \exists y (y \in x)$ liefert. Für den Fall mit primitivem $Set(x)$ konnten wir ein solches Axiom nicht dispensieren, denn dass Elementlosigkeit notwendige Bedingung dafür ist, keine Menge zu sein, folgt weder aus Ex_{neu} noch aus dem S-Komprehensionsschema. Wir könnten etwa bei der Einführung von $Set(x)$ an die Unterscheidung Menge/Klasse gedacht haben oder uns durch $Set(x)$ von irgendwelchen anderen ‘Unmengen’ (Äpfeln, Fettbohnen oder Giraffen) abgrenzen wollen.

Dennoch bleibt der ‘Makel’ bestehen; ob wir nun $\emptyset$ oder $Set(x)$ als primitive Symbole dazunehmen; in jedem Fall bedeutet das eine echte Spracherweiterung, und insofern ist es zumindest nicht verwunderlich, wieso man neue und ‘überraschende’ Sätze beweisen kann. (‘Neu und überraschend’ im Vergleich zu NF).

4.2 Geordnete Paare

Geordnete Paare sind wichtig. Genau genommen sind sie sogar entscheidend dafür, dass eine bestimmte Mengenlehre ihren Anspruch auf ‘Grundlegung der Mathematik’ einlösen kann. Will man Gruppen, topologische Räume, Graphen etc. - vor allem aber Relationen und Funktionen - als Mengen auffassen, so braucht man dafür eine mengentheoretische Definition des geordneten Paares.

Ein auffälliger Unterschied zwischen NFU und ZFC liegt im Umstand, dass ZFC tatsächlich ein *definiertes* geordnetes Paar verwendet, um die Theorie der Relationen und Funktionen zu entwickeln, während NFU dafür ein *primitives* geordnetes Paar benutzt. Aber warum ist das so und vor allem: *Muss* das so sein?

Die Antwort auf die zweite Frage lautet schlicht: Nein. Man *muss* nicht. Der Grund dafür, dass man es dennoch macht und nicht wie in ZFC die Definition von Kuratowski verwendet, ist im wesentlichen ein *pragmatischer*. Das Kuratowski-Paar $\{\{x\}, \{x, y\}\}$ bereitet im Aufbau der Theorie einfach zu viel ‘Ungemach’. Das liegt daran, dass die relativen Typen der von x und y im Kuratowski-Paar um zwei niedriger sind als der relative Typ des Paares selbst. Das hat zur Folge, dass z.B. der relative Typ einer Funktion um drei höher ist als der Typ der Argumente und Funktionswerte. Man braucht deshalb immer wieder die Singleton-Operation und die Vereinigungsmengenoperation um bestimmte Definition zu stratifizieren, bzw. um diesen Prozess wieder rückgängig zu machen.

Um all diese Probleme von vornherein zu vermeiden, fordert man vom primitiven geordneten Paar in NFU, dass es *type-level* ist, dass also der relative Typ eines Paares $z = (x, y)$ der gleiche ist wie der von x und y .²⁰ In der hier präsentierten Form von NFU wurde dem Rechnung getragen, indem wir in unserer Definition der stratifizierten Formeln eine eigene Bedingung aufgenommen haben, die regelt, wann eine Formel stratifiziert ist, wenn sie ein Paar enthält.

Einer der auffälligsten Unterschiede zwischen NFU (und auch NF) und ZFC besteht darin, dass es in NFU eine Menge gibt, die alle Dinge umfasst, während die Existenz einer solchen Menge in ZFC zu Widersprüchen führen würde. Aber wie kommt es dazu?

4.3 Aussonderung vs. S-Komprehension

Eines der charakteristischen *features* von ‘Zermelo-style’ Mengenlehren (manche halten es für *das* charakteristische Feature) ist das Aussonderungsschema, das in seiner Standardversion besagt, dass

$$\forall z \exists y \forall x (x \in y \leftrightarrow x \in z \wedge \varphi(x))$$

²⁰Wie wir sehen werden, bringt das zusätzliche ‘Stärke’ in unser System, weil es uns ein Unendlichkeitsaxiom ‘erspart’.

wobei $\varphi(x)$ eine Formel mit x frei ist, in der y nicht frei vorkommt. In Worten: von einer beliebigen Menge z kann man die Teilmenge von z , die aus allen Dingen x aus z besteht, die $\varphi(x)$ erfüllen, bilden.

Entstanden ist das Aussonderungssaxiom aus der Bemühung, das naive Komprehensionsprinzip in irgendeiner Form zu beschränken, da man mit diesem das Paradox von Russell ableiten kann.

Das naive Komprehensionsprinzip besagt, dass *jede* Elementbedingung, d.h. jede ‘Eigenschaft’ eine Menge bestimmt. Nehmen wir nun als Elementbedingung eine beliebige Formel $\varphi(x)$ in der Sprache der Mengenlehre, in der x frei vorkommt, so lautet es demnach:

$$\exists y \forall x (x \in y \leftrightarrow \varphi(x))$$

Noch deutlicher sieht man an der Second-Order-Variante worum es beim unbeschränkten Komprehensionsaxiom geht:

$$\forall F \exists y \forall x (x \in y \leftrightarrow Fx)$$

In Worten: Für jede Eigenschaft existiert die Menge aller Dinge, die diese Eigenschaft haben.²¹

Das Paradox von Russell (mengentheoretisch)

Das Paradox entsteht, wenn man via unbeschränktem Komprehensionsprinzip die Menge bildet, die aus allen Mengen besteht, die sich nicht selbst enthalten, i.e. die Menge $\{x : x \notin x\}$. Bezeichnet man diese Menge mit R , dann muss entweder $R \in R$ oder $R \notin R$ gelten. Falls ersteres der Fall ist, also $R \in R$, folgt nach Definition der Menge $R \notin R$, falls zweiteres der Fall ist, muss $\neg(R \notin R)$, also $R \in R$ gelten - in beiden Fällen ein Widerspruch.

Sehen wir uns an, wie die beiden Systeme das Paradox vermeiden:

NFU vermeidet es, weil es die Existenz einer Menge für eine Elementbedingung $\varphi(x)$ an x nur dann ‘garantiert’, wenn diese Bedingung stratifiziert ist, und die Formel $x \notin x$ offenbar nicht stratifiziert ist.

In ZFC gibt es zwar für alle z die Menge $\{x \in z : \varphi(x)\}$; Definiert man aber für beliebiges z $R_z =_{df} \{x \in z : x \notin x\}$, dann gilt: Falls $R_z \notin R_z$, dann ist $R_z \notin z$ oder $R_z \in R_z$. Das aber heißt einfach $R_z \notin z$. Hat man andererseits $R_z \in R_z$, dann ist nach Definition der Menge R_z $R_z \in z$ und $R_z \notin R_z$, was nicht möglich ist. Während in obigem Fall des unbeschränkten Komprehensionsprinzips beide Varianten zum Widerspruch geführt haben, bleibt uns für das Aussonderungsschema die Möglichkeit $R_z \notin R_z$, und wie bemerkt führt das einfach zu $R_z \notin z$ - und zwar für beliebiges z . Kein Widerspruch also.

²¹Oder in Frege’scher Terminologie: *Jeder Begriff hat einen Begriffsumfang.*

Die Allmenge

Die Allmenge V existiert in NFU, weil die Formel $x = x$ stratifiziert ist, und daher die Menge $V := \{x : x = x\}$ existiert.

Anders in ZFC: Das Aussonderungssaxiom verhindert gewissermaßen die Bildung von ‘zu großen’ Mengen²², insbesondere der Allmenge. Man kann das an folgendem Argument sehen:

Wie wir oben gezeigt haben, muss für beliebiges z $R_z \notin z$ sein (ansonsten droht, wie man oben sieht, das Russellparadox). Würde V existieren, so wäre $R_V \notin V$; das heißt, es gibt mindestens eine Menge, die nicht in V enthalten ist, was der Annahme widerspricht, V würde alles enthalten.

Anders formuliert: Angenommen es gäbe eine Menge $V := \{x : x = x\}$, dann gilt offenbar $\forall x (x \in V)$. Mit dem Aussonderungsschema kann man zeigen, dass die Menge $R_V := \{x \in V : x \notin x\}$ existiert. Weil $x \in V$ für alle x erfüllt ist, ist das aber einfach die Menge $R = \{x : x \notin x\}$, die defekte Russellklasse, die sofort zum Widerspruch führt.

In ZFC gibt es also keine Menge V , die alle Dinge umfasst, während die Existenz der Menge V in NFU qua S-Komprehension sozusagen die offensichtlichste Menge ist, die existiert.

Besonders beim Vergleich von verschiedenen Mengenlehren ist aber Vorsicht angebracht, wenn man von *der* Allmenge spricht.

In ZFC, wo es nur Mengen gibt, ist die ‘Allmenge’ V (die nicht existiert) die Menge (bzw. *Klasse*) aller Mengen. In NFU dagegen ist die Allmenge V die Menge aller Mengen *plus Urelemente* und V_{Set} (d.h. $\mathcal{P}(V)$) ist die Menge aller *Mengen*.

Dass in NFU die Allmenge V existiert, hat weitreichende Konsequenzen, die an dieser Stelle noch nicht abzuschätzen sind. Ein unbestreitbarer Vorteil ist die Tatsache, dass V mit Durchschnitt, Vereinigung und Komplement eine boolesche Algebra bildet, während in ZFC nur relative Komplemente (relativ zu einer existierenden Menge) tatsächlich ‘mengengroß’ sind (und damit ‘komplementfähig’).

Über Größe von Mengen, echte Klassen und dergleichen wird noch zu sprechen sein, man sollte aber auch hier schon im Auge behalten, dass, falls eine Menge in einem System, in dem V existiert, *nicht existiert*; das dann nicht daran liegen kann, dass diese Menge in irgendeinem Sinn ‘zu groß’ wäre. Mit anderen Worten: Wie kann eine Menge zu groß sein, wenn sogar die Menge aller Dinge existiert?

Trotzdem ist Vorsicht angebracht: Nur weil es in NFU viele ‘große’ Mengen gibt, die es in ZFC nicht gibt, heißt das nicht, dass *alle* Mengen existieren. Wie

²²Die Sprechweise ‘zu groß’, wenn man von Gesamtheiten spricht, die keine Mengen sind, lässt sich vollständig erst im Rahmen einer Klassen-Mengen-lehre wie NBG (Von Neumann-Bernays-Gödel) rechtfertigen. NBG ist eine konservative Erweiterung von ZFC, beweist also alle Sätze über *Mengen*, die auch in ZFC gelten. Zusätzlich erlaubt NBG aber auch Quantifikation über echte Klassen wie V und wir können dort neue Sätze über *Klassen* beweisen. Wir können dort etwa beweisen, dass *jede* echte Klasse (d.h. jede durch eine *beliebige* Eigenschaft bestimmte Gesamtheit) gleich groß wie das Universum V ist.

wir sehen werden, gibt es relativ einfache, ‘kleine’ Gesamtheiten, deren Existenz in NFU nicht bewiesen werden kann.

Ein kurze Bemerkung noch zur ‘Stärke’ des Aussonderungsschemas im Vergleich zum S-Komprehensionsschema. In einer bestimmten Hinsicht kann man das S-Komprehensionsschema mit gutem Recht ‘stärker’ nennen als das Aussonderungsschema. Es erlaubt uns nämlich die Ableitung *jeder stratifizierten Instanz des naiven Komprehensionsprinzips*. ZFC muss, weil es das naive Komprehensionsprinzip ganz aufgibt, *einige* Instanzen davon explizit als Axiome dazunehmen, um für Mengen zu sorgen, aus denen wir dann ‘aussondern’ können. Das S-Komprehensionsschema ist also ‘stärker’, weil wir uns etwa das Paarmentenaxiom oder das Potenzmengenaxiom von ZFC sparen, weil beides stratifizierte Fälle des unbeschränkten Komprehensionsprinzips sind.

Wie wir später aber noch sehen werden, bedeutet das nicht, dass das S-Komprehensionsschema in *jeder Hinsicht* ‘stärker’ ist als die Axiome der Aussonderung (d.h. nicht jeder Satz, der mit dem Aussonderungsschema bewiesen werden kann, kann auch mit dem S-Komprehensionsschema bewiesen werden). Auf eine - noch näher zu bestimmende Weise - sind sie sogar ‘unvergleichbar’, d.h. die beiden Schemata ‘wollen einfach verschiedene Dinge’.

4.4 Stratifizierung und heterogene Relationen

Ein wichtiger Unterschied zwischen ZFC und NFU betrifft die Zulässigkeit von bestimmten Relationen, nämlich den *heterogenen*.

Unter einer *heterogenen Relation* verstehen wir im Folgenden eine Relation zwischen Dingen von verschiedenem relativem Typ. Ein Beispiel für eine heterogene Relation wäre folgende Funktion f : Sei A die Menge $\{a, b\}$, dann ist die Potenzmenge $\mathcal{P}(A)$ die Menge $\{\emptyset, \{a\}, \{b\}, \{a, b\}\}$. Wir definieren $f : A \rightarrow \mathcal{P}(A)$ durch $f := \{(a, \{a\}), (b, \{b\})\}$, und erhalten für f die Funktion, die jedem Element aus A sein Singleton zuordnet.

Für ZFC ist die Bildung solcher Funktionen unproblematisch, weil es keinerlei Typenbeschränkungen gibt, d.h. $f(x) = \{x\}$ ist auf jeder Menge wohldefiniert. In NFU ist die Situation anders:

Nach unseren Stratifizierungsbedingungen von oben ist eine Formel, die eine Paarbedingung²³ enthält nur dann stratifiziert, wenn die beiden Dinge x und y des geordneten Paares (x, y) vom gleichen relativen Typ sind. Das soll nicht heißen, dass eine Formel, die eine nicht-stratifizierte Paarbedingung enthält, nicht existieren *kann*. Es heißt nur: Es *könnte* sein, dass eine Menge, die durch so eine Formel bestimmt ist, nicht existiert. Für jede endliche Funktion (jede Funktion, die auf einer endlichen Menge definiert ist) kann obige Funktion ohne Probleme definiert werden, indem einfach alle Paare ‘einzeln abgearbeitet’ werden. Weil die ‘Paarfunktion’ ein *totale* Funktion (im metasprachlichen Sinne von ‘Funktionszeichen in der Sprache von NFU’) ist, ist für alle ‘metasprachlichen Paare’

²³Unter einer Paarbedingung ist im Folgenden eine definierende Bedingung für eine Funktion oder Relation zu verstehen, d.h. etwas der Art $z \in f \leftrightarrow z = (x, f(x))$

(x, y) das geordnete Paar (x, y) (im mengentheoretischen Sinne) definiert, für beliebiges x existiert also das geordnete Paar $(x, \{x\})$.

Problematisch wird das aber für unendliche Mengen (z.B. den natürlichen Zahlen), die man nicht ‘einzeln abarbeiten’ kann. Für solche Mengen bräuchte man eine Eigenschaft, die dann ‘induktiv’ (im Falle der natürlichen Zahlen) auf alle Elemente dieser Menge ‘übertragen’ werden kann. Etwas der Art $x \in f \leftrightarrow x = (y, \{y\})$ oder allgemein $x \in f \leftrightarrow x = (y, f(y))$. Auf der rechten Seite dieser Äquivalenz muss eine Bedingung an x stehen, die die Funktion f *allgemein definiert*. So eine Bedingung bestimmt eine Funktion aber nur dann, wenn sie stratifiziert ist.

Auch Quine (der ‘Erfinder’ dieser Art Mengenlehre) war sich dieses Problems relativ bald bewusst (und zumindest *teilweise* der Konsequenzen) und spricht in Hinblick darauf auch von einem ‘Bann über heterogene Relationen’ als einem Relikt der Typentheorie Russells.²⁴

Ein konkretes Problem, das sich aus diesem Umstand ergibt ist die Unmöglichkeit einer ‘Standard-Definition’ des (unendlichen) Produktes von Mengen. In ZFC sieht das so aus:

Sei F eine Familie (i.e. eine Menge) von (möglicherweise unendlich vielen) nichtleeren Mengen. Das Produkt aller Mengen, die Element der Familie F sind, $\prod F$, definiert man als die Menge aller Funktionen f , die als domain F haben und deren Wert für jedes $A \in F$ ein Element $a \in A$ ist. Formal liest sich das so:

$$\prod F := \left\{ f : Funk(f) \wedge f : F \longrightarrow \bigcup F \wedge \forall A \in F (f(A) \in A) \right\}$$

Das Problem ist: Will man in NFU ein Objekt definieren, dann muss die definierende Bedingung stratifiziert sein, aber wegen $f(A) \in A$ kann obige Definition nicht stratifiziert sein!

Um dennoch eine analoge Definition zu erhalten, muss man versuchen, obige Formel zu stratifizieren. Das macht man, indem man die Funktionen von F nach $\mathcal{P}_1(\bigcup F)$ gehen lässt und statt $f(A) \in A$ fordert, dass $f(A) \subseteq A$. Die f ’s ordnen nun nicht mehr Mengen Elementen dieser Mengen zu, sondern Mengen Singletons dieser Mengen. Wir haben:

$$\prod^{NFU} F := \left\{ f : Funk(f) \wedge f : F \longrightarrow \mathcal{P}_1(\bigcup F) \wedge \forall A \in F (f(A) \subseteq A) \right\}$$

Die Bedingung $f(A) \subseteq A$ sagt, weil $f(A)$ ein Singleton ist, dass für ein Element $a \in \bigcup F$ gelten soll: $\{a\} \subseteq A$, also genau $a \in A$. Über die $\mathcal{P}_1$ -Operation kommen wir also über Umwege zum selben Ergebnis.²⁵

Indirekte Definitionen und Beweise dieser Art sind charakteristisch für NFU und verwandte Systeme der Mengenlehre. Immer wieder kommt es vor, dass man $\mathcal{P}_1$ - oder ähnliche Operationen braucht, um Aussagen zu stratifizieren.

²⁴Quine [1973] S. 212

²⁵Mehr zu diesem Thema erfährt man bei Holmes [1998] S. 74 ff.

4.5 Fundierungsaxiom contra $V \in V$

Historisch wirkt es bisweilen beinahe ironisch: Gerade in NFU, das aus der Typentheorie Russells entstanden ist, die Mengen, die sich selbst enthalten, als ‘ungrammatisch’ abtut, gibt es Mengen, für die genau *das* gilt. Andererseits gilt gerade das Fundierungsaxiom, das in der ersten Axiomatisierung von Zermelo noch gar nicht dabei war, und erst später von Von Neumann hinzugefügt wurde, als fundamentales Prinzip dieser Art Mengenlehre.

Der iterative Begriff der Menge

Wie man in Abschnitt 3 sehen konnte, schließt das Fundierungsaxiom (kurz: *Fund*) verschiedene Mengen aus. Dazu gehören zirkelhafte Mengensysteme, also Mengen für die gilt: $x_1 \in x_2 \in \dots \in x_n \in x_1$ und damit insbesondere Mengen, die sich selbst enthalten. Ganz allgemein verbietet *Fund* unendlich lange absteigende ‘ $\in$ -Folgen’.²⁶

Allgemein muss ja eine Relation auf einer Menge kein kleinstes Element haben. Für die ganzen Zahlen mit der üblichen Kleiner-Relation gilt das etwa nicht.

Das Fundierungsaxiom besagt nun, dass der Fall, dass eine Menge *kein* ‘ $\in$ -minimales’ Element hat, *nicht* eintreten kann. *Jede* Menge x besitzt ein Element y , für das gilt: $x \cap y = \emptyset$, d.h. es gibt eine Menge y in x derart, dass es kein $z \in x$ gibt, das auch in y wäre. In dem skizzierten Sinne kann man das Fundierungsaxiom ein *restriktives* Axiom nennen, denn es beschränkt das Universum V auf die Klasse der fundierten Mengen WF .

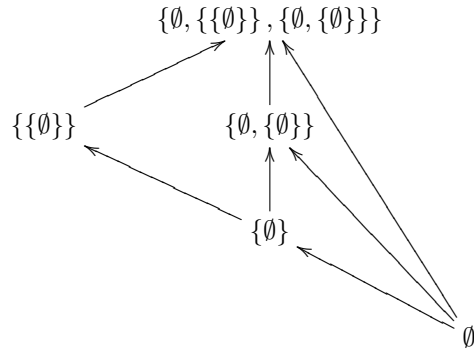
Wir wollen aber sehen, was für Effekte *Fund* außerdem noch hat und stellen uns die Frage: *Wie kommen wir eigentlich zu Mengen?*

Die einzige Menge, die uns in ZFC von Anfang an gegeben ist, ist die leere Menge $\emptyset$.²⁷ Die Axiome von ZFC erlauben uns nun die schrittweise Bildung von weiteren Mengen. Wir können etwa die Menge $\{\emptyset\}$ bilden oder $\{\{\emptyset\}\}$ oder $\{\emptyset, \{\emptyset\}\}$ als Vereinigung der beiden vorigen.

Betrachten wir als Beispiel die Menge $\{\emptyset, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$ und stellen wir uns den Aufbau dieser Menge wie einen Baum vor:

²⁶Ein Beispiel für eine ‘ $\in$ -Folge’ wäre etwa die Folge der natürlichen Von-Neumann Zahlen $0 \in 1 \in 2 \in 3 \in \dots$, die wir im Kapitel über Arithmetik in ZFC kennen lernen werden.

²⁷Alle anderen Axiome geben uns nur *weitere Mengen* unter der Annahme von schon *Bestehenden* oder sie ‘regeln’ einfach nur, wie sich die $\in$ -Relation verhält (Extensionalität und Fundierung).



Die Pfeile in diesem Diagramm deuten dabei die ‘Richtung’ der $\in$ -Relation an. Man sieht, dass die ‘Wurzel’ des Baums die leere Menge ist. Weil es in ZFC keine Atome gibt ist sie aber auch schon die einzige Wurzel.

Am Diagramm sieht man auch, dass jede Menge auf einer bestimmten Stufe gebildet wurde. Indizieren wir diese Stufen mit natürlichen Zahlen (bzw. allgemein *Ordinalzahlen*), so sehen wir, dass z. B. die Menge $\{\{\emptyset\}\}$ genauso wie die Menge $\{\emptyset, \{\emptyset\}\}$ auf der zweiten Stufe gebildet wurde.

Man kann dieses Bild verallgemeinern: Wir nehmen als 0. Stufe die leere Menge. Als 1. Stufe bezeichnen wir die Potenzmenge der leeren Menge, also $\{\emptyset\}$, die 2. Stufe soll die Potenzmenge der 1. sein, also die Menge $\{\emptyset, \{\emptyset\}\}$ usw.

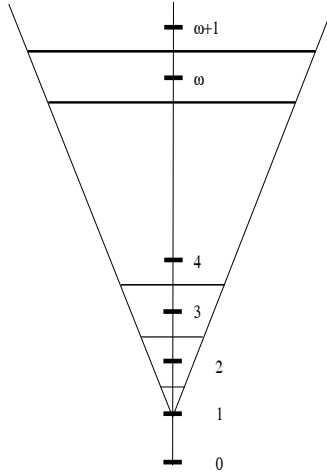
Formal definieren wir dabei über transfinite Rekursion die sogenannten Von Neumannschen Stufen wie folgt:

- $R_0 := \emptyset$
- $R_{\alpha+1} := \mathcal{P}(R_\alpha)$ (für Nachfolgerordinalzahlen)
- $R_\beta := \bigcup \{R_\gamma : \gamma < \beta\}$ (für Limesordinalzahlen)

Das Fundierungsaxiom besagt nun, dass unser Universum V durch die Von Neumannschen Stufen sozusagen ‘ausgeschöpft’ wird, d.h. dass jede Menge, die überhaupt in ZFC gebildet werden kann, in irgendeinem R_α vorkommt. Für eine beliebige Menge x nennt man das minimale α für das gilt $x \in R_{\alpha+1}$ den Rang von x , kurz $Rg(x)$.

Das Fundierungsaxiom besagt demnach, dass jede Menge einen eindeutig bestimmten Rang hat. Würde es unendlich absteigende $\in$ -Ketten geben, Mengen, die kein $\in$ -minimales Element haben, so könnten diese Mengen nicht auf einer bestimmten Stufe gebildet worden sein, also auch keinen eindeutig bestimmten Rang haben. Auch für zwei Mengen x, y mit $x \in y$ und $y \in x$ kann es keinen Rang geben, denn mit $x \in y$ legt man sich darauf fest, dass x in der linearen Von Neumannschen Hierarchie *vor* y gebildet wurde, mit $y \in x$ auf das Gegenteil.

Das Bild, das sich daraus ergibt, die sogenannte *Von Neumann’sche Hierarchie*, sieht so aus:



Man sieht auch, dass die Mengen-Hierarchie in ZFC *kumulativ* ist (im Gegensatz etwa zur Hierarchie der einfachen Typentheorie Russells). Jede Stufe enthält als Elemente und als Teilmengen alle vorherigen Stufen, denn es gilt für alle Ordinalzahlen α, β : $\beta < \alpha \rightarrow R_\beta \in R_\alpha$ und $\beta \leq \alpha \rightarrow R_\beta \subseteq R_\alpha$. Den durch dieses Bild repräsentierten Begriff der Menge nennt man auch den *iterativen Begriff der Menge*.

Anti-Fundiert?!

Wie sieht es demgegenüber in NFU aus? Eine lineare Hierarchie wie die Von-Neumannsche kann es nicht geben, denn mit $V \in V$ hat man sich darauf festgelegt, dass es zumindest *einen* Fall gibt, wo die $\in$ -Relation *nicht* irreflexiv ist. Macht man sich ein Pfeildiagramm wie oben, gibt es demnach Pfeile, die nicht bloß in eine Richtung zeigen.

Nimmt man für $x_1, x_2, \dots, x_n$ jeweils V , so sieht man auch, dass es Mengen gibt, für die $x_1 \in x_2 \in \dots \in x_n \in x_1$.²⁸

Es ist nicht ganz leicht, sich vorzustellen, wie solche Mengen ‘konkret’ aussehen könnten, ohne den Begriff ‘Menge’ und die $\in$ -Relation in einer Weise zu gebrauchen, die zumindest gewöhnungsbedürftig ist. Anschaulich besagt die Tatsache, dass das Fundierungsaxiom für NFU falsch ist nämlich, dass das Universum gerade kein ‘Baum’ ist, dass es sogar möglich ist - falls es außer V noch

²⁸Ein ‘anspruchsvolleres’ Beispiel würde sich ergeben, wenn wir statt des primitiven geordneten Paares das Paar von Kuratowski benutzt hätten. Wie wir später sehen werden, ist der Ordnungstyp der Menge aller Ordinalzahlen mit der üblichen Ordnung eine Ordinalzahl. Bezeichnen wir diese Ordinalzahl mit Ω , so gilt für eine weitere ‘größere’ Ordinalzahl α : $\Omega \in \{\Omega\} \in \{\{\Omega\}, \{\Omega, \alpha\}\} \in \dots \in \Omega$. Da wir in der hier präsentierten Form von NFU ein primitives Paar benutzen ist das kein ‘wirkliches’ Beispiel; dennoch bleibt ein ‘Beigeschmack’ von Zirkularität.

weitere zirkuläre Mengen gibt (und es gibt keinen *prima facie* Grund, dies *nicht* anzunehmen) - dass bestimmte Äste wieder zusammenwachsen. Das graphische Pendant zum Baum wäre also eher ein *Netz* oder ein Graph, der Zyklen enthält.

Es ist bezeichnend und ein Hinweis auf die verschiedenen Intuitionen, die hinter NFU und ZFC stecken, welches Verständnis Leute, die sich mit dem Fundierungsaxiom in 'ablehnender' Weise beschäftigen von Logik und Mengenlehre haben. Dazu ein Zitat von Jon Barwise im Vorwort zu einem Buch von Peter Aczel:

"[...] by reaching out and studying new phenomena, applied mathematics in general, and mathematical logic in particular, enriches mathematics, not only with new theorems, but also with new mathematical structures, structures for the mathematician to study and for others to apply in new domains.

The theory of circular and otherwise extra-ordinary sets is an excellent example for this synergetic process."²⁹

Die Sichtweise, dass man das Fundierungsaxiom nicht für 'wesentlich' für den Begriff der Menge hält, passt gut in dieses 'syntaktizistische' (mit Shapiro könnte man auch sagen: das 'modelltheoretische') Bild aus obigem Zitat. Mit 'syntaktizistischem' Bild ist dabei folgendes gemeint: Was wir tun, wenn wir Axiome aufstellen, ist nicht den Begriff der Menge und der Elementrelation in eine axiomatische Form bringen; wir 'explizieren' den Begriff nicht; vielmehr lassen wir uns gewissermaßen von den Zeichen leiten; das heißt wir sehen uns dabei an, wie sich das Zeichen $\in$ unter der Annahme von gewissen Axiomen eben verhält; und wenn dabei herauskommt, dass sich diese Relation (*als Zeichen*) eben *nicht* so verhält, wie wir die Relation 'in einer Menge Enthalten-Sein' 'umgangssprachlich', oder besser 'vor-theoretisch' verstehen (d.h. wir finden z.B. Modelle von ZFC^- , in dem das Fundierungsaxiom *nicht* gilt), so zeigt das nur, dass Fundiertheit eben nur 'wesentlich' für einen bestimmten Begriff der Menge, nämlich den iterativen, ist, und nicht den Begriff der Menge *schlechthin*.

Die Sichtweise, dass es einen kohärenten Begriff der Menge tatsächlich gibt, und dass es der 'iterative Begriff' ist, könnte man im Gegenzug die *ZFC-platonistische* nennen. Problematisch ist diese Sichtweise insofern, als nicht klar ist, was *genau* diesen Begriff nun ausmacht. Ist er eine sublimere Variante des 'umgangssprachlichen' Mengenbegriffs, oder ist damit der Begriff Cantors oder der Dedekinds gemeint?

In demselben Buch schreibt Aczel über nicht-fundierte Mengen (Mirimanoff nennt sie 'extra-ordinary'):

"What is extra-ordinary about such a set is that it would seem that it could never get formed; for in order to form the set we would first have its elements, and to form those elements we would have

²⁹ Aczel [1988] xi

to have previously formed their elements and so on leading to an infinite regress.”³⁰

Aczel wendet sich im Weiteren direkt gegen die ‘starke’ Sichtweise, dass es nur einen Begriff der Menge gäbe, nämlich den iterativen;³¹ die Vorstellung, dass man unter Mengen nur solche Dinge verstehen soll, die man aus vorher gegebenen Objekten irgendwie zu einer Einheit zusammenfasst, nennt er sogar ‘anthropomorph’.

In dieser Sichtweise ist der iterative Begriff der Menge einfach ein Spezialfall eines *allgemeineren* Mengenbegriffs, der eben auch nichtfundierte Mengen umfasst.

Es wird an dieser Stelle noch nicht näher auf diesen Punkt eingegangen und auf das letzte Kapitel verwiesen, das den Zusammenhang zwischen den verschiedenen Intuitionen - oder ‘Mengenbegriffen’ - die NFU und ZFC trennen, genauer klären soll. Aber eines ist klar: Nur weil sich Aczel gegen den ‘iterativen Begriff’ wendet, heißt das nicht, dass er einen Begriff der Menge vertritt, wie er durch NFU oder verwandte Systeme verkörpert wird. Martialisch ausgedrückt gilt also *nicht*: Der Feind meines Feindes ist mein Freund.

Das Zitat soll nur belegen, dass es auch ZFC-affine Leute gibt, die den ‘iterativen Begriff’ nicht für *den* Begriff der Menge halten, oder auch nur glauben, dass es *den* Begriff der Menge überhaupt gibt. Was diese Leute mit Verfechten anderer axiomatischer Mengenlehren wie NFU verbindet, ist vor allem ihre ‘liberale’ Sichtweise (ihr Relativismus) bzgl. verschiedener Mengenbegriffe.

4.6 Mengen und Klassen

Manchmal ist es in ZFC günstig, so zu tun, als habe man es z. B. bei V oder ON (der Klasse aller Ordinalzahlen) mit ‘realen’ Objekten zu tun. Reale Objekte sind in dieser Sprechweise einfach die Dinge - wir nennen sie Mengen - über die quantifiziert wird. Wir tun das in ZFC aber nur deshalb, weil wir uns davon einen ‘notationalen Vorteil’ erhoffen. So kann man für zwei Klassen C, N Sachen schreiben wie: $C \subseteq N$ als Abkürzung für die ausführliche und eigentlich korrekte Schreibweise: $\forall x(C(x) \rightarrow N(x))$. Das, was man im Rahmen von ZFC unter Klassen versteht, sind also in Wahrheit einfach Prädikate, *Formeln* in der Sprache $\{\in\}$ mit freien Variablen.

Manche dieser Prädikate bestimmen Mengen, manche nicht. Und solche, die keine Mengen bestimmen, nennt man *echte Klassen* (oder einfach *Klassen*). In ZFC können wir z. B. zeigen, dass es eine Menge gibt, die alle natürlichen Zahlen enthält. Wir bezeichnen das Prädikat, eine natürliche Zahl zu sein mit ‘ $Nat(x)$ ’ und können zeigen, dass die Menge $\{x : Nat(x)\}$ in unserem Mengenuniversum existiert. Andererseits können wir für das Prädikat, eine Ordinalzahl zu sein, kurz $On(x)$, zeigen dass die Menge $\{x : On(x)\}$ in ZFC nicht existiert, und d.h.:

³⁰Aczel [1988] xvii

³¹Die ‘schwache’ Sichtweise ist bei Aczel jene, in der man nichtfundierte Mengen deshalb ausschließt, weil sie zur ‘Grundlegung der Mathematik’ nicht gebraucht werden.

nicht Wert einer quantifizierbaren Variable ist, also nichts anderes als: *ON ist eine echte Klasse*. In ZFC fallen Existenz und die Eigenschaft ‘Menge zu sein’, zusammen, weil es nur eine Art von Objekten, nämlich Mengen gibt, über die quantifiziert wird. Es ist also nicht so, dass es Klassen gibt und Mengen, und bestimmte Klassen sind Mengen, und andere Klassen, die keine Mengen sind, sind echte Klassen. Der Begriff Klasse kommt im offiziellen Sprachgebrauch dieser Mengenlehre einfach nicht vor. Immer wenn man von Klassen spricht, spricht man eigentlich über Prädikate. Die Rede von Klassen in ZFC ist eine bloße *façon de parler*, die einem das Leben leichter machen soll.

Der Zusammenhang zwischen Prädikaten und Mengen ist in NFU - wie in ZFC - der, dass manchmal Prädikate Mengen bestimmen und manchmal nicht. Der Unterschied ist, dass man - axiomatisch qua S-Komprehension - von einer *bestimmten* Art von Prädikaten, nämlich den stratifizierten, fordert, dass sie *immer* Mengen definieren!

Dennoch bleibt der Rest der nichtstratifizierten Prädikate, von denen *prima facie* noch nicht klar ist, ob sie Mengen bestimmen oder nicht und es bleibt einem anheim gestellt, nicht-stratifizierte Prädikate keine Mengen bestimmen zu lassen, falls die Existenz dieser Mengen zu Widersprüchen führen würde.

Mengen, deren Existenz in NFU zu Widersprüchen führen würden kann man nun, analog zu ZFC, auch *Klassen* oder *echte Klassen* nennen.³² Der Unterschied ist, dass - anders als in ZFC - die Tatsache, dass es sich bei solchen Vielheiten um Klassen (im Gegensatz zu ‘real existierenden’ Mengen) handelt, nichts damit zu tun hat, dass sie ‘zu groß’ wären. In NFU existiert ja sogar die größte aller Mengen, V ; während eine relativ kleine Vielheit wie $\{\Omega, T(\Omega), T^2(\Omega), \dots\}$ - wie wir noch sehen werden - nicht existiert³³, obwohl sie - allem Anschein nach - sogar abzählbar ist (oder zumindest *wäre*, wenn sie existieren würde).

Einen wesentlichen Unterschied zwischen dem Aussonderungs- und dem S-Komprehensionsschema kann man in der Sprache der Klassen (als Prädikaten) kompakt ausdrücken: Das Aussonderungsschema besagt, dass jede Teilklasse einer Menge eine Menge ist. Dass das in NFU nicht gilt, sieht man an der *Teilklasse* $\{\Omega, T(\Omega), T^2(\Omega), \dots\}$ der *Menge* aller Ordinalzahlen, die aber *keine Menge* ist. In dieser Hinsicht kann man das Aussonderungsschema ‘stärker’ nennen als das S-Komprehensionsschema.

Man merkt hier - und das verweist auch schon auf *des Pudels Kern* - worauf die beiden Systeme ihre Schwerpunkte legen: ZFC lässt alle möglichen Prädikate Mengen bestimmen, sofern nur aus Objekten ‘gewonnen’, die nicht zu groß sind, d.h. die aus gegebenen Mengen ausgesondert werden können.

NFU dagegen verlagert die Betonung auf die Seite der Prädikate; es werden keine ‘Bedingungen’ an *Mengen* gestellt: Damit ein Prädikat eine Menge bestimmt, ist das einzige, das verlangt wird, dass das *Prädikat* stratifiziert ist.

³²Vopenka und Hajek nennen solche Vielheiten auch ‘semisets’. vgl. Forster [1992] S. 30 mit Fussnote

³³ $\Omega, T(\Omega), T^2(\Omega), \dots$ sind alles Ordinalzahlen.

5 Kardinal- und Ordinalzahlen in NFU

Ein grundlegender Aspekt von Mengen ist deren Größe. Die Theorie der Kardinalzahlen basiert in NFU auf der Intuition, dass alle Mengen mit einer bestimmten Anzahl von Elementen etwas ‘gemeinsam’ haben; nämlich die gleiche Anzahl von Elementen zu haben. Für die natürlichen Zahlen z.B. fasst man alle Mengen mit n Elementen zusammen und nennt die daraus resultierende Menge die Kardinalzahl n .³⁴

Ein anderer grundlegender Aspekt ist der des Zählens. Wichtig dabei ist nicht bloß die Größe (im Sinne der ‘Kardinalität’ einer Menge), sondern auch die Reihenfolge. Aus der Formalisierung dieses Aspekts entsteht der Begriff der Ordinalzahl. Grundlage für jede Theorie der Ordinalzahlen ist die Idee der *Wohlordnung*.

5.1 Kardinalzahlen: Grundlegende Begriffe

Bevor wir uns der Definition der Kardinalität und der Kardinalzahlen zuwenden können, noch ein paar Definitionen:

Definition 5.1. Eine Menge A heißt *höchstens so mächtig* wie eine Menge B , in Zeichen $A \preceq B$, wenn es eine injektive Funktion von A nach B gibt.³⁵

Eine Menge A heißt zu einer Menge B *gleichmächtig*, $A \approx B$, wenn es eine bijektive Funktion von A nach B gibt.

Eine Menge A heißt *schmächtiger* als eine Menge B , $A \prec B$, wenn $A \preceq B$, aber nicht gleichmächtig, ist.

Satz 5.1. 1. $A \preceq A$

2. $A \preceq B \wedge B \preceq C \rightarrow A \preceq C$

3. $A \preceq B \wedge B \preceq A \rightarrow A \approx B$

4. $A \preceq B \vee B \preceq A$

Um 1 und 2 einzusehen betrachte man die identische Funktion id_A auf A bzw. die Komposition von injektiven Funktionen $g : A \rightarrow B$ und $f : B \rightarrow C$.

Punkt 3 im Satz ist auch unter dem Titel ‘Satz von Cantor-Schröder-Bernstein’ bekannt und kann relativ *basic* geführt werden. 4 dagegen macht wesentlich Gebrauch vom Auswahlaxiom in der Variante des Wohlordnungssatzes.

Satz 4.1. besagt insgesamt, dass $\preceq$ eine (schwache) Totalordnung auf der Menge aller Mengen ist (mit ‘ $\approx$ ’ statt ‘=’).

³⁴Das klingt auf den ersten Blick zirkulär. Dieser Anschein sollte sich aber bald verflüchtigen, wenn man sieht, was bei der Definition tatsächlich passiert.

³⁵Man beachte, dass es - unter der Annahme des Auswahlaxioms - eine injektive Funktion von A nach B genau dann gibt, wenn es eine surjektive Funktion von B nach A gibt.

Die Gleichmächtigkeitsrelation ist eine Äquivalenzrelation auf der Menge aller Mengen, denn es gilt:

- Satz 5.2.** 1. $A \approx A$
 2. $A \approx B \wedge B \approx C \rightarrow A \approx C$
 3. $A \approx B \rightarrow B \approx A$

Für 1 und 2 kann man wieder die identische Funktion und die Komposition nehmen. Auch 3 ist klar, denn, falls $f : A \rightarrow B$ bijektiv ist, dann auch $f^{-1} : B \rightarrow A$.

Wir kommen nun zu einer der wichtigsten Definitionen in diesem Kapitel, der *Kardinalität* oder *Mächtigkeit* einer Menge. Bisher haben wir nur *relational* (über Funktionen) definiert, wann zwei Mengen gleichmächtig oder schwächer sind. Nun sind wir in der Lage, eine Definition der *Kardinalität* (oder *Kardinalzahl*) einer Menge zu geben:

Definition 5.2. Die *Kardinalität* $|A|$ einer Menge A ist die Menge aller Mengen B , die zu A gleichmächtig sind, i. e.
 $|A| := \{B : B \approx A\}$

Die Eigenschaft eine Kardinalzahl zu sein, wird also definiert durch $Kard(x) :\leftrightarrow \exists A(x = |A|)$

Da $\approx$ eine Äquivalenzrelation ist, definiert $\approx$ eine Partition auf der Menge aller Mengen. Definition 4.2. besagt also gerade, dass Kardinalzahlen Äquivalenzklassen unter der Beziehung der Gleichmächtigkeit sind. Wir verwenden in Zukunft kleine griechische Buchstaben $\kappa, \lambda, \mu, \dots$ (wenn nötig mit Indizes) für Mengen, die Kardinalzahlen sind und wir bezeichnen die nach S-Komprehension existierende Menge aller Kardinalzahlen als KARD.

Weiters definieren wir für Kardinalzahlen λ und κ , und für $A \in \lambda, B \in \kappa$

Definition 5.3. $\lambda \leq \kappa :\leftrightarrow A \preceq B$

Mit Satz 4.1. folgt, dass $\leq$ eine Totalordnung auf der Menge aller Kardinalzahlen definiert. Mehr noch. Es gilt:

Satz 5.3. $WO(\leq_{CARD})$, i.e.: Die *Kleiner-gleich-Relation* auf der Menge aller Kardinalzahlen ist eine Wohlordnung.

Exkurs: Die natürlichen Kardinalzahlen

Um die Menge aller natürlichen Zahlen definieren zu können, müssen wir noch einige Vorkehrungen treffen. Unser Ziel ist es, gemäß obiger Definition der Kardinalzahl einer Menge, die natürlichen Zahlen als Äquivalenzklassen von Mengen von endlicher Mächtigkeit aufzufassen, i.e. n ist die Menge aller Mengen mit Kardinalität n .

Jede beliebige (aber bestimmte) natürliche Zahl lässt sich ‘rein logisch’ durch

$$n := \{A : A = \{a_1, a_2, \dots, a_n\} \wedge a_i \neq a_j \text{ für } i \neq j\}$$

definieren.³⁶ Will man aber die Menge *aller* natürlichen Zahlen definieren, braucht man dafür eine Nachfolgerfunktion, die wir mit S bezeichnen. Davor noch eine allgemeinere

Definition 5.4. $A \oplus 1 := \{a \cup \{x\} : a \in A \wedge x \notin a\}$ falls es so ein x gibt (sonst $\emptyset$)

Die natürliche Zahl 0 definieren wir als $\{\emptyset\}$ und den Nachfolger $S(n)$ einer natürlichen Zahl als

Definition 5.5. $S(n) := n \oplus 1$

Insbesondere ist $S(0) = 1 = \mathcal{P}_1(V)$.

Die Menge $\mathcal{N}$ der natürlichen Zahlen können wir nun definieren durch

Definition 5.6. $\mathcal{N} := \bigcap Ind$

wobei Ind für die Menge aller *induktiven Mengen* steht, d.h. derjenigen Mengen, die

1. 0 enthalten und
2. bezüglich der Funktion $A \mapsto A \oplus 1$ abgeschlossen sind

$\mathcal{N}$ ist also die kleinste Menge, die 0 enthält und mit jedem n auch $S(n)$. Aus diesem Umstand erhalten wir

Satz 5.4. (Satz von der vollständigen Induktion - Eigenschaftsform)

Wenn $\varphi(x)$ ein stratifizierte Formel mit x frei ist, dann gilt: Wenn $\varphi(0)$ und für alle natürlichen Zahlen n folgt aus $\varphi(n)$ stets $\varphi(S(n))$, dann gilt $\forall n \varphi(n)$.

In der hier präsentierten Form besagt das Induktionsprinzip: Falls eine beliebige Eigenschaft auf 0 zutrifft und für jede beliebige natürliche Zahl n kann man zeigen, dass, falls die Eigenschaft auf n zutrifft so trifft sie auch auf $S(n)$ zu, so kann man schließen, dass diese Eigenschaft auf *alle* natürlichen Zahlen zutrifft; die Eigenschaft vererbt sich *induktiv* von 0 auf alle Nachfolger, i.e. auf ganz $\mathcal{N}$.

Wesentlich für diese Form des Induktionsprinzips ist die einschränkende Bedingung ‘falls $\varphi(x)$ stratifiziert’. Im Vergleich mit der *Mengenform* des obigen Satzes kann man auch sehen, wieso das so ist.

³⁶Z.B. ist die natürliche Kardinalzahl 2 die Menge aller Paarmengen, i.e. $2 := \{A : \exists a \exists b (a \neq b \wedge \forall x (x \in A \leftrightarrow x = a \vee x = b))\}$

Satz 5.5. (*Satz von der vollständigen Induktion - Mengenform*)

Wenn S eine Menge von natürlichen Zahlen ist, dann gilt: Wenn $0 \in S$ und für alle n folgt aus $n \in S$, dass auch $S(n) \in S$ ist, so ist $S = \mathcal{N}$ ³⁷

Der Grund für die Adaption des S-Komprehensionsschemas war ja, dass wir Dinge mit einer bestimmten Eigenschaft zu einer Menge zusammenfassen wollten. Das S-Komprehensionsschema erlaubt dies aber nur für stratifizierte ‘Eigenschaften’ (i.e. Formeln). Lesen wir nun in der Analogie von Elementbeziehung und Prädikation $n \in S$ als $\varphi(n)$, so bedeutet das, da das S in der Mengenform ja eine abquantifizierte Variable ist (d.h. über ‘existierende Mengen’ läuft), dass wir für $\varphi(n)$ nur stratifizierte Formeln verwenden dürfen, weil nur stratifizierte Formeln mit Sicherheit ‘existierende Mengen’ bestimmen.

Gerade dasjenige Axiom (S-Komprehension), das für ‘möglichst große Freiheit’ bei der Mengenbildung sorgen soll (S-Komprehension erlaubt die Bildung von V , ORD , $KARD$), schränkt uns also schon bei so ‘einfachen’ Mengen wie $\mathcal{N}$ (vorerst) gewaltig ein, denn wie wir sehen werden, gibt es recht einfache (nicht-stratifizierte) Sätze, die nicht beweisbar sind, von denen wir aber *gern hätten*, dass sie beweisbar *wären*.

Mit $A \oplus B := A \times \{0\} \cup B \times \{1\}$ kann man nun *Addition* und *Multiplikation* auf den natürlichen Zahlen wie folgt definieren:

Definition 5.7. Seien $n, m \in \mathcal{N}$ und $A \in n$ und $B \in m$, dann ist

1. $n + m := \iota k(A \oplus B \in k)$ ³⁸

2. $n \cdot m := \iota k(A \times B \in k)$

Man kann nun überprüfen, dass die derart definierten natürlichen Zahlen mit den darauf definierten Operationen die Peano-Axiome³⁹ erfüllen, d.h. es gelten:

1. $0 \in \mathcal{N}$
2. $n \in \mathcal{N} \rightarrow n + 1 \in \mathcal{N}$
3. $n \in \mathcal{N} \rightarrow n + 1 \neq 0$
4. $n \in \mathcal{N} \wedge n \neq 0 \rightarrow \exists m(n = m + 1)$
5. $n + 1 = m + 1 \rightarrow n = m$
6. $n + 0 = n$

³⁷Man beachte, dass, anders als in ZFC das Prinzip der vollständigen Induktion ein unmittelbares Korollar aus der Definition von $\mathcal{N}$ ist. Gelten nämlich für $S \subseteq \mathcal{N}$, dass $0 \in S$ und $\forall n(n \in S \rightarrow S(n) \in S)$, so ist S offenbar induktiv. Da aber $\mathcal{N}$ die kleinste induktive Menge ist, muss auch $\mathcal{N} \subseteq S$, d.h. $S = \mathcal{N}$ gelten.

³⁸Die Definition von $n \oplus 1$ stimmt mit der von $n + 1$ überein, d.h. $n + 1 = n \oplus 1 = S(n)$.

³⁹Oft meint man mit den Peano-Axiomen nur die Axiome 1 bis 5 zusammen mit dem Induktionsprinzip. 2 und 4 werden oft durch ‘Jede natürliche Zahl hat einen Nachfolger, der ebenfalls eine natürliche Zahl ist’ zusammengefasst.

$$7. m + (n + 1) = (m + n) + 1$$

$$8. m \cdot 0 = 0$$

$$9. m \cdot (n + 1) = m \cdot n + m$$

Das Modell $(\mathcal{N}, +, \cdot)$, das wir von den natürlichen Zahlen gegeben haben, leistet also tatsächlich, was wir von ihm erwarten.

Ein Problem, das auftauchen kann, ist, dass es möglicherweise ‘nicht genügend’ Dinge im Universum gibt. Wenn das so ist, dann kann es passieren, dass es kein x mehr gibt, sodass $x \notin A$ für $A \in n$, weil schon alle Dinge in A ‘drinnen’ sind, d.h. $A = V$. Dann aber wäre $S(n) = S(S(n)) = \dots = \emptyset$. Dass dieser Fall nicht eintreten kann, ergibt sich aus folgendem

Satz 5.6. (*Unendlichkeitstheorem*)

$\emptyset$ ist keine natürliche Zahl.

Dass das Universum unendlich viele Dinge enthält ist also eine Konsequenz aus den Axiomen und braucht nicht extra als Axiom hinzugefügt werden. Das liegt im wesentlichen daran, dass wir den Axiomen ein primitives type-level geordnetes Paar zugrunde gelegt haben, das uns die Existenz von unendlich vielen Dingen für jeden ‘Typ’ garantiert. Ohne dieses spezielle *feature* kann man zwar immer noch die Existenz potentiell unendlich vieler Dinge garantieren, kann diese aus Stratifikationsgründen aber nicht in ‘eine Menge packen’.

Axiom of Counting

Ein Problem, das oben angedeutet wurde, ergibt sich, wenn man sich folgendes überlegt: Angenommen wir haben eine Menge $A \in n$. Intuitiv würden wir erwarten, dass, da A eine n -elementige Menge ist, auch die Menge aller Singletons von Elementen von A , $\mathcal{P}_1(A)$ eine n -elementige Menge ist, und daher auch $\mathcal{P}_1(A) \in n$. Man hätte also gerne: $\forall A(A \in n \rightarrow \mathcal{P}_1(A) \in n)$. Das Problem ist: Versucht man das (durch vollständige Induktion) zu beweisen, dann sieht man, dass das nicht möglich ist, weil die Bedingung $\forall A(A \in n \rightarrow \mathcal{P}_1(A) \in n)$ nicht stratifiziert ist, und Induktion nur über stratifizierte Bedingungen möglich ist.

Ähnlich ergeht es uns mit der Menge $\{1, \dots, n\}$, die offensichtlich n Elemente hat, von der wir aber nicht zeigen können, dass sie Element von n ist.

Um solche Unschönheiten zu vermeiden, definiert man zuerst

Definition 5.8. Für jede natürliche Zahl n und $A \in n$ sei

$$1. T(n) := \iota m(\mathcal{P}_1(A) \in m) = |\mathcal{P}_1(A)|$$

$$2. T^{-1}(n) = \iota m(T(m) = n)$$

Die Abbildung $n \mapsto T(n)$ definiert eine Operation auf den natürlichen Zahlen, und bildet natürliche Zahlen auf natürliche Zahlen ab. T ist aber keine *Funktion* (im Sinne einer Menge von geordneten Paaren), denn weil $T(n)$ alle $\mathcal{P}_1(A)$ für $A \in n$ enthält, ist $T(n)$ von einem höheren relativen Typ als n , und die definierende Paarbedingung $x \in T \leftrightarrow x = (n, T(n))$ deshalb nicht stratifiziert.

Um die Probleme von oben zu beseitigen fordert man nun *per Axiom*, dass für alle n gilt:

- $T(n) = n$ (Axiom of Counting)

Man sieht sofort, dass das erste der oben erwähnten Probleme durch dieses Axiom gelöst wird, weil nun mit $A \in n = T(n)$ auch $\mathcal{P}_1(A) \in T(n) = n$ ist, und daher $A \in n$ genau dann der Fall ist, wenn $\mathcal{P}_1(A) \in n$. Auch das zweite Problem löst sich in Luft auf. Außerdem erhält man folgendes

Theorem. (*Restricted Subversion Theorem*) Falls $n \in \mathcal{N}$ und $\varphi(n)$ eine Formel in der Sprache von NFU mit n frei ist (nicht notwendigerweise stratifiziert), dann existiert die Menge $\{n : \varphi(n)\}$.

Das heißt, wenn eine Variable auf die natürlichen Zahlen beschränkt ist, braucht man sich über etwaige ‘Typenbeschränkungen’ keine Sorgen machen, weil uns die T-Operation ein freies *shiften* hinauf oder hinab erlaubt. Die Einschränkungen, von denen im Zusammenhang mit der vollständigen Induktion die Rede war, werden durch das Axiom of Counting gewissermaßen ‘zurückgenommen’. Induktion ist nun auch über nicht-stratifizierte Bedingungen erlaubt.

Soviel zum Thema natürliche (Kardinal-)Zahlen. Bevor wir uns dem Thema Kardinalzahlen weiter widmen können, wollen wir uns die grundlegenden Begriffe zur Theorie der Ordinalzahlen ansehen.

5.2 Ordinalzahlen: Grundlegende Begriffe

Intuitiv stellt man sich Ordinalzahlen als *Maßzahlen* für die *Länge* von Wohlordnungen vor: sieht man sich z.B. die Menge aller positiven ganzen Zahlen in ihrer üblichen Reihenfolge $1 < 2 < 3 < \dots$ an, dann ist das eine Wohlordnung. Sieht man sich die gleiche Menge mit der neuen Ordnung $1 < 3 < 5 < \dots < 2 < 4 < 6 < \dots$ an, dann ist das noch immer eine Wohlordnung. Der Unterschied ist, dass die zweite *zwei* Elemente ohne Vorgänger hat (1 und 2), während die erste bloß *ein* solches Element hat. Bei der zweiten Wohlordnung bekommen wir die positiven ganzen Zahlen sozusagen *zweimal* in ihrer üblichen Reihenfolgen hinein. Wir stellen uns vor, wir würden die ungeraden Zahlen der Reihe nach mit positiven ganzen Zahlen indizieren und dann die geraden noch einmal, also $1_1 < 3_2 < 5_3 < \dots < 2_1 < 4_2 < 6_3 < \dots$. Die Indizes spiegeln dann den ‘Typ’ dieser Wohlordnung. Es ist gewissermaßen so, als hätten wir die Reihe der positiven ganzen Zahlen in ihrer üblichen Reihenfolge dupliziert und *dazu addiert*.

Die Theorie der (unendlichen) Ordinalzahlen und deren Arithmetik ist eine Präzision dieser Ideen.

Wie auch die Kardinalzahlen entwickeln wir die Ordinalzahlen in der Art von Frege und Russell, i. e. wir betrachten Wohlordnungen und stecken Wohlordnungen, die sich *ähnlich* (i.e. *ordnungsisomorph*) sind, in eine Menge und nennen diese Menge dann die Ordinalzahl dieser Wohlordnungen.

Definition 5.9. Für je zwei Wohlordnungen R auf A und S auf B gilt: R ist *ordnungsisomorph* zu S , in Zeichen $R_A \cong S_B$, genau dann wenn es eine bijektive Funktion f von A nach B ist, für die gilt: $\forall a, b \in A : aRb \leftrightarrow f(a)Sf(b)$.

Falls es so einen Ordnungsisomorphismus f gibt, dann ist dieser eindeutig bestimmt. Ordnungsisomorphie ist eine Äquivalenzrelation auf der Menge aller Wohlordnungen und partitioniert diese in Äquivalenzklassen. Wir nennen diese Äquivalenzklassen auch *Ordnungstypen* und schreiben für den Ordnungstyp einer Wohlordnung R auf A $o.t.(R_A)$.

Das Prädikat eine Ordinalzahl zu sein, $On(x)$ definieren wir - obigem folgend - mit

Definition 5.10. $On(x) :\leftrightarrow \exists R \exists A (WO(R_A) \wedge x = o.t.(R_A))$

Ordinalzahlen sind also Ordnungstypen von Wohlordnungen. Wir verwenden in Zukunft kleine griechische Buchstaben $\alpha, \beta, \gamma, \dots$ (womöglich mit Indizes) für Ordinalzahlen und bezeichnen die nach S-Komprehension existierende Menge aller Ordinalzahlen mit ORD .

Um Wohlordnungen ‘ihrer Länge nach’ vergleichen zu können, definieren wir noch die *Fortsetzung* einer Wohlordnung.

Definition 5.11. Für zwei Wohlordnungen R auf A und S auf B gilt: S ist eine *Fortsetzung* von R , wenn gilt: $\exists b \in B (R_A \cong S_{Pred(b, S, B)})$

Damit können wir eine Ordnung $<$ auf ORD definieren, indem wir für je zwei Ordinalzahlen $\alpha, \beta \in ORD$, $R_A \in \alpha$ und $S_B \in \beta$ festlegen: $\alpha < \beta$ genau dann wenn S_B eine Fortsetzung von R_A ist.

Satz 5.7. Die Ordnungsrelation $<$ auf ORD ist eine strikte Wohlordnung.

Mit $\alpha \leq \beta :\leftrightarrow \alpha < \beta \vee R_A \cong S_B$ für $R_A \in \alpha$ und $S_B \in \beta$ haben wir mit $\leq$ eine schwache Wohlordnung auf ORD .

Definition 1.4. folgend ist das durch die Ordinalzahl α eindeutig bestimmte (echte) Anfangsstück der Reihe der Ordinalzahlen $Pred(\alpha, \leq, ORD) := \{\beta : \beta < \alpha\}$,

Man beachte, dass, falls die Ordinalzahlen α und β den relativen Typ n haben, $Pred(\alpha, \leq, ORD)$ den relativen Typ $n+1$ hat. Da auch $\leq_{Pred(\alpha, \leq, ORD)}$ eine Wohlordnung ist, hat $\leq_{Pred(\alpha, \leq, ORD)}$ einen Ordnungstyp, i.e. eine Ordinalzahl.

Die Ordinalzahl, von der $\leq_{Pred(\alpha, \leq, ORD)}$ ein Element ist, hat den relativen Typ $n+2$.

Auf den Ordinalzahlen können wir nun eine Addition und Multiplikation definieren durch

Definition 5.12. Seien α und β Ordinalzahlen und R_A, S_B Wohlordnungen mit $R_A \in \alpha$ und $S_B \in \beta$. Dann ist

$$\alpha +_o \beta := o.t.(P_{A \oplus B})$$

Die Ordnungsrelation P ist dabei die Menge $\{((a, 0), (a', 0)) : a, a' \in A \wedge aRa'\} \cup \{((b, 1), (b', 1)) : b, b' \in B \wedge bSb'\} \cup \{((a, 0), (b, 1)) : a \in A \wedge b \in B\}$ ⁴⁰

Auf ähnlich Weise kann man auch die ordinale Multiplikation $\cdot_o$ definieren und folgende Fakten über Ordinalzahlen zeigen:

Satz 5.8. 1. $\alpha +_o 0 = \alpha$

$$2. \alpha +_o (\beta +_o \gamma) = (\alpha +_o \beta) +_o \gamma$$

$$3. \alpha \cdot_o 0 = 0$$

$$4. \alpha \cdot_o (\beta \cdot_o \gamma) = (\alpha \cdot_o \beta) \cdot_o \gamma$$

$$5. \alpha \cdot_o (\beta +_o \gamma) = \alpha \cdot_o \beta +_o \alpha \cdot_o \gamma$$

Man beachte, dass weder die ordinale Multiplikation noch Addition im Allgemeinen kommutativ sind.⁴¹

Exkurs: Die natürlichen Ordinalzahlen

Die natürlichen Ordinalzahlen $\mathcal{N}_o$ sind genau die Ordnungstypen von Wohlordnungen endlicher Länge, d.h. von Wohlordnungen die auf Mengen definiert sind, deren Kardinalitäten natürliche Kardinalzahlen sind. Die natürliche Ordinalzahl 3 z.B. ist die Menge aller Wohlordnungen auf 3-elementigen Mengen. Man kann nun in der offensichtlichen Weise die natürlichen Kardinalzahlen mit ihren ordinalen Gegenstücken korrelieren⁴²; *identifizieren* kann man sie *nicht*! Ordinalzahlen sind Mengen von Mengen von geordneten Paaren; Kardinalzahlen sind Mengen von Mengen von *irgendwelchen* Dingen.

Auch die Menge aller natürlichen Zahlen (egal ob Kardinal- oder Ordinalzahlen) mit der natürlichen Ordnung bilden eine Wohlordnung. Den Ordnungstyp dieser Wohlordnung bezeichnen wir mit ω . Wie oben schon gesagt wurde, ist die ordinale Addition nicht kommutativ. Ein Beispiel dafür liefert ω , denn es ist $1 + \omega = \omega$ aber $\omega \neq \omega + 1$.

⁴⁰Die Unordnung oben besagt, dass wir die Menge $A \oplus B$ so ordnen, dass die ersten α Elemente so geordnet sind wie R_A , die nächsten β Elemente so wie S_B und überdies alle Elemente von $B \times \{1\}$ größer sind als die von $A \times \{0\}$. Suggestiver sieht das z.B. so aus: $(a, 0) <_P (b, 0) <_P (c, 0) <_P (a, 1) <_P (b, 1)$ für die Wohlordnungen $a <_R b <_R c$ und $a <_S b$.

⁴¹Im folgenden verwenden wir wieder $+$ statt $+_o$. Aus dem Kontext sollte immer klar sein, wann ordinale und wann kardinale Addition bzw. Multiplikation gemeint sind.

⁴²Das heißt, $\langle \mathcal{N}, +, \cdot, < \rangle$ und $\langle \mathcal{N}_o, +_o, \cdot_o, <_o \rangle$ sind isomorphe Strukturen.

5.3 Mehr ordinale Arithmetik

Ein besonders wichtiges Werkzeug, um Sätze über Ordinalzahlen beweisen zu können, ist das Prinzip der *transfiniten Induktion*. In der Form für Ordinalzahlen⁴³ besagt es: Falls $\alpha, \beta \in ORD$, dann gilt:

$$[(\forall \beta < \alpha (\varphi(\beta))) \rightarrow \varphi(\alpha)] \rightarrow \forall \alpha (\varphi(\alpha))$$

Wenn man also für eine beliebige Ordinalzahl α zeigen kann, dass auf sie eine Eigenschaft zutrifft, falls sie auf alle ihre Vorgänger zutrifft, dann trifft diese Eigenschaft auf *alle* Ordinalzahlen zu. Wie im Falle der vollständigen Induktion für die natürlichen Zahlen ist aber auch hier zu beachten, dass Induktionsbeweise nur über stratifizierte Bedingungen möglich sind.

Ein einfaches, aber wichtiges Ergebnis ist folgender

Satz 5.9. *Sei R eine Wohlordnung auf einer Menge A und $a \in A$; dann gilt: $o.t.(R_{Pred(a,R,A)}) < o.t.(R_A)$.*

Jedes echte Anfangsstück einer Wohlordnung ist echt kürzer als die Wohlordnung selbst. (Das gilt natürlich insbesondere für die Menge aller Ordinalzahlen mit der Ordnung $\leq$.)

Halten wir kurz inne und sehen uns an, wie die Folge der Ordinalzahlen aussieht:

$$0 < 1 < 2 < \dots < \omega < \omega + 1 < \omega + 2 < \dots < \omega + \omega < \dots$$

Wenn man sich das so ansieht, könnte man auf folgende Vermutung kommen

- $o.t.(\leq_{Pred(\alpha, \leq, ORD)}) = \alpha$

Zu Deutsch: Die Länge der Wohlordnung $\leq$ auf der Menge aller Ordinalzahlen kleiner α , ist *gleich* α . Mit anderen Worten: Jede Ordinalzahl *zählt* die Menge aller ihrer Vorgänger.

Dass das zunächst nicht unplausibel klingt, sieht man ja: Der Ordnungstyp der Menge $\{0, 1, 2, 3\}$ mit der üblichen Ordnung auf den Ordinalzahlen ist offensichtlich 4. Der Ordnungstyp der Menge aller Ordinalzahlen kleiner ω (das heißt der Menge aller natürlichen Zahlen) mit der üblichen Ordnung ist ω . Gilt obige Vermutung also für *alle* Ordinalzahlen?

Wie im Anschluss an Satz 5.7. festgestellt wurde, ist auch $\leq$ auf ORD eine Wohlordnung und hat daher einen Ordnungstyp, i. e. eine Ordinalzahl. Bezeichne also Ω diesen Ordnungstyp. Es ist Ω eine Ordinalzahl, also $\Omega \in ORD$. Nach

⁴³Es gibt allgemeinere Varianten, wo die Induktion über beliebige Wohlordnungen geführt wird, bzw. - noch allgemeiner - über *fundierte Strukturen*.

Satz 5.9. gilt nun: $o.t.(\leq_{Pred(\Omega, \leq, ORD)}) < o.t.(\leq_{ORD}) = \Omega$. Nach unserer ‘Vermutung’ von oben gilt aber auch, dass $o.t.(\leq_{Pred(\Omega, \leq, ORD)}) = \Omega$, insgesamt also $\Omega < \Omega$!

Die Ableitung dieser Antinomie, auch bekannt als *Paradox von Burali-Forti*, ist nun aber in NFU (glücklicherweise) nicht möglich, und zwar deshalb, weil die ‘Vermutung’ von oben nicht bewiesen werden kann (tatsächlich sogar *falsch* ist). Die einzige Möglichkeit diese Vermutung zu beweisen würde über transfinite Induktion erfolgen müssen. Sieht man sich aber die Induktionsbedingung $o.t.(\leq_{Pred(\alpha, \leq, ORD)}) = \alpha$ genauer an, merkt man, dass sie nicht stratifiziert ist. Transfinite Induktion ist aber nur über stratifizierte Bedingungen möglich.

Bevor wir uns mit dem Thema weiter beschäftigen, definieren wir eine weitere T-Operation

Definition 5.13. Für eine Ordinalzahl α und eine Wohlordnung $R_A \in \alpha$ sei $T(\alpha) := o.t.(SI(R_A))$ bzw. allgemein $T^n(\alpha) := o.t.(SI^n(R_A))$

Wie auch die T-Operation, die für natürliche Zahlen im Zusammenhang mit dem *Axiom of Counting* eingeführt wurde, ist auch diese eine *typen-erhöhende Operation*, d.h.: Hat die Wohlordnung R_A den relativen Typ n , so hat $SI(R_A)$ den Typ $n + 1$; falls $R_A \in \alpha$, so hat $T(\alpha)$ den Typ $n + 2$.

Sehen wir uns nun noch einmal die ‘Vermutung’ von oben an, so sehen wir, dass wir, wenn wir auf der rechten Seite der Gleichung α durch $T^2(\alpha)$ ersetzen, eine stratifizierte Bedingung haben, über die wir einen Induktionsbeweis führen können. Wir können nun zeigen, dass

Satz 5.10. $o.t.(\leq_{Pred(\alpha, \leq, ORD)}) = T^2(\alpha)$

Mit den Überlegungen zum Paradox von Burali-Forti zeigt das folgenden:

Satz 5.11. $T^2(\Omega) < \Omega$

Ω ist also eine jener Ordinalzahlen, für es gerade *nicht* der Fall ist, dass sie die Menge ihrer Vorgänger ‘zählt’. Dazu kommt es, weil die beiden Ordinalzahlen, die man mit einem echten Anfangsabschnitt der Folge der Ordinalzahlen verbinden kann, nämlich 1.) den Ordnungstyp dieses Abschnitts und 2.) die kleinste Ordinalzahl, die nicht in diesem Abschnitt ist, *nicht identifiziert* werden. Während nämlich α größer als β ist für alle $\beta \in Pred(\alpha, \leq, ORD)$, gilt das *nicht* offensichtlich für $o.t.(\leq_{Pred(\alpha, \leq, ORD)})$.

Satz 5.11. zeigt im Gegenteil, dass für bestimmte Ordinalzahlen - in der Regel sind das sehr große - 1.) und 2.) *nicht übereinstimmen*. Es gibt Ordinalzahlen α , die ein Anfangsstück der Reihe

$$0 < 1 < 2 < 3 < \dots < \omega < \omega + 1 < \dots$$

bestimmen, das ‘kürzer’ ist als alle Wohlordnungen vom Ordnungstyp α .

Das klingt beinahe absurd und schrammt nur haarscharf am Paradoxen vorbei. Wir können das Bild, das Satz 5.11. von großen Ordinalzahlen gibt, nämlich erweitern und erhalten eine absteigende Folge von Ordinalzahlen $\Omega > T^2(\Omega) > T^4(\Omega) > \dots$. Die T-Operation bildet für große Argumente (wie Ω) Ordinalzahlen auf kleinere Ordinalzahlen ab. Unsere Ordinalzahl-Reihe sieht also so aus:

$$0 < 1 < 2 < \dots < \omega < \dots < \dots < T^4(\Omega) < \dots < T^2(\Omega) < \dots < \Omega < \Omega + 1 < \dots$$

Das Verstörende an dieser Situation ist, dass etwa $T^2(\Omega)$ in der Reihe der Ordinalzahlen offenbar vor Ω kommt und man erwartet, dass $T^2(\Omega)$ in $Pred(\Omega, \leq, ORD)$ liegt, weil $Pred(\Omega, \leq, ORD)$ alle Ordinalzahlen kleiner Ω beinhaltet. Andererseits ist $T^2(\Omega)$ ja gerade der Ordnungstyp der üblichen Wohlordnung auf $Pred(\Omega, \leq, ORD)$. Das heißt die Ordinalzahl $T^2(\Omega)$ misst gewissermaßen eine Wohlordnung auf einer Menge, von der sie selbst Element ist.

Das einzige, was uns daran hindert, die paradoxen Konsequenzen zu ziehen, ist die Unmöglichkeit in der Aussage $T^2(\Omega) \in Pred(\Omega, \leq, ORD)$ Typenindizes derart anzubringen, dass die Aussage wohlgeformt nach den Regeln der Typentheorie wäre.

Eine weitere seltsame Eigenschaft der Folge $\Omega > T^2(\Omega) > T^4(\Omega) > \dots$ ist, dass sie offenbar kein kleinstes Element hat. Das ist insofern beachtlich, als ja die Menge aller Ordinalzahlen mit der üblichen Ordnung eine Wohlordnung ist, i.e. *jede Teilmenge von ORD hat ein kleinstes Element*. (Limiten dürfen nur ‘nach rechts’ auftreten.) Die Menge $OK := \{\Omega, T^2(\Omega), T^4(\Omega), \dots\}$ hat aber offenbar kein kleinstes Element!

Wieder rettet uns nur die Tatsache, dass OK keine stratifizierte Definition hat, vor Widersprüchen. OK ist ein Beispiel für eine *echte Klasse* im Sinne von NFU. Dennoch bleibt der Umstand, dass in einem Model von NFU die Ordinalzahlen nicht wohlgeordnet sind, d.h. in einem bestimmten Sinne kann das System NFU nicht ‘erkennen’, dass die Ordinalzahlen nicht wohlgeordnet sind. Das System kann sozusagen ‘definitive Entscheidungen’ nur für stratifizierte Sätze als ‘Input’ treffen.

Sehen wir uns Satz 5.11. noch einmal an: Was er - abgesehen von obigen Implikationen die Arithmetik betreffend - aussagt, ist doch, dass die Wohlordnung auf der Menge aller SSingletons von Ordinalzahlen $\{\{0\}\} < \{\{1\}\} < \{\{2\}\} < \dots$ echt kürzer ist als die übliche Wohlordnung auf der Menge aller Ordinalzahlen $0 < 1 < 2 < \dots$. Weil man mit ähnlichen Argumenten wie oben auch zeigen kann, dass $T(\Omega) < \Omega$ ist auch die Wohlordnung auf der Menge aller Singletons von Ordinalzahlen echt kürzer als die übliche Wohlordnung auf ORD. Das aber heißt nichts anderes als

Satz 5.12. *Es gibt keine Funktion $f : ORD \longrightarrow \mathcal{P}_1(ORD)$ mit $f(\alpha) = \{\alpha\}$ (und damit erst recht keine Funktion $g : V \longrightarrow V$ mit $g(x) = \{x\}$).*

Die Singleton-Operation verdient ihren Namen zu Recht! Sie ist keine Funktion im Sinne einer Menge von geordneten Paaren, sondern eine ‘echte’ Operation. Das sollte uns aus verschiedenen Gründen nicht überraschen. Ein ‘falscher’

Grund wäre aber, nicht überrascht zu sein, weil die Singleton-Operation auch in ZFC eine ‘echte’ Operation ist. In ZFC ist ja der Grund, dass die Singleton-Operation keine Funktion ist, weil damit die Domain dieser ‘Funktion’ (das Universum V) eine Menge wäre. Die Ursache für die ‘Nicht-Funktionalität’ der Singleton-Operation in ZFC hat also etwas mit ‘Größe’ zu tun. Da Größe (zumindest *prima facie*) keine Rolle spielt in NFU könnte man also mit einigem Recht erwarten, dass auch Funktionen zwischen großen Mengen in Ordnung sind.

Ein ‘guter’ Grund für Zweifel an der Existenz der Singleton-Operation *als Funktion* wäre der, dass so eine Funktion keine stratifizierte Bedingung haben kann. Es war also, weil ja nur stratifizierte Bedingungen mit Sicherheit Mengen definieren, von vornherein zweifelhaft, ob es so eine Funktion geben kann.

Bevor wir uns wieder mit dem Thema Kardinalzahlen beschäftigen, noch eine Bemerkung:

Satz 5.11. machte unsere Hoffnungen zunichte, für beliebige Ordinalzahlen $T(\alpha) = \alpha$ zu bekommen. Dennoch sind wir an dieser Eigenschaft von Ordinalzahlen interessiert, und erwarten, dass Fälle wie $T(\alpha) < \alpha$ eher die Ausnahme als die Regel sind. Ordinalzahlen, für die $T(\alpha) = \alpha$ gilt, nennen wir *Cantor’sch* (eine Bezeichnung, die sich auf den nächsten Seiten klären wird).

Dass bestimmte große Ordinalzahlen (wie Ω) nicht Cantor’sch sind, wissen wir. Aber welche *sind* es? Mit dem *Axiom of Counting* können wir zeigen, dass jede natürliche Ordinalzahl Cantor’sch ist, außerdem die Ordinalzahlen $\omega, \omega + 1 \dots$ Wie *genau* das ‘Cantor’sche Teiluniversum’ in NFU aussieht, hängt im wesentlichen davon, durch welche zusätzlichen Axiome man NFU verstärkt.⁴⁴ Ohne das Axiom of Counting, konnten wir ja nicht einmal beweisen, dass für alle natürlichen Kardinalzahlen $T(n) = n$ ist (hier ist die T-Operation auf den natürlichen Kardinalzahlen gemeint). Wenn wir also $A \in n$ nicht bijektiv auf $\mathcal{P}_1(A) \in T(n)$ abbilden können, dann können wir erst recht keinen Ordnungsisomorphismus finden. Wir könnten also nicht einmal für alle natürlichen Ordinalzahlen zeigen, dass $T(n) = n$ ist (hier ist wieder die T-Operation auf den Ordinalzahlen gemeint).

Doch zurück zum Thema Kardinalzahlen

5.4 Mehr kardinale Arithmetik

Gemäß unserem Plan, Kardinalzahlen als Mengen von Mengen gleicher Mächtigkeit aufzufassen, definieren wir nun die erste unendliche Kardinalzahl

Definition 5.14. $\aleph_0 := \{A : A \approx \mathcal{N}\}$

$\aleph_0$ ist also die Kardinalzahl, die jeder Menge zukommt, die gleich groß wie die Menge aller natürlichen Zahlen ist. Wir nennen Mengen der Mächtigkeit $\aleph_0$ auch *abzählbar* oder *abzählbar unendlich*. Hier einige einfache Sätze über

⁴⁴Vgl. etwa Holmes [1998] (insbesondere S. 112 ff., S. 197 ff. und Holmes [1994] zum Thema *Axiom of small ordinals*, *Axiom of large ordinals* und anderen Verstärkungen.

$\aleph_0$, die in NFU beweisbar sind. (Die Definitionen der kardinalen Addition und Multiplikation bleiben auch für unendliche Kardinalzahlen bestehen.)

Satz 5.13. 1. $\aleph_0 + n = \aleph_0$ für jede natürliche Zahl n .

2. $\aleph_0 + \aleph_0 = \aleph_0$

3. $\aleph_0 \cdot \aleph_0 = \aleph_0$

Allgemein können wir für unendliche Kardinalzahlen zeigen, dass gilt:

Satz 5.14. 1. $\lambda + \kappa = \kappa + \lambda = \kappa$ falls $\lambda \leq \kappa$

2. $\lambda \cdot \kappa = \kappa \cdot \lambda = \kappa$ falls $\lambda \leq \kappa$

3. $\lambda + \kappa = \lambda \cdot \kappa = \max\{\lambda, \kappa\}$

Doch wie sieht es nun mit Mengen größerer Mächtigkeit als $\aleph_0$ aus? Eine gute Idee wäre, zu versuchen, zu Mengen größerer Mächtigkeit zu gelangen, indem wir den Potenzmengenoperator auf $\mathcal{N}$ anwenden.

Der Satz von Cantor in NFU

Der Satz von Cantor besagt, dass für *alle* Mengen A gilt: $A \prec \mathcal{P}(A)$. Das heißt, die Potenzmenge *jeder* Menge ist *echt größer* als die Menge selbst. Das klingt ziemlich plausibel. Die Menge aller Teilmengen einer Menge A enthält ja *mindestens* alle Singletons der ursprünglichen Menge, außerdem aber noch viele mehr. Für die natürlichen Zahlen etwa haben wir neben den Singletons von natürlichen Zahlen zusätzlich noch die Menge aller geraden, aller ungeraden, die Menge aller Primzahlen usw.

Um zu sehen, wieso dieser Satz trotz seiner Plausibilität in NFU dennoch *nicht* gilt, sehen wir uns den ‘Beweis’ dazu an:

Wir nehmen an, es gibt eine surjektive Funktion $f : A \rightarrow \mathcal{P}(A)$. Nun bilden wir die sogenannte Diagonalmenge $D \subseteq A$, die alle Elemente $a \in A$ enthält, für die gilt: $a \notin f(a)$. Da D eine Teilmenge von A , also Element von $\mathcal{P}(A)$ ist, muss D Wert der surjektiven Funktion f sein, d.h. es muss ein $a_0 \in A$ geben sodass $f(a_0) = D = \{a \in A : a \notin f(a)\}$. Für dieses a_0 gilt aber $a_0 \in f(a_0) \leftrightarrow a_0 \notin f(a_0)$ nach Definition der Menge D . Widerspruch.

Im Gegensatz zur Annahme müssen wir also schließen, dass es keine surjektive Funktion von A nach $\mathcal{P}(A)$ gibt. Das heißt, jede Menge ist kleiner als ihre Potenzmenge.

Dass dieser Satz in NFU nicht gelten kann, liegt eigentlich auf der Hand: Weil in NFU V eine Menge ist, und zwar eine, die alles enthält, kann es keine Menge geben, die noch größer wäre. Aber wie entgeht man dieser Konsequenz ohne das ‘Vernünftige’ an diesem Satz über Bord zu werfen?

Der Punkt, wo dieser ‘Beweis’ in NFU zusammenbricht, ist der, wo wir die Diagonalmenge D bilden, denn die definierende Bedingung für die Menge D war $a \notin f(a)$, und diese ist nicht stratifiziert.

Ein weiterer Fall von ‘trügerischer Intuition’ liegt vor, wenn man versucht über die ‘offensichtliche’ Äquivalenz $A \approx \mathcal{P}_1(A)$ zum gleichen Ergebnis zu gelangen; denn während man nun die Diagonalmenge bzgl. $\mathcal{P}_1(A)$ bilden kann, gibt es keinen ‘offensichtlichen’ Weg A bijektiv auf $\mathcal{P}_1(A)$ abzubilden, denn wie im Zusammenhang mit dem Paradox von Burali-Forti bemerkt wurde, gibt es im Allgemeinen keine *Funktion*, die ein Element von A auf sein Singleton abbildet. Die ‘offensichtliche’ Äquivalenz ist also gar keine Äquivalenz (zumindest nicht für *alle* Mengen A).

Aus diesen Bemerkungen ergibt sich erstens

Satz 5.15. (*Satz von Cantor*)⁴⁵

Für alle Mengen A gilt: $\mathcal{P}_1(A) \prec \mathcal{P}(A)$

Außerdem aber geben sie Anlass zu einer

Definition 5.15. Eine Menge A heißt *Cantorsch*, wenn gilt: $A \approx \mathcal{P}_1(A)$

Cantor’sche Mengen sind also gerade die Mengen, für die der Satz von Cantor in seiner ursprünglichen Form gilt. Erweitern wir unsere T-Operation für die natürlichen Kardinalzahlen auf Kardinalzahlen im Allgemeinen, heißt das, dass eine Menge Cantorsch ist, wenn für ihre Kardinalzahl κ gilt: $T(\kappa) = \kappa$.⁴⁶

Dass nicht alle Mengen diese Eigenschaft haben, sieht man am Beispiel V . Denn während nach Satz 5.15. $\mathcal{P}_1(V)$ kleiner als $\mathcal{P}(V)$ ist, gilt *nicht*, dass $\mathcal{P}_1(V)$ gleich groß wie V ist (und damit $V \prec \mathcal{P}(V)$). Das Universum V ist ein Beispiel für eine *Nicht-Cantorsche Menge*.

Es stellt sich die Frage, welche Größe die Potenzmenge des Universums nun *hat*? Noch spricht nichts dagegen, dass beide gleich groß sind. Folgender Satz gibt die Antwort (und es fallen dabei Zusatzinformationen ab):

Satz 5.16. (*Satz von Specker*)

$\mathcal{P}(V) \prec V$

Die Potenzmenge des Universums ist kleiner als das Universum selbst. Die Zusatzinformation ist nun: Es gibt Urelemente! Weil die Potenzmenge des Universums die Menge aller *Mengen* ist, folgt aus $\mathcal{P}(V) \prec V$, dass es Urelemente geben muss!

Das ist insofern nicht überraschend, als wir ja Urelemente in unserem System explizit einkalkuliert haben. Trotzdem: Dass wir ihre Existenz *beweisen* können, stand im vorhinein *nicht* zu erwarten.

Eine der wichtigsten arithmetischen Operationen - zumal im Bereich der transfiniten Zahlen - ist die Exponentiation. Zahlreiche Fragen der Mengenlehre stehen in engem Kontakt zu ihr. Als bekannteste sei nur das sogenannte

⁴⁵Der Beweis dazu verläuft ganz analog zum ursprünglichen Beweis.

⁴⁶Hier erklärt sich die Bezeichnung *Cantor’sche Ordinalzahl*: *Cantor’sche Ordinalzahlen* sind nämlich genau die Ordnungstypen von Wohlordnungen, die auf *Cantor’schen Mengen* definiert sind.

Kontinuumproblem erwähnt, das die frühe Phase der Mengenlehre so dominiert hat. Grob gesagt, ist das (allgemeine) Kontinuumproblem die Frage, ob die Reihe der unendlichen Kardinalzahlen durch die Exponentiation 2^κ vollständig festgelegt ist, d.h. ob 2^κ der kardinale Nachfolger von κ ist (i.e. die kleinste Kardinalzahl größer κ), oder ob es Kardinalzahlen ‘dazwischen’ gibt.

Ohne auf dieses Problem im folgenden einzugehen, wollen wir sehen, wie wir unsere Exponentiation für Kardinalzahlen überhaupt definieren.

Exponentiation

Bevor wir auf die kardinale Exponentiation im Allgemeinen zu sprechen kommen, wollen wir uns den Spezialfall für natürliche Zahlen ansehen für die Basis 2.

Eine angemessene Definition bereitet uns in diesem Fall keine Schwierigkeiten: Wir können 2^n rekursiv definieren mit der Anfangsbedingung $2^0 := 1$ und dem Rekursionsschritt $2^{n+1} := 2 \cdot 2^n$. Das Gute an dieser Definition ist, dass der Ausdruck 2^n vom gleichen relativen Typ wie n selbst ist, wir können also ohne Probleme zeigen, dass z.B. für alle natürlichen Zahlen n gilt: $n < 2^n$. Vor allem aber ‘funktioniert’ diese Definition, weil klar ist, was der kardinale Nachfolger von n ist; nämlich $S(n) = n + 1$. Für unendliche Kardinalzahlen κ ist aber zunächst nicht klar, was überhaupt der kardinale Nachfolger von κ sein soll: Auf jeden Fall wissen wir, dass es nicht $\kappa + 1$ ist, denn $\kappa + 1 = \kappa$ nach Satz 5.13.

Rekursion (analog zu den natürlichen Zahlen) fällt als Definitionsmethode von 2^κ also weg. Wir müssen einen anderen Weg finden und machen uns dabei folgende Beobachtung zunutze: Für jede Menge A gilt:

$$\mathcal{P}(A) \approx \{f : Funk(f) \wedge f : A \longrightarrow \{a, b\}\}$$

für irgendwelche Mengen $\{a, b\}$, für die $a \neq b$. Dabei entspricht die Menge auf der rechten Seite der Menge aller charakteristischen Funktionen auf A ⁴⁷. Weil das nun für beliebige Zielmengen $\{a, b\}$ gilt, besagt obiges im wesentlichen, dass für beliebiges $N \in 2$ gilt:⁴⁸

$$\mathcal{P}(A) \approx N^A$$

Nach dem bisher gesagten scheint es also vernünftig unsere Exponentiation zu definieren durch $2^{|A|} := |\mathcal{P}(A)|$. Das Problem, das wir haben, ist dass, falls $|A|$ vom relativen Typ $n + 1$ ist, $|\mathcal{P}(A)|$ vom relativen Typ $n + 2$ ist. Wir hätten aber gerne, dass $\kappa \mapsto 2^\kappa$ eine Funktion definiert. So wie sie da steht, ist die Definition von 2^κ aber nicht stratifiziert, bestimmt also nicht unbedingt eine Funktion.

Kurz: Wir müssen darauf achten, dass unsere Definition die Zahl 2^κ vom gleichen Typ macht wie κ selbst.

⁴⁷Man korreliert also 1-1 mit jeder Teilmenge B von A eine charakteristische Funktion f_B , die alle Elemente von B auf a abbildet und jedes Element von B^C_A auf b .

⁴⁸Wir wollen im Folgenden für zwei Mengen A und B unter A^B die Menge aller Funktionen von B nach A verstehen. Aus dem Kontext sollte immer klar sein, wann im Gegensatz dazu die eigentliche Exponentiation gemeint ist.

Um das zu erreichen bedienen wir uns der auf unendliche Kardinalzahlen erweiterten T-Operation aus dem vorigen Abschnitt und definieren

Definition 5.16. $2^{|A|} := T^{-1}(|\mathcal{P}(A)|)$

Diese Definition ist stratifiziert, denn mit A vom relativen Typ n , hat $|\mathcal{P}(A)|$ den relativen Typ $n + 2$, und weil T^{-1} eine typen-senkende Operation ist, hat $T^{-1}(|\mathcal{P}(A)|) = 2^{|A|}$ den relativen Typ $n + 1$ - genau wie $|A|!$. Zu beachten ist, dass mit dieser Definition der Exponentiation 2^κ nicht für *alle* κ definiert ist. Zumindest für sehr große Kardinalzahlen (wie $\kappa = |V|$) ist nämlich $T^{-1}(\kappa)$ nicht definiert.⁴⁹ Mit diesen Einschränkungen im Hinterkopf können wir nun die Exponentiation für allgemeine Kardinalzahlen definieren.

Definition 5.17. Seien κ, λ Kardinalzahlen und $A \in \kappa, B \in \lambda$. Dann ist

$$\kappa^\lambda := T^{-1}(|A^B|)$$

falls $T^{-1}(|A^B|)$ für $|A^B|$ definiert ist.

Für Kardinalzahlen, für die 2^κ definiert ist, können wir nun zeigen, dass $\kappa < 2^\kappa$. Insbesondere liefert uns die Exponentiation die Existenz von Kardinalzahlen, die größer als $\aleph_0$ sind, denn es ist $\aleph_0 = |\mathcal{N}| < |\mathcal{P}(\mathcal{N})| = 2^{\aleph_0}$.⁵⁰ Die *kleinste* Kardinalzahl größer $\aleph_0$ (den kardinalen Nachfolger von $\aleph_0$) bezeichnen wir mit $\aleph_1$. Die Kontinuumshypothese besagt also gerade, dass $\aleph_1 = 2^{\aleph_0}$.

⁴⁹Denn $T^{-1}(V)$ ist diejenige Kardinalzahl κ , für die gilt $T(\kappa) = |V|$. Für eine bestimmte Menge B mit $|B| = \kappa$ muss also gelten $|\mathcal{P}_1(B)| = T(|B|) = |V|$. Es gibt aber keine Menge B , die diese Gleichung erfüllt. Die einzige dafür in Frage kommende Menge, V selbst, fällt aus, denn wie gezeigt wurde, ist V *nicht* Cantor'sch.

⁵⁰Man beachte, dass - auch ohne das Axiom of Counting - $\mathcal{N}$ Cantor'sch ist (Unendlichkeitsaxiom vorausgesetzt).

6 Kardinal- und Ordinalzahlen in ZFC

Bevor wir mit dem Aufbau der Theorie der Ordinal- und Kardinalzahlen in ZFC beginnen, sollten wir uns ins Gedächtnis rufen, dass in ZFC das Universum V keine Menge, kein Wert einer quantifizierten Variable ist. Aus diesem Grund ist es auch nicht möglich, Ordinal- und Kardinalzahlen so aufzufassen wie in NFU, im Falle der natürlichen Zahl n etwa als die Menge aller Mengen mit n Elementen. Schon die Menge aller Mengen mit einem Element, also die 1 in NFU, ist nämlich in ZFC ‘gleich groß’ wie das Universum.

Da also die Entwicklung der Ordinal- und Kardinalzahlen a la Frege-Russell in ZFC nicht möglich ist, muss ein anderer Weg gefunden werden, Kardinal- und Ordinalzahlen irgendwie zu ‘implementieren’. Wie sich gezeigt hat, eignet sich dabei die Variante von Von Neumann besonders gut. Die Grundidee besagt ungefähr folgendes: Wenn wir zählen, dann gehen wir schrittweise immer von einer Zahl zur nächsten. Sollen wir also Ordinalzahlen irgendwie mengenmäßig auffassen, so erscheint es plausibel, eine bestimmte Ordinalzahl gerade als die Menge aller ihrer Vorgänger zu bestimmen. Kardinalzahlen kann man dann in weiterer Folge mit bestimmten Ordinalzahlen identifizieren.

Diese Gedanken näher auszuführen ist das Ziel dieses Kapitels.

6.1 Ordinalzahlen: Grundlegende Begriffe

Die Definition eines Ordnungsisomorphismus zwischen Wohlordnungen R_A und S_B ist dieselbe wie diejenige, die wir für NFU verwendet haben. Weiters ist dieser Isomorphismus eindeutig bestimmt und es gilt:

Satz 6.1. *Für Wohlordnungen R auf A und S auf B gilt genau einer der folgenden Fälle:*

1. $R_A \cong S_B$
2. $R_A \cong S_{Pred(b, S, B)}$ für ein $b \in B$
3. $R_{Pred(a, R, A)} \cong S_B$ für ein $a \in A$

Wie auch in NFU sagen wir im zweiten Fall, dass S eine Fortsetzung von R ist. Außerdem ist keine Wohlordnung ordnungsisomorph zu einem echten Anfangsstück.

Die Theorie der Wohlordnungen in ZFC unterscheidet sich im Prinzip nicht wesentlich von der in NFU. Die Unterschiede beginnen erst dort, wo wir Ordinalzahlen als Maße für die Länge dieser Wohlordnungen einführen. Unser Ziel ist also auch hier, eine Definition der Ordinalzahl zu finden, die es uns gestattet, z.B. der Wohlordnung $a_1 < a_2 < a_3 < \dots < b_1 < b_2 < b_3 < \dots$ einen Ordnungstyp (eine Ordinalzahl) zuzuordnen (und dieser Typ sollte $\omega + \omega$ sein).

Doch vorher noch einige Definitionen.

Definition 6.1. *eine Menge A heißt **transitiv**, (kurz $Trans(A)$) falls gilt:*
$$\forall x(x \in A \rightarrow x \subseteq A)$$

Einfache Beispiele für transitive Mengen sind die Mengen $\{\emptyset, \{\emptyset\}\}$ oder $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$. (Ein anspruchsvolleres Beispiel haben wir im Zusammenhang mit dem Fundierungsaxiom in Abschnitt 3 bereits kennen gelernt. Wir hatten dort mittels transfiniter Rekursion über die Ordinalzahlen (informell) die Von Neumann-schen Stufen R_α definiert. Jede dieser Stufen ist ebenfalls ein Beispiel für eine transitive Menge.)

Weiters definieren wir:

Definition 6.2. Eine Menge A heie durch $\in$ wohlgeordnet, falls die Relation $\in_A := \{(x, y) \in A \times A : x \in y\}$ eine Wohlordnung auf A ist.⁵¹

Ordinalzahlen sind in unserem System nun genau alle transitiven $\in$ -Wohlordnungen, i.e. wir definieren das Prädikat, eine Ordinalzahl zu sein durch

Definition 6.3. $On(x) :\leftrightarrow Trans(x) \wedge WO(\in_x)$

Dass diese Definition ‘gut’ ist, wird sich im weiteren zeigen. Folgender Satz sollte erste Bedenken zerstreuen. (Wir verwenden im Folgenden wie in NFU Variablen $\alpha, \beta, \gamma, \dots$ für Ordinalzahlen.)

Satz 6.2. Für Ordinalzahlen α, β und γ gilt:

1. Wenn $\in_\alpha \cong \in_\beta$, dann ist $\alpha = \beta$
2. Es gilt genau einer der folgenden Fälle: $\alpha \in \beta$, $\beta \in \alpha$ oder $\alpha = \beta$
3. Wenn $\alpha \in \beta$ und $\beta \in \gamma$, dann ist $\alpha \in \gamma$
4. Wenn C eine nichtleere Menge von Ordinalzahlen ist, dann gibt es ein $\in$ -minimales Element, i.e. $\exists \alpha \in C (\forall \beta \in C (\alpha \in \beta \vee \alpha = \beta))$

Der erste Punkt dieses Satzes rechtfertigt gewissermaßen das ‘Weglassen’ der Wohlordnung $\in_\alpha$ beim Sprechen über Ordinalzahlen α , denn aus Isomorphie der Wohlordnungen $\in$ auf α und $\in$ auf β folgt immer schon Gleichheit der Ordinalzahlen (i.e. der ‘blanken Mengen’) α und β . Wir werden deshalb oft einfach $R_A \cong \alpha$ schreiben, statt $R_A \cong \in_\alpha$.

Die Punkte 2 bis 4 besagen intuitiv, dass die Ordinalzahlen mit der Ordnung $\in$ eine Wohlordnung bilden.⁵² 2 stellt dabei sicher, dass die $\in$ -Relation auf den Ordinalzahlen linear ist, 3 die Transitivität und 4 die Fundiertheit (i.e. die Wohlordnungseigenschaft).⁵³

Eine spezielle Eigenschaft der derart definierten Ordinalzahlen ist, dass sich eine größer-kleiner-Beziehung besonders leicht darauf erklären lässt. Der vorige

⁵¹Wir werden in Zukunft wieder $\in_A$ schreiben, falls $\in_A$ eine Wohlordnung auf A ist (und nicht etwa $\in_{A_A}$)

⁵²bzw. bilden würden, wenn die Menge aller Ordinalzahlen existieren würde (als Menge).

⁵³Dass überdies Reflexivität gilt, i.e. für alle Ordinalzahlen $\alpha \notin \alpha$ ist, ist eine Konsequenz aus dem Fundierungsaxiom, würde für Ordinalzahlen aber auch ohne dieses gelten.

Satz besagt ja, dass sich die $\in$ -Relation auf den Ordinalzahlen genau so verhält wie wir das von einer (strikten) Ordnung erwarten.

Wir setzen für Ordinalzahlen α, β also einfach

Definition 6.4. $\alpha < \beta :\leftrightarrow \alpha \in \beta$

Weil jede Ordinalzahl transitiv ist, also aus $\alpha \in \beta$ immer $\alpha \subseteq \beta$ folgt können wir durch $\alpha \subseteq \beta$ eine schwache Ordnung auf den Ordinalzahlen festlegen.

Um zu sehen, wie die Ordinalzahlen konkret aussehen, hier einige Beispiele: $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ usw. Um einzusehen, wieso diese Mengen Ordinalzahlen sind, braucht man nur die definierenden Eigenschaften von Ordinalzahlen zu prüfen. Letztere Menge etwa ist eine Ordinalzahl, weil sie offenbar transitiv ist (alle Elemente sind auch Teilmengen) und außerdem $\emptyset \in \{\emptyset\} \in \{\emptyset, \{\emptyset\}\}$, d.h. $\in$ ist eine Wohlordnung auf $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$.

Wir definieren nun

Definition 6.5. $0 := \emptyset$ und $S(\alpha) := \alpha \cup \{\alpha\}$

und erhalten so ein Anfangstück der Reihe der Ordinalzahlen

$$\emptyset \in \{\emptyset\} \in \{\emptyset, \{\emptyset\}\} \in \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \in \dots$$

oder mit den Definitionen $0 := \emptyset, 1 := S(0), 2 := S(1), 3 := S(2)$, usw. die hübschere Variante

$$0 < 1 < 2 < 3 < \dots$$

Eines der wichtigsten Resultate über Ordinalzahlen in ZFC ist folgender

Satz 6.3. Für jede Wohlordnung R_A gibt es eine Ordinalzahl α , sodass $R_A \cong \alpha$

Den besten Eindruck über das ‘Wesen’ dieses Satzes erhält man, wenn man sich den Beweis dazu ansieht, der Gebrauch von *Hartogs’ Theorem*⁵⁴ macht.

Satz 6.4. (*Hartogs’ Theorem*)

$$\forall A \exists \alpha \neg \exists f (f : \alpha \xrightarrow{\text{inj.}} A)$$

Satz 6.4. besagt, dass, falls uns irgendwie ein Verfahren gegeben ist, eine beliebige Menge A mit Hilfe von Ordinalzahlen ‘aufzulisten’; es dann nicht möglich ist, dass uns die Ordinalzahlen irgendwann ‘ausgehen’. D.h. wenn wir eine Folge $a_0, a_1, a_2, \dots, a_\omega, \dots$ bilden können, dann kann es nicht passieren, dass ‘zu wenige’ Indizes da sind.

Hartogs’ Theorem garantiert uns, dass es genug Ordinalzahlen gibt, um Listen von allen möglichen Mengen zu erstellen (sofern uns ein Verfahren dazu gegeben ist).

⁵⁴Es gibt auch direkte Beweise, ohne diesen Umweg.

Dass daraus leicht Satz 6.3. folgt, sieht man, wenn man sich daran erinnert, dass alle Wohlordnungen vergleichbar sind bzgl. Isomorphie, d.h. für alle Wohlordnungen gilt: Die eine ist Fortsetzung der anderen oder umgekehrt. Beachtet man nun, dass jede Ordinalzahl α ja schon eine Wohlordnung *ist* (nämlich bzgl. $\in$), so sieht man, dass für eine beliebige Wohlordnung R_A gelten muss $\alpha \cong R_{Pred(a, R_A)}$ für ein $a \in A$, $\alpha \cong R_A$ oder $\beta \cong R_A$ für ein $\beta \in \alpha$. Nach Hartogs' Theorem können wir nun aber α so groß wählen, dass die beiden ersten Fälle ausscheiden.

Egal wie lang eine Wohlordnung also ist, wir können sie immer mit einer Ordinalzahl messen.

Satz 6.3. garantiert uns, dass wir Ordinalzahlen als Repräsentanten von Wohlordnungen eines bestimmten Ordnungstyps verwenden können. M.a.W.: Mit den Ordinalzahlen ist es möglich *alle* Wohlordnungen, die es überhaupt gibt, eindeutig bis auf (Ordnungs-)Isomorphie zu charakterisieren.

Satz 6.3. rechtfertigt somit folgende Definition des *Ordnungstyps* einer Wohlordnung

Definition 6.6. $o.t.(R_A) := \iota\alpha(\alpha \cong R_A)$

Satz 6.3. lässt sich dadurch prägnant formulieren: Jede Wohlordnung hat einen Ordnungstyp.⁵⁵

Satz 6.3. machte implizit auch von einem weiteren Resultat Gebrauch, das bisher unterschlagen wurde und in NFU falsch ist. Wir hatten oben gemeint, dass $\beta \cong R_A$ für ein $\beta \in \alpha$ gelten muss. Korrekterweise hätte es heißen müssen: $\in_{Pred(\beta, \alpha, \in_\alpha)} \cong R_A$. Die Schreibweise $\beta \cong R_A$ ließe sich nur dann rechtfertigen, wenn $\beta = Pred(\beta, \alpha, \in_\alpha)$, oder mit anderen Worten $o.t.(\in_{Pred(\beta, \alpha, \in_\alpha)}) = \beta$. Nach dem bisher gesagten ist das aber beinahe trivial. Unser Plan war ja, Ordinalzahlen aufzufassen als Menge all ihrer Vorgänger im Sinne der Elementbeziehung (also genau der $<$ -Beziehung). Es ist für jede Ordinalzahl $\alpha = \{\beta : \beta \in \alpha\} = Pred(\alpha, \alpha, \in_\alpha)$, für $\beta \in \alpha$ also $\beta = Pred(\beta, \alpha, \in_\alpha)$.

Weil jedes Element einer Ordinalzahl wieder eine Ordinalzahl ist, rechtfertigt das aber die obige Argumentation.

Genau jene ‘Vermutung’, die uns in NFU solches Kopfzerbrechen bereitete, und die sich schließlich als falsch herausgestellt hat, wird hier im Rahmen von ZFC gewissermaßen zur Motivation für die Definition der Ordinalzahl.

Aus dieser Beobachtung (zusammen mit Satz 6.2) lässt sich aber noch eine weitere Information herausholen.

⁵⁵Der Beweis von Satz 6.3. (ob nun in der direkten Variante oder über Hartogs' Theorem) macht wesentlich vom Ersetzungsaxiom Gebrauch und kann in ZFC ohne dieses nicht bewiesen werden. Man beachte, dass dieser ‘Satz’ in NFU mit den Frege-Russell-Ordinalzahlen trivial ist. Anders als in ZFC heißt das in NFU aber *nicht*, dass es für alle Mengen ‘genug’ Ordinalzahlen gibt, um alle Elemente ‘aufzuzählen’.

Satz 6.5. (*Paradox von Burali-Forti*)

$$\neg \exists y \forall x (x \in y \leftrightarrow On(x))$$

Es gibt keine Menge, die alle Ordinalzahlen enthält. Sehen wir uns an, wieso das so ist:

Angenommen, es gäbe eine solche Menge ON . Wie bemerkt, ist jede Ordinalzahl eine Menge von Ordinalzahlen (nämlich die Menge aller *kleineren* Ordinalzahlen). Das aber heißt nichts anderes als $\forall \alpha (\alpha \in ON \rightarrow \alpha \subseteq ON)$, d.h. ON ist transitiv. Außerdem besagt Satz 6.2., dass $\in$ eine Wohlordnung auf der Menge ON aller Ordinalzahlen ist. ON ist also eine transitive $\in$ -Wohlordnung, i.e. eine Ordinalzahl, und damit gilt $ON \in ON$. Widerspruch.

Es gibt also keine *Menge*, die alle Ordinalzahlen enthält. Das hindert uns aber nicht daran von ON (informell) als *Klasse* zu sprechen (solange wir uns dabei bewusst sind, dass damit eigentlich ein Prädikat - nämlich $On(x)$ - gemeint ist).

Bevor wir uns weiter mit Ordinalzahlen beschäftigen, wollen wir uns ansehen, wie es um die natürlichen Zahlen steht.

6.2 Exkurs: Die natürlichen Zahlen

Anders als in NFU sprechen wir hier auf einmal von *den* natürlichen Zahlen (und nicht etwa von den natürlichen *Ordinalzahlen*). Wieso ist das so? Wir werden später sehen, dass Kardinalzahlen ganz bestimmte Ordinalzahlen sind. Für alle endlichen Ordinalzahlen n wird sich zeigen, dass $|n| = n$, i.e. die Kardinalzahl einer natürlichen Ordinalzahl ist diese selbst. Das rechtfertigt die Rede von *den* natürlichen Zahlen.

Die Idee, die hinter der Definition der natürlichen Zahlen steht ist ziemlich simpel: Wir erhalten die natürliche Zahl n durch n -malige Anwendung der Nachfolgeroperation S (die oben allgemein für Ordinalzahlen definiert wurde) und erhalten so die Reihe

$$0 \quad 1 := S(0) \quad 2 := S(S(0)) \quad 3 := S(S(S(0))) \quad \dots$$

die oben schon vorweggenommen wurde. Die Eigenschaft, eine natürliche Zahl zu sein definieren wir durch

Definition 6.7. $Nat(x) :\leftrightarrow On(x) \wedge \forall \beta < x (\beta = 0 \vee \exists \gamma (\beta = S(\gamma)))$

Mit anderen Worten: x ist eine natürliche Zahl genau dann wenn x eine Ordinalzahl ist und alle Ordinalzahlen kleiner x entweder 0 oder Nachfolger von irgendwelchen Ordinalzahlen sind (es gibt keine ‘Limiten’, die übersprungen werden).

Entscheidend für die Definition der Menge ω aller natürlichen Zahlen, ist nun die Existenz einer induktiven Menge, aus der die natürlichen Zahlen dann geeignet ‘ausgesondert’ werden können. Genau das leistet das Unendlichkeitsaxiom $\exists y(\emptyset \in y \wedge \forall x(x \in y \rightarrow x \cup \{x\} \in y))$. Das Unendlichkeitsaxiom wurde augenscheinlich gerade so gewählt, dass es bei der Definition der natürlichen Zahlen mit dem durch $S(\alpha) := \alpha \cup \{\alpha\}$ definierten Nachfolger am wenigsten ‘Probleme’ bereitet. Wir können nun für gegebenes induktives S per Aussonderungsaxiom definieren:

Definition 6.8. $\omega := \{x \in S : \text{Nat}(x)\}$

Man beachte, dass, da für alle $n \in \omega$ gilt: $n = \{0, 1, 2, \dots, n-1\} \subseteq \omega$, und $\in_\omega$ eine Wohlordnung ist, d.h. ω selbst ist eine Ordinalzahl (und nicht ‘bloß’ eine Menge). Das erklärt auch, wieso wir hier das Zeichen ‘ ω ’ nicht bloß für den *Ordnungstyp* der natürlichen Ordnung auf den natürlichen Zahlen verwenden, sondern *für die Menge selbst*.⁵⁶ Da ω offenbar nicht Nachfolger einer Ordinalzahl ist, führt uns das zu folgender

Definition 6.9. Für Ordinalzahlen α, β gilt:

1. α ist eine *Nachfolgerordinalzahl* gdw. $\exists \beta(\alpha = S(\beta))$
2. α ist eine *Limesordinalzahl* gdw. α keine Nachfolgerordinalzahl ist.

Wie in NFU für $\mathcal{N}$ können wir auch hier zeigen, dass ω die Peano-Axiome 1 bis 5 aus dem vorigen Abschnitt erfüllt.

Mit einer geeigneten Definition der ordinalen Addition und Multiplikation kann überdies gezeigt werden, dass die Struktur $(\omega, +_o, \cdot_o)$ alle Axiome aus dem vorigen Abschnitt erfüllt. Diese Definitionen funktionieren im Wesentlichen wie jene in NFU mit dem Unterschied, dass wir in ZFC ‘direkt’ vorgehen können. In NFU mussten wir ja, weil Ordinalzahlen Äquivalenzklassen waren, zur Definition von $\alpha +_o \beta$ Repräsentanten verwenden.⁵⁷ Nun, da wir de facto Repräsentanten als Ordinalzahlen verwenden (vgl. Satz 6.3.) fallen mögliche ‘Umwege’ weg und wir können die Operationen, die wir in NFU für Repräsentanten erklärt haben für die Ordinalzahlen selbst verwenden.

Ganz allgemein können wir mit dieser Definition der ordinalen Addition und Multiplikation nun allgemeine Sätze über Ordinalzahlen (etwa Satz 5.8. aus dem vorigen Abschnitt) beweisen.

Wie auch in NFU können wir in ZFC das *Prinzip der transfiniten Induktion* beweisen, das entscheidend dazu benötigt wird, um rekursive Definitionen über die Ordinalzahlen zu ermöglichen. Rekursionen über die Ordinalzahlen spielen

⁵⁶In NFU hatten wir dafür - aus gutem Grund - zwei verschiedenen Bezeichnungen: $\mathcal{N}$ für die (‘bloße’) Menge aller natürlichen Zahlen und ω für den *Ordnungstyp* der natürlichen Ordnung $<_{\mathcal{N}}$ auf den natürlichen Zahlen - zwei verschiedene Objekte.

⁵⁷Das heißt, wir haben für beliebige Wohlordnungen $R_A \in \alpha$ und $S_B \in \beta$ irgendeine Operation erklärt.

eine nicht zu überschätzende Rolle bei der Strukturierung unseres Universums - wie wir im Zusammenhang mit dem iterativen Begriff der Menge schon gesehen haben, wo wir die Folge der Von-Neumannschen Stufen über transfinite Rekursion definiert haben.

Wir werden später noch sehen, dass das für NFU nicht unbedingt gilt.

6.3 Kardinalzahlen: Grundlegende Begriffe

Wie in NFU sollen auch hier Kardinalzahlen als Maßzahlen für die Größe von Mengen fungieren. Wir lassen uns wie in NFU von der Vorstellung leiten, dass zwei Mengen gleich groß sind, wenn sie eineindeutig aufeinander abbildbar sind.

Wir übernehmen also die Definitionen der Begriffe ‘höchstens so mächtig’ ($\preceq$), ‘gleichmächtig’ ($\approx$) und ‘schmächtiger’ ($\prec$) aus dem vorigen Abschnitt.

Da keiner der darauf folgenden Sätze darauf Bezug nimmt wie Kardinalzahlen definiert sind, heißt das, dass wir auch die Sätze 5.1. und 5.2. bzgl. der Ordnungs- bzw. Äquivalenzeigenschaften von $\preceq$ ($\prec$) bzw. $\approx$ übernehmen können.

Bevor wir uns an die Definition der Kardinalität einer Menge machen, noch einige Beobachtungen: Für alle endlichen Ordinalzahlen n gilt offenbar: $n \approx m \rightarrow n = m$, d.h. gleichmächtige natürliche Zahlen sind gleich. Das muss aber nicht für unendliche Ordinalzahlen gelten. So ist z.B. $\omega \approx \omega + \omega$ aber es ist offenbar *nicht* $\omega = \omega + \omega$. Mit diesem Beispiel im Hinterkopf legen wir fest:

Definition 6.10. *Für eine Menge A ist $|A|$ (sprich: die Kardinalität von A) die kleinste Ordinalzahl α , für die $A \approx \alpha$.*

Diese Definition ist insofern für *alle* Mengen gerechtfertigt, als sich (dank des Auswahlaxioms) *jede* Menge wohlordnen lässt, und daher jede Menge zu irgendeiner Ordinalzahl isomorph (und also gleichmächtig) ist.

Die Intuition hinter dieser Definition besagt ungefähr folgendes: Wenn wir die Größe einer Menge bestimmen wollen, dann zählen wir einfach ihre Elemente (ordinaler Aspekt); die kleinste Ordinalzahl, die für diesen Zählprozess ausreicht, nennen wir dann die Kardinalität dieser Menge.

Unabhängig vom Auswahlaxiom ist für Ordinalzahlen $|\alpha|$ immer definiert und wir bestimmen das Prädikat, eine Kardinalzahl zu sein durch

Definition 6.11. *Für eine Ordinalzahl α gilt $\text{Kard}(\alpha) :\leftrightarrow |\alpha| = \alpha$*

Aus dieser Definition sehen wir etwa, wieso $\omega + \omega$ keine Kardinalzahl ist, während ω und alle natürlichen Zahlen $n \in \omega$ schon welche sind. Offenbar gilt nämlich für jede natürliche Zahl $|n| = n$, denn alle kleineren natürlichen Zahlen sind nicht mehr gleichmächtig zu n . Wie angekündigt rechtfertigt das den bestimmten Artikel bei der Rede von *den* natürlichen Zahlen. Für Kardinalzahlen verwenden wir in Zukunft wieder kleine griechische Buchstaben $\kappa, \lambda, \mu \dots$

Die kleiner-größer-Relation $<$ können wir nun festlegen durch

Definition 6.12. $\kappa < \lambda :\leftrightarrow \kappa \prec \lambda$ und
 $\kappa \leq \lambda :\leftrightarrow \kappa \preceq \lambda$

Auch kardinale Addition und Multiplikation können wir in naheliegender Weise definieren durch

Definition 6.13. $\kappa + \lambda := |\kappa \times \{0\} \cup \lambda \times \{1\}|$ und $\kappa \cdot \lambda := |\kappa \times \lambda|$

Definieren wir zusätzlich $\aleph_0 := \omega$, dann können wir ohne Probleme die Sätze 5.13. und 5.14. aus dem vorigen Abschnitt beweisen. Wie in NFU nennen wir Mengen der Mächtigkeit $\aleph_0$ abzählbar unendlich oder einfach abzählbar. Und wie auch in NFU wissen wir bis jetzt noch nicht, ob es größere Kardinalzahlen gibt als $\aleph_0$, d.h. wir wissen nicht, ob es Mengen gibt, die größer sind als jede abzählbare Menge. Anders aber als in NFU bereitet uns der Beweis dieser Tatsache in ZFC keine Probleme. Es gilt

Satz 6.6. (Satz von Cantor) $\forall A (A \prec \mathcal{P}(A))$

Der Beweis - nun ohne Anführungszeichen - für diesen Satz wurde im vorigen Abschnitt schon gegeben und geht so wie er da steht durch. Das Potenzmengenaxiom, das uns für beliebige Mengen deren Potenzmenge garantiert, garantiert uns mit dem Satz von Cantor demnach ebenfalls, dass es zu *jeder* Menge eine noch größere gibt. Die Reihe der Kardinalzahlen kennt keine Grenzen und die Gesamtheit der Kardinalzahlen ist - wie auch die der Ordinalzahlen - wenig überraschend eine echte Klasse.

Da wir nun auf Stratifizierung und derartiges keine Rücksicht mehr nehmen müssen können wir - den Beobachtungen aus dem vorigen Abschnitt folgend - die kardinale Exponentiation definieren durch

Definition 6.14. $\kappa^\lambda := |\kappa^\lambda|$

Außerdem gilt folgender Satz (der *nicht* in NFU gilt):

Satz 6.7. Für alle Kardinalzahlen κ ist $2^\kappa = |\mathcal{P}(\kappa)|$

Wir wollen hier den kurzen Einblick in die Arithmetik der Kardinal- und Ordinalzahlen abbrechen und uns die beiden Systeme noch einmal im Vergleich ansehen.

7 Mengenlehre und Arithmetik: Ein Vergleich

Wir konnten an einigen Stellen schon beobachten, dass NFU und ZFC ziemlich verschiedene Vorstellungen von den Zahlen und den Sätzen, die für sie gelten, haben. Die Zahlen in ZFC sehen völlig anders aus als die in NFU; das Paradox von Burali-Forti wird auf verschiedene Weisen vermieden; der Satz von Cantor gilt in ZFC, nicht aber in NFU. Wir werden versuchen im Folgenden einiges davon näher zu beleuchten.

7.1 Was sind und was sollen die Zahlen?

Wie man in den vergangenen Abschnitten sehen konnte, werden Ordinal- und Kardinalzahlen in NFU und ZFC auf verschiedene Weisen definiert. Dabei stellt sich die natürliche Frage, ob das so sein *muss*. Gibt es Alternativen? Oder wird durch die Art des Systems auf irgendeine Weise ‘festgelegt’, wie die Zahlen auszusehen haben?

Nehmen wir z. B. die natürlichen Kardinalzahlen; Eine Art die Dinge zu sehen ist folgende: Im Rahmen der Mengenlehre, in der wir uns gerade befinden, ist es möglich, verschiedene Mengen zu bilden (Singletons, Vereinigungen, Potenzmenge usw.). Wenn es uns auf irgendeine Weise gelingt, die Zahl 0 zu definieren und weiters eine Operation S derart, dass 0 und S zusammen mit einer irgendwie definierten Addition und Multiplikation die Peano-Axiome erfüllen, so nennen wir die dadurch festgelegte Struktur im jeweiligen System ‘die’ natürlichen Zahlen. Oder wie Quine es formuliert:

“Es steht in unserem Belieben, als 0 irgend etwas, was wir gern möchten, zu nehmen und S als irgendeine Funktion nach unserem Geschmack zu definieren, sofern diese Funktion die Eigenschaft hat, dass sie, iteriert auf 0 angewandt, bei jeder Anwendung etwas Neues ergibt.”⁵⁸

Quine betont in obigem Zitat einen Umstand, der charakteristisch für die Mengenlehre im Allgemeinen ist: Was wir tun, wenn wir bestimmte Mengen mit bestimmten mathematischen Objekten - z. b. natürlichen Zahlen - ‘identifizieren’, kann man vergleichen mit dem Implementieren eines ‘Programms’ durch irgendeine Hardware.

Der Hardware entsprechen in dieser Analogie die im jeweiligen System gegebenen Mengen (für die 0 in ZFC z. B. die leere Menge) und das Programm oder die ‘Funktion’ ist das, was diese Mengen *leisten* müssen, nämlich die Peano-Axiome für die natürlichen Zahlen zu erfüllen.

In gewisser Weise kommt diese Strategie einem ‘Ausweichen’ auf die Frage: “Was sind die Zahlen?” gleich. Man zieht sich auf den ‘strukturellen’ Standpunkt zurück, und sagt: Alle Strukturen desselben Typs wie die naive Struktur $(N, 0, S, +, \cdot)$, die die Peano-Axiome erfüllen, taugen gleich gut als ‘die’

⁵⁸Quine [1973] S. 60

natürlichen Zahlen. Eine bestimmte Implementation der natürlichen Zahlen in einer bestimmten Mengenlehre muss einfach bringen, was man von ihr erwartet.

Sollte man eine ‘Antwort’ dieser Art *nicht* als Antwort auf die Frage “Was sind die Zahlen?” akzeptieren, muss man sich überlegen, was man überhaupt als Antwort gelten lassen würde. Geht man diesen Weg weiter, kommt man nämlich schnell in die Situation, dass man als Antwort nur mehr mit “Die Zahlen sind irreduzibel” oder “Die Zahlen sind Ideen im platonischen Himmel” zufrieden ist.

Im Prinzip handeln die Frage “Was sind die natürlichen Zahlen?” und die Antwort: “Die natürlichen Zahlen sind die Struktur $(\omega, 0, S, +, \cdot)$, derart, dass gilt...” (im Falle von ZFC) nämlich gar nicht vom gleichen Thema. Im Rahmen einer Mengenlehre kann nämlich die Antwort auf *jede* Frage “Was ist ...?” *immer* nur lauten: “Die Menge...” oder “Das geordnete Paar...”, weil es außer Mengen keine anderen Objekte gibt.

Fragen nach dem ‘Wesen’ oder der ‘Wirklichkeit’ der natürlichen Zahlen haben hier einfach keinen Platz.

Trotzdem suggeriert das Quine-Zitat von oben mehr Freiheit bei der Implementation als man in der Praxis in einem bestimmten System tatsächlich hat. Wie wir sehen konnten, war die natürliche Zahl 1 in NFU definiert als die Menge aller Singletons $\mathcal{P}_1(V)$. Dass diese Gesamtheit in ZFC keine Menge definiert, ist offensichtlich, denn dann würde folgen, dass auch V eine Menge wäre, was in ZFC nicht möglich ist.

Umgekehrt ist es in NFU nicht ohne weiteres möglich, die Menge aller natürlichen Von Neumann Zahlen ω zu bilden.⁵⁹ Zur Definition dieser Menge braucht man nämlich die induktive und nicht-stratifizierte Bedingung $n \in \omega \rightarrow n \cup \{n\} \in \omega$.

Solche Einschränkungen kommen nicht von ungefähr und es stellt sich die Frage nach den Gründen. Positiv formuliert: Warum stehen verschiedenen Systemen verschiedene Umsetzungen der (natürlichen) Zahlen näher als andere? Das heißt, gibt es eine Möglichkeit für die Bevorzugung einer bestimmten Variante zu argumentieren, oder sind wir in dieser Hinsicht, wie Quine es nahe legt, der Beliebigkeit ausgeliefert?

In dieser Hinsicht scheint NFU mit seinen Frege-Russell-Zahlen klar im Vorteil zu sein: Wie schon angesprochen wurde - und später noch ausführlicher behandelt wird - gibt es in NFU eine starke Verbindung zwischen Prädikation und Elementrelation. Die Intuition hinter den natürlichen Zahlen nach Frege-Russell ist ja, dass alle Mengen mit einer bestimmten Anzahl von Dingen etwas ‘gemeinsam’ haben. Alle n -elementigen Mengen teilen das Prädikat ‘hat die Anzahl n ’. Der engen Anbindung von Prädikation und Elementrelation folgend, interpretiert man das Prädikat ‘hat die Anzahl n ’ selbst als Menge und die Aus-

⁵⁹Noch weniger die Menge aller Von Neumann *Ordinalzahlen*.

sage ‘Die Menge A hat n Elemente’ wird gelesen als ‘Die Menge A ist Element der Menge aller n -elementigen Mengen’.⁶⁰

In ZFC ist scheint es dagegen eher so zu sein, dass man als ‘Zahlen’ das nimmt, was eben am besten *funktioniert*. Auf den ersten Blick wirkt es beinahe so, als habe man sich - dem Quine’schen Rat folgend - *irgendetwas* als natürliche Zahlen herausgepickt; und es hat sich eben herausgestellt, dass man mit diesen Zahlen ‘gut arbeiten’ kann. Es scheint, also wäre die einzige Rechtfertigung eine *pragmatische*. Dennoch: Gerade die pragmatischen Gründe, die für Von Neumann Zahlen sprechen sind vielfältig.

Ein Beispiel wäre, dass die Größer-Kleiner- Beziehung auf den Von Neumann-Zahlen einfach die Elementbeziehung selbst ist. Für Ordinalzahlen gilt allgemein, dass eine Ordinalzahl α kleiner als eine Ordinalzahl β ist, wenn α *Element von* β ist. Es sind keine zusätzlichen Definitionen über Fortsetzungen und dergleichen notwendig.

Ein weiterer Grund, der für Von Neumann-Zahlen spricht, ist, dass es ziemlich einfach ist Kardinalzahlen im Allgemeinen mit Ordinalzahlen zu *identifizieren* (und nicht bloß zu *korrelieren*)⁶¹.

Ein Vorteil gegenüber den Frege-Russell Zahlen liegt auch in folgendem Umstand: Bei den natürlichen Zahlen in NFU haben wir ja gesehen, dass, wenn wir irgendetwas mit ihnen anstellen wollten (Definitionen oder Beweise), *Repräsentanten* benutzen mussten. So konnten wir die Addition und Multiplikation $n + m$ und $n \cdot m$ nur definieren über Repräsentantenmengen $A \in n$ und $B \in m$. Ähnlich erging es uns mit Ordinalzahlen, wo wir immer mit *bestimmten* (wenn auch beliebigen) Wohlordnungen arbeiten mussten; genauso bei der Definition der kardinalen Exponentiation.

In ZFC haben wir diese Probleme nicht: Hier arbeiten wir *de facto* mit Repräsentanten. Jede Ordinalzahl *ist* eine Menge mit einer bestimmten darauf erklärten Wohlordnung.

Dennoch: Es gibt eine auch *inhaltliche* Rechtfertigung, die für die Von Neumann’schen Ordinal-zahlen spricht.

Was bei der Definition der Ordinalzahlen nach Von Neumann passiert, ist ja folgendes: Man lässt sich von der ‘intuitiv einsichtigen’ Beobachtung leiten, dass, gegeben eine Ordinalzahl α (im ‘naiven’ Sinn), immer gilt: Der Ordnungstyp der Menge aller Ordinalzahlen, die kleiner als α sind, zusammen mit der

⁶⁰Diese Motivation ist intuitiv sogar so überzeugend, dass manche Autoren meinen, sich damit so beschäftigen zu müssen, als wäre damit intendiert, das *Wesen der Zahl* zu erfassen. Vgl. dazu Benacerraf [1965]

⁶¹Wie man in NFU unendliche Kardinalzahlen mit bestimmten Ordinalzahlen *korreliert*, werden wir im Kapitel ‘Die Struktur des Unendlichen’ noch sehen. Für natürliche Kardinal- und Ordinalzahlen ist die Korrelation offensichtlich.

üblichen Ordnung auf den Ordinalzahlen ist eine Wohlordnung vom Ordnungstyp α , wobei der Ordnungstyp relational über die Beziehung der Ordnungsisomorphie definiert ist. Aus dieser Beobachtung ergibt sich nun die Definition der Von-Neumann Ordinalzahlen einfach dadurch, dass man sagt: Die Menge aller Ordinalzahlen kleiner als α ist die Ordinalzahl α selbst! Fasst man Ordinalzahlen in dieser Weise auf, folgen daraus unmittelbar die Eigenschaften, die wir oben zur formalen Definition der Ordinalzahlen in ZFC benutzt haben (Ordinalzahlen sind genau die transitiven Wohlordnungen bzgl. der $\in$ -Relation).⁶²

Das Problem ist: Diese ‘Rechtfertigung’ nimmt nicht unbedingt direkten Bezug auf das System ZFC und *dessen* Motivation (die, wie wir schon gesehen haben und noch sehen werden, vor allem im ‘iterativen Begriff der Menge’ liegt), d.h. es gibt keine ‘intrinsische’ Motivation die für die Verwendung von Von Neumann Zahlen spricht.

7.2 Unendlichkeit

In jeder Mengenlehre, die beansprucht, eine ‘Grundlage’ für die Mathematik zu sein, muss es möglich sein, mathematische Objekte wie die Menge der natürlichen Zahlen, die Menge der reellen Zahlen usw. zu repräsentieren. Dafür braucht man Mengen mit unendlich vielen Elementen. Um die Existenz von unendlichen Mengen zu garantieren braucht man sowohl in ZFC als auch in NFU entweder ein explizites Axiom oder es muss durch die anderen Axiome irgendwie sichergestellt werden, dass es unendliche Mengen gibt.

In ZFC stellt ein eigenes Axiom sicher, dass es zumindest *eine* induktive Menge gibt, also eine Menge, die die leere Menge enthält und die bezüglich der von Neumann Nachfolger-Operation abgeschlossen ist. Dass ein eigenes Axiom notwendig ist, folgt aus dem Umstand, dass uns zwar *potentiell* unendlich viele Mengen gegeben sind, wir diese potentiell unendlich vielen Mengen aber nicht zu einer Menge ‘zusammenfassen’ können, sodass es nicht möglich ist, die Existenz *irgendeiner* induktiven Menge zu garantieren, wenn man nicht schon eine unendliche Menge *gegeben* hat. ‘Zusammenfassen’ könnten wir diese potentiell unendlich vielen Mengen (z.b. die Mengen $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \dots$) nur mit einem unbeschränkten Komprehensionsprinzip. Genau dieses Prinzip wurde aber als inkonsistent enttarnt, und war einer der wesentlichen Gründe für die Entwicklung des Systems von Zermelo mit seinem Aussonderungsschema.

In NFU hat man *dieses* Problem nicht, aber eine andere Schwierigkeit, die mit dem NFU-Pendent zum Aussonderungsschema, dem S-Komprehensionsschema, zu tun hat.

Wie in ZFC gibt es auch in NFU potentiell unendlich viele Mengen wie die leere Menge, das Singleton mit einzigem Element $\emptyset$ usw. Der Grund dafür, dass wir diese potentiell unendlich vielen Mengen nicht in eine Menge stecken können, liegt darin, dass die Definition einer derartigen Menge nicht stratifiziert sein kann.

⁶²Vgl. dazu Deiser [2002] S. 251 und S. 257 ff.

Dass eine induktive Menge existiert, ist auch in NFU nicht sichergestellt, nur weil wir (wie in ZFC) keine größere Menge brauchen, aus der wir aussondern müssten. ‘Induktiv’ heißt ja in NFU, abgeschlossen bzgl. des Frege-Russell Nachfolgers, und wie wir in Zusammenhang mit der Definition der natürlichen Zahlen in NFU sehen konnten, kann es vorkommen, dass einem schlicht die Elemente desselben ‘Typs’ ausgehen.

In der hier präsentierten Variante von NFU stellt die Tatsache, dass wir für das geordnete Paar angenommen haben, dass es ‘type-level’ ist, sicher, dass es unendlich viele Dinge desselben ‘Typs’ gibt. ‘Type-level’ heißt ja einfach nur, dass die rechte und linke Projektion eines geordneten Paares denselben relativen Typ hat wie das geordnete Paar selbst. Sei etwa $z = (x, y)$ und pr_1 bezeichne die linke und pr_2 die rechte Projektion, dann ist der relative Typ von $pr_1(z) = x$ gleich dem Typ von $pr_2(z) = y$ gleich dem Typ von z selbst. Wir können nun über vollständige Induktion zeigen, dass für alle $n \in \mathcal{N}$ gilt: $A \in n \rightarrow A \times \{0\} \in n$. Mit dieser Beobachtung im Hinterkopf können wir zeigen, dass jeder Nachfolger einer nichtleeren natürlichen Zahl n selbst wiederum nichtleer ist, indem wir zeigen, dass für alle $n \in \mathcal{N}$ aus $A \in n$ immer folgt, dass $A \times \{0\} \cup \{(0, 1)\} \in n + 1$.⁶³

Wie wir sehen konnten, sind beide betrachteten Systeme - aus verschiedenen Gründen - nicht in der Lage aus den potentiell unendlich vielen Mengen eine ‘Gesamtheit’ zu erzeugen, die selbst unendlich ist. Zweierlei Betrachtungen sind zum Verständnis dieses Problems hilfreich: Erstens die bekannte philosophische Unterscheidung zwischen *aktual* und *potentiell* Unendlichem, und zweitens das Modell der sogenannten *erblich-endlichen* Mengen.

Die Unterscheidung zwischen dem aktual und dem potentiell Unendlichen bezeichnet grob gesagt folgende Positionen zum Unendlichen: Erstere besagt, dass es unendliche ‘Objekte’ oder ‘Gesamtheiten’ gibt, die uns *fertig* gegeben sind. Die zweite Sichtweise widerspricht genau dieser Auffassung vom Unendlichen. Für diese Sichtweise ist das wesentliche am Unendlichen nicht, dass sie eine Gesamtheit bildet, sondern, dass es mit den ‘neuen’ Dingen einfach *kein Ende* nimmt. Ein berühmter Exponent dieser Auffassung ist Gauss:

“So protestiere ich zuvörderst gegen den Gebrauch einer unendlichen Größe als einer vollendeten, welches in der Mathematik niemals erlaubt ist. Das Unendliche ist nur eine facon de parler, indem man eigentlich von Grenzen spricht, denen gewisse Verhältnisse so nahe kommen als man will, während andere ohne Einschränkung zu wachsen gestattet ist.”⁶⁴

Es liegt auf der Hand, dass Mengenlehre ganz allgemein mit so einer Position nichts anfangen kann; Mengenlehre *ist de facto* die *Lehre vom (aktual)*

⁶³Vgl. dazu Holmes [1998] S. 82

⁶⁴Zitiert nach Deiser [2002] S. 24

Unendlichen. Alle ‘wichtigen’ Objekte, die im Zusammenhang mit der Arithmetik behandelt wurden, wie die Menge der natürlichen Zahlen in ZFC oder die Menge der Ordinalzahlen in NFU, *sind* ‘vollendete Größen’.

In dieser Terminologie kann man das Problem, das ein Unendlichkeitsaxiom lösen soll, so reformulieren: Gibt es eine *aktual* unendliche Menge im Mengenumversum? (Im Gegensatz zur Frage: Enthält das Universum potentiell unendlich viele Mengen?)

Man kann sich das auch anhand eines einfachen Modells von ZFC ohne Unendlichkeitsaxiom begreiflich machen: Wir sehen uns dazu die ω -te Von Neumann Stufe R_ω , also die Menge, die genau aus den Mengen $\emptyset, \mathcal{P}(\emptyset), \mathcal{P}(\mathcal{P}(\emptyset)), \dots$ besteht, an. Man kann sich leicht davon überzeugen, dass jedes Axiom von ZFC, bis auf das Unendlichkeitsaxiom, in diesem Modell erfüllt ist. Während diese Menge selbst eine aktual unendliche ist, so sind doch ihre Elemente - also das, worüber quantifiziert wird - allesamt *endlich*. Für dieses Modell muss man obige Fragen demnach so beantworten: Ja. Das Universum enthält potentiell unendlich viele Mengen. Aber Nein. Das Universum enthält keine aktual unendliche Menge.⁶⁵

7.3 Das Paradox von Burali-Forti

Das Paradox von Burali-Forti ist eine der ersten Antinomien der ‘naiven’ Mengenlehre und als es das erste mal (in einer falschen Form von Cesare Burali-Forti) explizit zur Sprache gebracht wurde, schlug es große Wellen. Es lohnt sich genauer anzusehen, was eigentlich das Paradoxe an diesem Paradox ist, oder in Form einer Frage: Was genau, welcher Begriff der Menge und der Ordinalzahl, welche Sätze, die man allgemein akzeptabel findet, führen zur Antinomie? Die Antworten auf diese Fragen werfen auch ein Licht auf die Verschiedenheit zwischen NFU und ZFC.

Wir werden uns dieser Frage von der intuitiven Seite nähern, und dann versuchen herauszuarbeiten, wie das in den betrachteten Systemen NFU und ZFC aussieht.

Das Paradox ergibt sich unter naiv-mengentheoretischen Konditionen aus der Zusammenfassung aller Ordinalzahlen zu einer Menge. Ordinalzahlen fasst man auf als Ordnungstypen von Wohlordnungen. Cantor formuliert das so:

“... hat jede einfach geordnete Menge M einen bestimmten Ordnungstypus $\bar{M}$; es ist dies der Allgemeinbegriff, welcher sich aus M ergibt, wenn unter Festhaltung der Rangordnung ihrer Elemente von der Beschaffenheit der letzteren abstrahiert wird, so dass aus ihnen lauter Einsen werden, die in einem bestimmten Rangverhältnis zu einander stehen. Allen einander ähnlichen Mengen, und nur solchen,

⁶⁵So wie die Menge aller natürlichen Zahlen unendlich ist, aber jede auch noch so große natürliche Zahl selbst endlich ist.

kommt ein und derselbe Ordnungstypus zu. Den Ordnungstypus einer wohlgeordneten Menge F nenn wir die ihr zukommende ‘Ordnungszahl’.”⁶⁶

Cantor spricht hier von einem ‘Allgemeinbegriff’, der sich durch einen Prozess der ‘Abstraktion’ ergibt. Es ist nicht ganz klar, was das bedeuten soll, aber der nächste Satz im Zitat legt nahe, dass es sich dabei um eine ähnliche Auffassung der Ordinalzahlen wie die für NFU präsentierte handelt.⁶⁷

Die natürliche Ordnung auf dieser Menge, $0 < 1 < 2 < \dots < \omega_0 < \omega_1 < \dots \gamma < \dots$, ist eine Wohlordnung, muss also einen Ordnungstyp haben. Der Ordnungstyp der Menge aller Ordinalzahlen mit der üblichen Ordnung ist also selbst eine Ordinalzahl (wir nannten sie Ω) und daher in der Menge ORD selbst wieder vertreten, gleichzeitig aber größer als alle Ordinalzahlen.

Eine erste (korrekte) Formulierung liefert Cantor in seinem berühmten Brief an Dedekind von 1899:

“Es kann Ω' (und daher auch Ω) keine konsistente Vielheit sein: wäre Ω' konsistent, wo würde ihr als einer wohlgeordneten Menge eine Zahl δ zukommen, die größer wäre als alle Zahlen des Systems Ω ; im System Ω kommt aber, weil es alle Zahlen umfasst, auch die Zahl δ vor; es wäre also δ größer als δ , was ein Widerspruch ist, Also

A. Das System Ω aller Zahlen ist eine inkonsistente, eine absolut unendliche Vielheit.”⁶⁸

wobei Cantor unter einer ‘inkonsistenten’ oder ‘absolut-unendlichen Vielheit’ eine Vielheit versteht, deren Zusammenfassung zu einem Ganzen zum Widerspruch führt; eine konsistente Vielheit, oder auch ‘Menge’ nennt er eine solche, die man ohne Widerspruch zusammengefasst denken kann.

Es ist nützlich, sich anzusehen, was an diesem Argument von Cantor man - als notwendig für Mengen und Ordinalzahlen - akzeptieren muss, und was man dispensieren kann. An entscheidenden Stellen kommt es nämlich darauf an

1. was man mit diesem Argument überhaupt zeigen möchte
2. wie man die Begriffe ‘Menge’ und ‘Ordinalzahl’ auffasst
3. und welche Sätze über Ordinalzahlen man behalten will und welche man bereit ist, aufzugeben.

⁶⁶Cantor in Cantor [1895], § 14

⁶⁷Sei es wie es sei: Hier arbeiten wir naiven Ordinalzahlen. Es kann uns also egal sein, wie Cantor Ordinalzahlen ‘definiert’ haben würde. Manches spricht dafür, dass er sie gar nicht ‘definieren’ wollte, sondern als primitive Objekte oder Universalien, betrachtete. Vgl. dazu auch Wang [1974], wiederabgedruckt in ‘Philosophy of mathematics’ S. 539.

⁶⁸Es ist hier zu beachten, dass Cantor mit Ω die Gesamtheit aller Ordinalzahlen meint - wir hingegen haben darunter den Ordnungstyp der Menge aller Ordinalzahlen mit der üblichen Ordnung verstanden. Vgl. Cantor [1899]

Aus dem ZFC-Blickwinkel ist die Interpretation dieser Stelle bei Cantor *straightforward*: Wir ersetzen ‘inkonsistente Vielheit’ durch ‘echte Klasse’ und das Argument verwandelt sich in einen Beweis, dass die Menge aller Ordinalzahlen nicht existiert, dass sie eine echte Klasse ist.

Um das Cantor’sche Argument in NFU zu interpretieren, muss man sich ins Gedächtnis rufen, dass in NFU - anders als in ZFC und auch anders als für Cantor - nicht immer gilt, dass der Ordnungstyp eines durch α bestimmten Anfangsstücks der Ordinalzahlen mit der üblichen Ordnung gleich α selbst ist. Im Gegenteil: das ‘Burali-Forti-Argument’⁶⁹ zeigt gerade, dass das für den Ordnungstyp Ω der Wohlordnung auf der Menge *aller* Ordinalzahlen *nicht* gilt (falls wir einen Widerspruch vermeiden wollen).

Das ‘Burali-Forti-Argument’ zeigt also nicht, dass ‘das System Ω (ORD in unserer Notation) eine inkonsistente, eine absolut unendliche Vielheit’ ist (was in NFU ja einen Widerspruch bedeuten würde, weil ORD eine stratifizierte Definition hat und daher existiert), sondern widerlegt gewissermaßen die ‘Motivation’ für die Von-Neumannschen Ordinalzahlen.

Insbesondere zeigt das Argument, dass man in sinnvoller Weise ordinale Arithmetik betreiben kann ohne den für die Von Neumann’schen Zahlen charakteristischen Sachverhalt *o.t.* $(\prec_{Pred(\alpha, \prec)}) = \alpha$ (und den Cantor für selbstverständlich zu halten schien). Man kann sich die Ordinalzahlen in NFU in dieser Hinsicht einfach als *verallgemeinerte* Ordinalzahlen vorstellen. D.h. viele ‘kleine’ (die *Cantor’schen*) Ordinalzahlen verhalten sich ‘gesittet’, während NFU doch auch ‘große’ Zahlen (wie Ω) zulässt, die das *nicht* tun.

Analoges gilt auch für die Kardinalzahlen in NFU, die man als verallgemeinerte Kardinalzahlen bzgl. des Satzes von Cantor betrachten kann. Die ‘kleinen’ Kardinalzahlen sind dabei die Kardinalzahlen jener Mengen A , für die der Satz von Cantor $A \prec \mathcal{P}(A)$ gilt, die ‘großen’ (wie $|V|$) sind jene, für die er *nicht* gilt.

7.4 Induktive Definitionen, Indizierungen und lange Wohlordnungen

Ein Punkt, in dem sich ZFC und NFU stark voneinander unterscheiden, betrifft die Rolle der Ordinalzahlen in einem weiteren Sinn.

In ZFC spielen die Ordinalzahlen nicht bloß eine Rolle in der Arithmetik wie in NFU, sondern wirken gewissermaßen ‘strukturierend’ auf das gesamte durch die Axiome beschriebene Mengenuniversum. Die Ordinalzahlen in ZFC sind z.B. (und *vor allem*) insofern ‘tragend’, als sie die *Achse* bilden, um die sich das Mengenuniversum dreht; ein Bild, das man - mit der Von Neumannschen Hierarchie im Hinterkopf - beinahe wörtlich verstehen kann, denn *Existenz* einer Menge x fällt in ZFC mit der Existenz einer Ordinalzahl α zusammen, sodass $x \in R_\alpha$. Das entscheidende Mittel, um unserem Mengenuniversum diese Form geben zu können, war die Möglichkeit *beliebig lange Folgen* über transfinite

⁶⁹Holmes nennt das Argument, dass durch *reductio ad absurdum* zeigt, dass $\Omega < T^2(\Omega)$ ist, so.

Rekursion über die Ordinalzahlen definieren zu können.⁷⁰

Wir wollen uns das genauer ansehen.

Schon in der Schule wird man mit Folgen vertraut gemacht. Meistens behandelt man dort Folgen natürlicher oder reeller Zahlen. Man schreibt etwa für eine reelle Folge $(a_n)_{n \in \mathbb{N}}$ und meint damit eine Funktion $f : \mathbb{N} \rightarrow \mathbb{R}$. Für einen bestimmten Wert n einer Funktion schreibt man dann statt $f(n)$ üblicherweise a_n . Eine rekursive Folge (d.h. eine Folge, die über Rekursion über die natürlichen Zahlen definiert ist) ist nun eine solche, wo uns die Funktionswerte nicht ‘alle auf einmal’ gegeben sind, sondern uns nur eine ‘Anleitung’ gegeben ist, wie wir den Funktionswert an der Stelle $n + 1$ berechnen können, gesetzt den Fall, wir haben den Wert für n schon berechnet. ‘Induktiv’ folgt dann, dass so eine Folge für *alle* natürlichen Zahlen definiert ist.

‘Mathematisch’ (d.h. wenn man von der ‘formalen’ Verschiedenheit absieht) ist die Situation für die Methode der transfiniten Rekursion über die Ordinalzahlen völlig analog:⁷¹ Mit der einzigen zusätzlichen Bedingung, dass man, um den Wert dieser ‘Funktion’ an einer bestimmten Stelle β bestimmen zu können, die Werte aller Vorgänger benötigt, weil nicht alle Ordinalzahlen Nachfolgerzahlen sind (wie die natürlichen Zahlen).

‘Formal’ jedoch müssen wir in ZFC aber (zumal im Vergleich mit NFU) streng zwischen den beiden Rekursionsprinzipien unterscheiden; während das eine über eine *Menge* (die Menge der natürlichen Zahlen) und *Funktionen* spricht, behandelt das andere *Klassen* (nämlich *ON*) und *Operationen* (zweistellige funktionale Prädikate). Das Rekursionsprinzip ist ein beweisbarer Satz in ZFC, das Prinzip der transfiniten Rekursion ein Schema, das unendlich viele Sätze umfasst.⁷²

Am Beispiel der Von Neumannschen Hierarchie kann man das sehr gut beobachten: Wir definieren dabei über die Potenzmengenoperation (bzw. die Vereinigungsoperation im Falle der Limesordinalzahlen) die $\alpha + 1$ -te Stufe, gesetzt den Fall wir haben alle vorherigen Stufen schon definiert; und wieder schreiben wir die Werte der durch diese Operationen entstehenden Operation auf *ON* als Indizes.

Dass die Methode der transfiniten Rekursion für ZFC so brauchbar ist, liegt vor allem daran, dass es so viele Ordinalzahlen gibt⁷³: Es kann keine Objekte geben, die nicht durch Ordinalzahlen ‘erreichbar’ wären. Jedes noch so große

⁷⁰Ein Umstand, den wir wiederum dem Ersetzungsschema verdanken.

⁷¹Zur Erinnerung: Das Prinzip der transfiniten Rekursion (das aus dem Prinzip der transfiniten Induktion folgt) gestattet uns rekursive Definitionen über die Ordinalzahlen, d.h. wenn wir z.B. eine Folge von Objekten definieren wollen, können wir das tun, indem wir eine Anfangsbedingung angeben und weiters erklären, wie wir über eine Operation (bzw. eine *Funktion* in NFU) zu weiteren Objekten kommen, falls die Folge für alle Vorgänger schon definiert ist.

⁷²Für Rekursion über die natürlichen Zahlen stimmt ersteres für beide Systeme. Für die Ordinalzahlen gilt das offenbar nicht: In NFU ist auch der Satz über die transfinite Rekursion über die Ordinalzahlen ein beweisbarer Satz und kein *Schema*.

⁷³Vgl. dazu Satz 6.5. (Burali Forti) und Satz 6.4. (Satz von Hartogs)

Objekt ist rekursiv ‘abtragbar’.

In NFU hingegen haben wir zweierlei Probleme mit rekursiven Definitionen und Indizierungen:

Erstens müssen wir bei Definitionen über ORD (wie bei jeder Definition) darauf achten, dass unsere Bedingung stratifiziert ist. Das ist im Allgemeinen aber nicht gewährleistet, wenn man unter einer rekursiven Definition, wie oben erklärt, eine ‘Indizierungsfunktion’ verstehen möchte. Weil für eine Menge A und eine Wohlordnung R_A für irgendeine Ordinalzahl α gilt, dass $R_A \in \alpha$, können wir nicht immer eine *Funktion* (und in NFU brauchen wir eine *Funktion*, und keine echte *Operation*) finden, die jedem ordinalen Index dasjenige $a \in A$ zuordnet, sodass der durch α bestimmte Anfangsabschnitt der Reihe der Ordinalzahlen gleich lang ist, wie der durch a bestimmte Anfangsabschnitt von A .

Das zweite - und gewichtigere - Problem besteht darin, dass im Allgemeinen nicht sicher gestellt ist, dass es wirklich genug Ordinalzahlen gibt um *alle* Elemente einer beliebig langen Wohlordnung zu indizieren (vorausgesetzt, man weiss schon, was eine Indizierung *ist* (wenn sie keine Funktion mit Definitionsbereich ORD oder einem Anfangsabschnitt von ORD ist))⁷⁴ Definieren wir nämlich eine Folge von durch Ordinalzahlen indizierten Objekten, dann sind wir dabei durch den Ordnungstyp Ω aller Ordinalzahlen mit der üblichen Ordnung beschränkt. Es gibt aber offenbar längere Wohlordnungen als die natürliche Ordnung auf ORD ! Weil in NFU ORD eine Menge ist, erhalten wir eine längere Wohlordnung als $<_{ORD}$ z.B. indem wir zu ORD einfach irgendein $x \notin ORD$ dazunehmen und es für ‘größer’ als alle Ordinalzahlen erklären.

In ZFC dagegen gibt es nichts ‘Größeres’ (oder unter dem geordneten Aspekt ‘Längeres’) als die Gesamtheit der Ordinalzahlen - alle Klassen sind ‘gleich groß’. Man muss sich in NFU deshalb anderer Methoden bedienen, um ‘große’ oder besonders ‘lange’ Objekte oder Folgen von Objekten zu definieren.⁷⁵

All das zeigt, dass die Ordinalzahlen in NFU nicht dieselbe Rolle spielen, und auch nicht spielen *können*, wie in ZFC.

Eine ganz bestimmte durch Ordinalzahlen indizierte Folge - diejenige der Alephs - werden wir im nächsten Abschnitt kennen lernen.

7.5 ‘Die Struktur des Unendlichen’

In den jeweiligen Abschnitten über Kardinalzahlen in ZFC und NFU hatten wir allgemein von Kardinalzahlen $\kappa, \lambda, \kappa^+, \dots$ gesprochen und bemerkt, dass die Folge der unendlichen Kardinalzahlen ihrer Größe nach eine Wohlordnung darstellen (weil die Folge aller Kardinalzahlen ihrer Größe nach eine Wohlordnung ist). Dieser Folge können wir nun eine konkrete Gestalt in Form der Folge der

⁷⁴Man definiert eine Indizierungsfunktion indem man festlegt, dass für eine Menge A und eine Wohlordnung R auf A gilt: a_α ist dasjenige Element $a \in A$, sodass die Ordnung $R_{Pred(a, R, A)} \in \alpha$.

⁷⁵Vgl. dazu Forster [1992] S. 54 ff.

Alephs $\aleph_0, \aleph_1, \aleph_2; \dots$ geben. Das erste Glied dieser Reihe - $\aleph_0$ - wurde für beide Systeme bereits definiert durch $\aleph_0 := \omega$ (für ZFC) und $\aleph_0 := |\mathcal{N}|$ (für NFU).

Formal können wir in ZFC nun über transfinite Rekursion über die Ordinalzahlen definieren:⁷⁶

$$\aleph_{\alpha+1} := (\aleph_\alpha)^+$$

Um die Probleme, die im vorigen Abschnitt erwähnt wurden, zu vermeiden definiert man nun für NFU:

$$\aleph_\alpha := \iota \kappa (\prec_{Pred(\kappa, <^c, KARD')}^c \in \alpha)$$

wobei unter $KARD'$ die Menge aller unendlichen Kardinalzahlen zu verstehen ist, und unter $<^c$ die Wohlordnung auf dieser Menge. Für den Fall $\alpha = 1$ liefert diese Definition, dass $\aleph_1$ diejenige Kardinalzahl κ ist, für die der Ordnungstyp der Wohlordnung auf $Pred(\kappa, <^c, KARD')$ gleich 1 ist.⁷⁷ Das ist aber genau $(\aleph_0)^+$. Die Definition in NFU hat also denselben Effekt wie jene in ZFC; für jede Ordinalzahl α ist die zu $\alpha + 1$ gehörige Kardinalzahl λ der kardinale Nachfolger der zu α gehörigen Kardinalzahl κ , i.e. $\lambda = \kappa^+$.

Die Folge der Alephs enthält also in beiden Fällen *alle* Kardinalzahlen - es gibt keine Lücken.⁷⁸ In ZFC, wo Kardinalzahlen mit bestimmten Ordinalzahlen identifiziert werden, ergibt sich so ein hübsches Bild der Reihe der Ordinalzahlen, wo jedes Aleph einen neuen Abschnitt (eine sogenannte ‘Zahlklasse’) eröffnet:⁷⁹

$$0 <^o 1 <^o 2 <^o \dots \aleph_0 = \omega <^o \omega + 1 <^o \dots \aleph_1 = \alpha <^o \alpha + 1 <^o \dots \beta = \aleph_2 <^o \dots$$

In NFU können wir Kardinalzahlen nur mit bestimmten Ordinalzahlen *korrelieren*. Das macht man in augenscheinlicher Weise, indem man für $R_A \in \alpha$ setzt:

$$card(\alpha) := |A|$$

und

$$init(\kappa) := \min \{ \alpha : card(\alpha) = \kappa \}$$

Mit diesen Definitionen erhalten wir ein weitgehend analoges Bild für die Reihe der Ordinalzahlen in NFU

$$0 <^o 1 <^o 2 <^o \dots \omega (= init(\aleph_0)) <^o \dots \alpha (= init(\aleph_1)) <^o \dots \beta (= init(\aleph_2)) <^o \dots$$

⁷⁶Für Limesordinalzahlen

$$\aleph_\gamma := \bigcup \{ \aleph_\alpha : \alpha < \gamma \}$$

⁷⁷Weil $<_{\{\aleph_0\}}^c \in 1$

⁷⁸Und unter der Annahme der verallgemeinerten Kontinuumshypothese können sie sogar sehr einfach berechnet werden.

⁷⁹Mit $<^o$ ist natürlich die ordinale Kleiner-Relation gemeint.

Wir haben in den obigen Folgen sowohl für ZFC als auch für NFU Pünktchen verwendet, um anzudeuten, dass die Folge $\aleph_0, \aleph_1, \aleph_2, \dots$ kein Ende nimmt. Nun ist aber zu fragen: Wie lässt sich das rechtfertigen?

Wieder einmal ist der Fall für ZFC relativ klar: Alle Kardinalzahlen sind Alephs; weil die Vielheit der Kardinalzahlen eine echte Klasse ist (sie ist ja offensichtlich gleich groß wie diejenige der Ordinalzahlen, die Alephs wurden ja gerade über die Ordinalzahlen definiert), nimmt es mit den Kardinalzahlen tatsächlich kein Ende.

In NFU ist das anders: Klarerweise ist $|V|$ größer als alle anderen Kardinalzahlen, d.h. es gibt zumindest eine Kardinalzahl ohne Nachfolger;⁸⁰ Aber wie groß *ist* $|V|$? Für welchen Index gilt: $\aleph_\gamma = |V|$ (wenn es überhaupt der Fall ist, dass $|V|$ der Form $\aleph_\gamma$ ist)?

Die Antwort auf diese Frage hängt davon ab, durch welche zusätzlichen Axiome man NFU verstärkt. Ohne das *Axiom of Counting* ist es z.B. konsistent, dass $|V| = \aleph_n$ für eine natürliche Zahl n . Andererseits folgt mit dem *Axiom of Counting*, dass $\aleph_n$ für alle n existiert und ungleich $|V|$ ist.

Verstärkungen dieser Art verschieben die Grenze, die $|V|$ bzgl. der Kardinalzahlen bildet, immer weiter nach ‘außen’; für *irgendeine* Ordinalzahl α muss aber $|V| = \aleph_\alpha$ gelten und diese Ordinalzahl muss eine Nicht-Cantorsche Ordinalzahl sein. Unsere Folge der unendlichen Kardinalzahlen sieht also so aus:

$$\aleph_0 <^c \aleph_1 <^c \aleph_2 <^c \dots <^c \aleph_\alpha = |V|(\text{ENDE!})$$

Auch wenn wir die ‘Grenze’ $|V|$ durch stärkere Axiome immer weiter nach ‘außen’ verschieben können: eine Grenze bleibt! In dieser Hinsicht gleicht NFU jemandem, der sich jederzeit beliebig lange Urlaub nehmen kann, aber dann wieder arbeiten muss. Sie kann sich zwar einbilden, alle Möglichkeiten zu haben; alles tun zu können wonach ihr der Sinn steht - im Hinterkopf wird doch immer der Gedanke stehen, dass nach dieser Zeit wieder die Arbeit wartet. ZFC vergleicht man am besten mit einer jungen Heißsporn, die alle Verbindungen kappt, die Arbeit kündigt und völlig frei von irgendwelchen Bindungen lebt. Auch sie muss Mittel suchen (große Kardinalzahlaxiome), um sich ihre Traumreisen finanzieren zu können (große Kardinalzahlen), aber es gibt keine *prinzipielle* Schranke für sie.

Soll heißen: In intuitiven Modellen von ZFC finden wir *immer* beliebig große Kardinalzahlen, aber in jedem intuitiven Modell von NFU sind wir immer durch $|V|$ beschränkt.

In gewisser Weise ist es so, dass ZFC dadurch, dass es große Vielheiten wie V , ORD oder $KARD$ verbietet, *Weite schafft*, während NFU dadurch, dass es aus so großen Vielheiten Mengen macht, das Universum wie einen Käfig aussehen lässt.⁸¹

⁸⁰Dass $|V|$ die einzige ist, folgt aus dem Auswahlaxiom.

⁸¹Einen guten Eindruck von dieser ‘Enge’ erhält man, wenn man sich die großen Ordinal- und Kardinalzahlen ansieht. Die Potenzmengenoperation, die uns tendenziell größere Objekte liefern soll, angewandt auf V , *wirft uns wieder zurück* in der Reihe der Kardinalzahlen. Man prallt an $|V|$ wie an einer Mauer wieder ab.

Aber genug von den Zahlen und deren Arithmetik. Wir wollen uns nun, wie angekündigt, ansehen, wie es um die Intuitionen steht, die die beiden Systeme motivieren.

8 Schluss

Es wurde an mehreren Stellen darauf hingewiesen, dass den Systemen ZFC und NFU möglicherweise grundverschiedene Intuitionen darüber, was eine Menge ist, zu Grunde liegen. Wir werden in diesem Kapitel versuchen, dies näher zu begründen - zum Teil historisch, zum Teil inhaltlich.

Dass hinter ZFC und NFU verschiedene Mengenbegriffe stehen, liegt zunächst nicht unbedingt auf der Hand: Beides sind first-order Theorien und beide enthalten als primitive zweistellige Relation ‘ $\in$ ’ und in beiden Systemen liest man ‘ $x \in y$ ’ in einem bestimmten Sinne ‘zu Recht’ als ‘ x ist ein Element von y ’. In beiden Systemen gilt, dass Mengen allein durch ihre Elemente bestimmt sind (extensionaler Aspekt der Mengen) und beide gestatten die Bildung von Mengen aus vorher gegebenen ‘Dingen’ (ob nun Urelementen oder wiederum Mengen), z.B. des Singletons $\{x\}$ für beliebiges x oder der Paarmenge $\{x, y\}$, der Potenzmenge $\mathcal{P}(x)$ usw. Auch die Tatsache, dass sich die Sprache, in der NFU formuliert ist, von der, in der ZFC formuliert ist, unterscheidet, sollte keinen Anlass zu grundlegenden Erwägungen geben: Die primitive Paarfunktion, die wir in NFU verwendet haben, kann man sich mit einigem Aufwand sparen und die Definition von Kuratowski verwenden und Urelemente kann man durch Abschwächung des Extensionalitätsaxioms und durch Einführung zusätzlicher nichtlogischer Konstanten auch in die Theorie von ZFC einkalkulieren ohne den dadurch bestimmten Mengenbegriff zu verändern. Sowohl in ZFC als auch in NFU kann man sich, wie wir gesehen haben, die ‘üblichen’ Cantor’schen Fragen nach Ordinalzahlen, Wohlordnungen, transfiniter Arithmetik usw. stellen.

Was also macht ZFC so verschieden von NFU?

Um zu sehen, wieso sie sich tatsächlich wesentlich voneinander unterscheiden, werfen wir einen Blick zurück auf das System NF, aus dem NFU entstanden ist. Dass das gerechtfertigt für die Untersuchung des intuitiven Mengenbegriffs von NFU ist, liegt daran, dass durch Hinzufügen von Urelementen nichts an der Vorstellung der Mengenbildung geändert wird. Das einzige *tool* zur Generierung von Mengen ist wie auch in NFU das S-Komprehensionsschema. Der einzige Grund für die Wahl von NFU als Objekt der Gegenüberstellung zu ZFC bestand ja darin, dass NF insofern ‘unzureichend’ ist, als es einerseits das Auswahlaxiom widerlegt und es andererseits auch nach 80 Jahren noch immer keine befriedigenden (relativen) Konsistenzbeweise gibt.⁸²

8.1 Von Frege über Russell zu TT

Damit wir den Mengenbegriff von Quines NF besser verstehen können, werfen wir zunächst einen Blick zurück zu Frege. Um sich dem Problem der Mengen bei Frege nähern zu können, muss man zuerst verstehen, was ‘Logik’ bei

⁸²Relative Konsistenzbeweise für NFU gibt es zuhauf. Vgl. dazu Forster [1992] S. 66 ff. Der erste Konsistenzbeweis für NFU (ohne Auswahl- und Unendlichkeitsaxiom) stammt von Jensen und ist ein Theorem der Peano-Arithmetik.

Frege heißt. Das Programm Freges war ja, die Arithmetik (also die Theorie der natürlichen Zahlen) auf die Logik zu gründen. Eine Beurteilung der ‘Güte’ dieser Rückführung muss (lässt man die Inkonsistenz des Frege’schen Systems beiseite) Kriterien angeben können, inwiefern diese ‘Logik’ tatsächlich ‘logisch’ (oder eben *nicht* ‘logisch’) ist. Ist man etwa prinzipiell der Meinung, die Rückführung sei gelungen, muss man erklären können, inwiefern die verwendeten Grundbegriffe Freges (wie *Begriff*, *Gegenstand*, *Begriffsumfang*, *Wertverlauf* usw.) ‘logischer Natur’ sind.

Ist man auf der anderen Seite der Meinung, die Rückführung sei nicht gelungen, muss man Kriterien angeben können, an welcher Stelle in der (vermeintlichen) Rückführung nicht-(oder *nicht ausschließlich*) logische Prinzipien zum Tragen kommen.

Es zeigt sich, dass die zweite Sichtweise zunächst klare Vorteile in der Interpretation hat - eine Sichtweise, die besonders von Quine forciert wurde, und in der Folge sehr wirkmächtig war. In diesem Bild meint man mit ‘Logik’ die Prädikatenlogik erster Stufe und eine *mathematische* Theorie beginnt dort, wo nichtlogische Konstanten und nichtlogische Axiome ins Spiel kommen.

Die Tatsache, dass in Freges Kalkül auch über *Begriffe* quantifiziert wird (tendenziell second-order), reicht in dieser Betrachtungsweise schon aus, um die Rückführung als ‘nicht gelungen’ zu brandmarken. Aber auch an anderer Stelle wird man die Rückführung auf ‘rein logische Prinzipien’ als verfehlt betrachten müssen. Vereinfacht ausgedrückt: Dort, wo Mengen und Axiome, die die Interaktion von Mengen regeln, ins Spiel kommen (was bei Frege in Form seiner *Begriffsumfänge* und des notorischen Grundgesetzes V, passiert), hört *reine Logik* (i.e. *Prädikatenlogik erster Stufe*) auf und Mathematik beginnt.⁸³

Der Grund, wieso Frege glauben konnte, dass sein Rückführungsprojekt gelungen sei, ist, dass er meinte, die Rede von Mengen oder ‘Begriffsumfängen’ sei eine *abgeleitete*. Die Elementrelation ist bei Frege *definiert*, und zwar durch $x \in y :\Leftrightarrow \exists F(y = \hat{e}F\epsilon \wedge Fx)$, wobei das Zeichen $\hat{e}F\epsilon$ gelesen wird als ‘die Menge aller epsilons mit der Eigenschaft F ’. Es ist demnach ‘ x Element von y ’ genau dann wenn es einen Begriff F (ein *Prädikat* in moderner Terminologie)⁸⁴ gibt, derart, dass y der Umfang dieses Begriffes ist und F auf x zutrifft.

⁸³Um es klar zu sagen: Das Frege’sche System aus den *Grundgesetzen* ist *keine Mengenlehre* (auch wenn manche so tun, als wäre es eine *Quasi-Mengenlehre*). Eine Mengenlehre im eigentlichen Sinne ist ein System, in dem man über Mengen quantifiziert und das die Elementrelation als primitive Relationskonstante enthält. Das disqualifiziert sowohl Freges System, als auch das von Russell als Mengenlehre - auch wenn Mengen in der ein oder anderen Form auftreten. Freges Theorie geht eher in Richtung *second-(oder higher-)order Arithmetik*. (Vgl. dazu auch neuere Arbeiten von Crispin Wright z.B. [2007].) Die Werte der ‘Begriffe’, über die Frege - in einer ‘klassischen’ second-order Interpretation seines Systems - quantifiziert, sind dann *Mengen*. (Freilich bedeutet eine ‘klassische’ Second-Order-Interpretation, dass das Frege’sche System insgesamt ‘geglättet’ wird.) Insofern hat Quine nicht ganz unrecht, wenn er die Second-order logic “Mengenlehre im Schafspelz” nennt. Vgl. Quine [2005] S. 93 ff.

⁸⁴Eigentlich quantifiziert Frege über Begriffe - der Grund, wieso wir hier *Prädikat in moderner Terminologie* in Klammer notiert haben, ist, dass in Freges System jedes Prädikat (i.e. jede Formel mit einer freien Variable), einem Begriff korrespondiert: d.h. falls $\varphi(x)$ eine Formel mit x frei ist, dann gilt: $\exists F\forall x(F(x) \leftrightarrow \varphi(x))$.

Die Elementrelation wird also *de facto* über die *logische Prädikation* definiert. Mengen sind keine eigenständigen Objekte, sondern nur insofern ‘existent’, als sie Umfänge *von Begriffen* sind; und die Elementrelation ist keine ‘irreduzible’ Grundbeziehung, sondern für jede beliebige Eigenschaft ‘separat’ definiert einfach als das Zutreffen dieser Eigenschaft auf einen Gegenstand.

Der entscheidende Schritt beim Übergang von Begriffen zu deren Umfängen passiert bei Frege nun in Form des notorischen Grundgesetzes V, das besagt, dass sich jede Gleichheit von Begriffen in eine Gleichheit von deren Umfängen umwandeln lässt und umgekehrt. Aus Grundgesetz V ergibt sich leicht, dass jeder Begriff einen Begriffsumfang hat, in Frege’scher Terminologie $\forall F \exists y (y = \hat{\epsilon} F \epsilon)$ oder leicht umformuliert $\forall F \exists y \forall x (x \in y \leftrightarrow Fx)$, das naive Komprehensionsprinzip, aus dem man, wie wir schon gesehen haben, sehr einfach das Russell-Paradox ableiten kann.

Mit einigem Recht kann man also Frege den sogenannten ‘*naiven Begriff der Menge*’ unterlegen, welchen wir durch das Extensionalitätsaxiom und das naive Komprehensionsschema kodifiziert betrachten.

Es wurde viel über das Scheitern Freges nachgedacht, und die Gründe sind vielfältig, ist doch das Frege’sche System ein Sammelsurium aus Higher-Order Logik, Mengenlehre und eigenwilliger Ontologie und Semantik. Auf einen Punkt sollte dennoch – auch im Hinblick auf die Weiterentwicklung Fregeanischer Ideen durch Russell – explizit hingewiesen werden: Noch in ‘*Funktion und Begriff*’ bemerkt Frege, dass man den Quantifikationsbereich von Begriffen erster Stufe (das sind Eigenschaften, die von Fregeschen Gegenständen präzisieren⁸⁵) streng von solchen zweiter Stufe (das sind Begriffe, die von Begriffen erster Stufe präzisieren) trennen muss, und den Quantifikationsbereich von Begriffen dritter Stufe von dem zweiter und erster Stufe usw. Ein Gedanke, der schon in Richtung Typentheorie verweist.⁸⁶

De facto wird diese angepeilte *Hierarchisierung* aber durch GG V verhindert. GG V erlaubt uns gewissermaßen, Begriffe beliebiger Stufen so zu behandeln, als wären sie ‘Gegenstände’.

Wir können etwa vom Begriff ‘*x ist ein Mensch*’ (oder ‘*Menschsein*’) übergehen zur *Menge* aller Menschen, einem ‘Gegenstand’. Während die (unformalisierte) Funktion-und-Begriff-Sichtweise Fragen verbietet wie ‘Ist der Begriff des Menschseins eine Mensch?’, wird das in der (formalisierten) Grundgesetze-der-Arithmetik-Sicht (zumindest *zum Teil*) wieder zurückgenommen, weil man sich nun die Frage stellen kann ob die *Menge* aller Menschen ein Mensch ist. Weil Begriffe wie $M(x)$ der Forderung nach ‘scharfer Begrenztheit’ (Begriffe müssen für *alle* Gegenstände einen Wahrheitswert haben) unterliegen, muss $M(\hat{\epsilon} M \epsilon)$ einen Wahrheitswert haben. Die Frage ‘Enthält sich die Menge aller Mengen, die sich nicht selbst enthalten, selbst?’ ist also völlig in Ordnung.

Der Punkt ist: Während die *Begriffe* tatsächlich hierarchisch geordnet sind,

⁸⁵Der Hinweis, dass es sich hier um *Fregesche* Gegenstände handelt, ist entscheidend.

⁸⁶Vgl. dazu Wang [1974], wiederabgedruckt in Benacerraf/Putnam [1983] S. 560

werden deren *Umfänge* samt und sonders auf ‘Typ 0’ heruntergebrochen. In diesem Sinne ist GG V insofern schuld an der Inkonsistenz von Freges System (zusammen mit seiner ontologischen Unterscheidung zwischen *Begriff* und *Gegenstand*), als es eine ‘effektive’ Typisierung verhindert. Den Widerspruch erhält man dann, wenn man den Begriff bildet ‘Der Umfang eines Begriffes zu sein, auf den dieser Begriff nicht zutrifft’, d.h. den Begriff ‘ $H(x) := \exists F(x = \hat{\epsilon}F\epsilon \wedge \neg Fx)$ ’. Für diesen Begriff erhält man nun sowohl für $H(\hat{\epsilon}H\epsilon)$ als auch für $\neg H(\hat{\epsilon}H\epsilon)$ einen Widerspruch.

Für Russell ergibt sich der Widerspruch von oben (unter anderem) deshalb, weil die Variable x in der obigen Formel $H(x)$ auch *die Menge aller x für die $H(x)$ gilt*, als Wert annehmen kann. Für Russell stellt sich deshalb das Problem zunächst auch in der Form: ‘Über welche Gesamtheit kann man widerspruchsfrei quantifizieren?’⁸⁷

Erst Russell gelingt es, die Begriffe (oder *propositional functions* wie er sie nennt) in eine ‘effektive’ Typenhierarchie zu bringen, weil er die Variablen, die an den Argumentstellen einer propositional function stehen, nicht mehr über das gesamte Universum (Die Gesamtheit der Frege’schen *Gegenstände*) laufen lässt. $F(x)$ ist also nicht mehr für *alle* Werte sinnvoll oder ‘signifikant’, sondern nur für solche eines *bestimmten Typs*. Z.B. werden durch die Typentheorie Russells Ausdrücke wie $F(F)$ ausgeschlossen und weil Gleichheit nur zwischen Dingen desselben Typs bestehen kann, ist auch der Ausdruck $\exists F(x = \hat{\epsilon}F\epsilon \wedge \neg Fx)$ schlicht ungrammatisch. Weil Mengen bei Russell *propositional functions in einem bestimmten Gebrauch* sind, legt man sich mit $x = \hat{\epsilon}F\epsilon$ typenmäßig darauf fest, dass die Werte von x denselben Typ haben wie F , mit $\neg F(x)$ aber darauf, dass der Typ von x kleiner ist als der von F .⁸⁸

Wir wollen kurz skizzieren, wie das bei Russell aussieht:⁸⁹

Auf der untersten Stufe oder dem unterstem ‘Typ’ stehen Individuen. Individuen sind nicht näher bestimmt, aber man sollte sich darunter Dinge vorstellen wie Steine, Punkte, oder wie auch immer nicht näher analysierbare Einheiten.⁹⁰ Ausgehend von diesem Bereich der Individuen bilden wir nun propositional functions dieser Individuen. Wenn z.B. x eine Variable ist, die Individuen als Werte annimmt, dann ist ‘ x ist rot’ eine propositional function. Über den Bereich der propositional functions erster Stufe (oder ersten ‘Typs’) können wir nun wieder quantifizieren. Steht z.B. ‘ $F(x)$ ’ für eine propositional function ers-

⁸⁷Vgl. dazu auch Ramharter/Rieckh [2006] S. 26 ff.

⁸⁸“If α represents a class and $\varphi\hat{x}$ is one of its determining functions [so that $\alpha = \hat{z}(\varphi z)$], it is not sufficient that $\varphi(\alpha)$ be a false proposition, it must be nonsense.” Russell/Whitehead [1910] S.25 Eigentlich ist Russells Vorschlag zu Vermeidung der Paradoxien eine Vermischung von zwei Ideen: 1. Der Bezug auf die *Wohlgeformtheitsbedingungen* und 2. Die *Beschränkung des Wertebereichs* einer Variable. (Ein Umstand, der aus seinem *vicious circle principle* folgt.)

⁸⁹Wir beschränken uns hier auf den für unsere Zwecke einzig relevanten Fall einer einfachen (nicht-verzweigten) Typentheorie. Eine erste Version einer einfachen Typentheorie findet sich in Russell [1903]

⁹⁰Russell vermeidet eine tatsächliche Festlegung auf den jeweiligen Typ durch seine Konvention der *systematic ambiguity* - d.h. wichtig sind die Typenzuschreibungen nur in einem bestimmten Kontext (eines Beweises oder einer Definition etwa) und nicht in einem ‘absoluten’ Sinn. Was man z.B. unter einem ‘Individuum’ zu verstehen hat, ist demnach kontextabhängig.

ter Stufe (d.h. eine propositional function von Individuen), so können wir fragen ob diese propositional function eine bestimmte andere propositional function erfüllt. Z.B können wir uns fragen, ob auf den Begriff ‘Rot sein’ (auf die propositional function ‘ x ist rot’) der Begriff zweiter Stufe ‘Farbbegriff sein’ (‘ F ist ein Farbbegriff’) zutrifft. Usw.

Ein Nachteil dieser Art Typentheorie ist, dass wir dort über Objekte quantifizieren, die nicht besonders ‘klar’ sind. Propositional functions sind eine Art ‘Zwitterwesen’ zwischen *Formeln* mit freien Variablen und ‘Objekten’ im Sinne von Freges *Begriffen*.⁹¹

Zu einer einfachen Typentheorie⁹², die als Grundobjekte Mengen und Individuen hat und als Grundrelation die Elementbeziehung, gelangen wir, wenn wir die Prädikation als Elementrelation deuten; d.h. wir deuten $F(x)$ (wie Russell es durch seine Kontextdefinition der Elementrelation nahelegt) als $x \in F$ (sprich: ‘ x ist Element der durch F bestimmten Menge’)⁹³. Die Hierarchie der typentheoretisch geordneten (*prädikativen*) propositional functions⁹⁴ und die Wohlgeformtheitsbedingungen ‘parallelisieren’ wir, indem wir von einer Variable, die rechts von einem $\in$ in einer Formel $x \in y$ steht, verlangen, dass sie genau einen Typ höher ist, als jene Variable links vom $\in$ und wir deuten diese Typenhierarchie durch natürliche Zahlen als Indizes an.

Die Axiome dieser Theorie (wir nennen sie TT) sind:

$$\forall x_{n+1} \forall y_{n+1} ((\forall z_n (z_n \in x_{n+1} \leftrightarrow z_n \in y_{n+1}) \rightarrow x_{n+1} = y_{n+1})$$

$$\exists y_{n+1} \forall x_n (x_n \in y_{n+1} \leftrightarrow \varphi(x_n))$$

TT⁹⁵ ist in vielerlei Hinsicht einfacher als die Typentheorie Russells. Zum einen ist TT in einer mehrsortigen Logik formuliert und es bedarf keiner *higher-order-logic*, deren Semantik um einiges komplizierter ist (Bei Russell selbst sind *Syntax* und *Semantik* nicht klar getrennt). Wir können durch Hinzufügen von Prädikaten $T_1, T_2, T_3, \dots$ (‘Typ 1’, ‘Typ 2’, ‘Typ 3’, ...) aus TT sogar eine normale First-Order-Theorie machen, indem wir statt der Indizes eine bestimmte

⁹¹An Beispielen, wie dem vom Begriff ‘Rot-Sein’ und ‘Farbbegriff sein’ wird einem sehr schnell die Mehrdeutigkeit des Begriffs ‘propositional function’ bewusst und warum Leute wie Quine ihre Probleme damit haben. Die Frage ist, was genau man meint, wenn man über propositional functions quantifiziert? Meint man damit eine Formel mit freien Variablen oder ist damit eine eigene Art ‘Ding’ gemeint. Das spiegelt sich bei Russell auch typographisch. Wenn Russell den Formel-Charakter betonen will, schreibt er etwa $\varphi(x)$ oder φx ; für abquantifizierte propositional functions verwendet er andererseits Schreibweisen wie $\varphi \hat{x}$, um zu betonen, dass wir es nun mit einem ‘Objekt’ zu tun haben. Siehe Russell/Whitehead [1910] S.16 ff.]

⁹²‘Einfach’ im doppelten Sinne von ‘nicht verzweigt’ und im Sinne von ‘einsichtiger’

⁹³Genau in diesem Sinne wurde auch der $\in$ -Formalismus von Peano eingeführt. $x \in \alpha$ liest sich bei Peano ‘ x ist ein α ’. Vgl. Peano [1889]

⁹⁴*Prädikativ* nennt Russell eine propositional function, wenn der Typ des Arguments genau um einen Typ niedriger ist als der Typ der propositional function selbst.

⁹⁵Es ist zu beachten dass beide Axiome eigentlich Axiomenschemata sind. Durch Einsetzung einer *bestimmten* natürlichen Zahl für den schematischen Buchstaben n erhalten wir dann im ersten Fall ein Axiom (Extensionalität), im zweiten Fall wiederum ein Axiomenschema (Komprehension), das dann erst durch Einsetzung einer *bestimmten* Formel für das schematische $\varphi(x)$ zu einem Axiom wird. Als Urheber von TT gelten (nach Quine) Gödel und Tarski.

Variable auf das entsprechende Typenprädikat beschränken.⁹⁶ TT zeichnet sich im Vergleich zu Russells Theorie durch einen klaren Aufbau aus - durch die Einführung von $\in$ als nichtlogischer Konstante und dem Umbau in eine ‘normale’ first order Theorie können wir die üblichen semantischen Fragen nach Modellen usw. stellen. Und statt der Zwitterwesen *propositional functions*, quantifizieren wir über Individuen, Mengen von Individuen, Mengen von Mengen von Individuen, usw.

Dennoch bleiben einige Probleme, die Russell hatte, auch für TT bestehen:⁹⁷ Das Komprehensionsschema für TT erlaubt uns die Bildung der ‘Allmenge’ $V_n := \{x_n : x_n = x_n\}$; V_n ist aber nicht die Menge *aller* Dinge, sondern bloß die Menge aller Dinge vom *Typ* n . Wir haben also für *jeden Typ* eine eigene Allmenge. Ähnlich ergeht es uns mit allen in diesem System definierbaren Objekten. Es gibt eine Leermenge für jeden Typ, die natürlichen Zahlen, die wir in der schon beschriebenen Art nach Frege-Russell entwickeln (und wofür wir natürlich ein geeignetes für TT adaptiertes Unendlichkeitsaxiom benötigen), treten ebenfalls für jeden Typ ≥ 2 auf, genauso die reellen Zahlen, π , e und die Quadratwurzel aus 2.

Das alles ist nicht sehr wünschenswert. Man möchte *ein* Universum, *ein* π und *eine* Quadratwurzel aus 2.

8.2 Quine’s NF und der modifizierte logische Begriff der Menge

Die Idee Quines, die dieses Problem löst wurde in ‘*New foundations for mathematical logic*’ entwickelt und sieht so aus:

Wir nehmen als Axiome für unsere neue Theorie NF genau die Axiome von TT, nur ohne Typenindizes (d.h. genau die Axiome der naiven Mengenlehre). Die Variablen sind universell und nicht in Typen eingeteilt. Da wir uns vor den Antinomien schützen wollen, machen wir uns aber die Idee, die hinter den Indizes von TT steht derart zunutze, dass wir sagen: eine Formel $\varphi(x)$ *bestimmt* eine Menge nur dann, wenn es möglich ist, Typenindizes derart anzubringen, dass diese Formel wohlgeformt wäre im Sinne von TT.

Wir treffen die Schutzvorkehrungen gegen die Paradoxien (allen voran der Russellschen) also nicht wie Russells Typentheorie oder TT auf der *Ebene der zulässigen Sätze* (den Wohlgeformtheitsbedingungen), sondern erst *beim Übergang von Formeln (oder Prädikaten) zu Mengen*. Für TT ist $x_n \in x_n$ einfach kein sinnvoller Ausdruck: $x \in x$ ist in NF wunderbar - kommt aber als Instanz für die NF-Form des Komprehensionsschemas nicht in Frage, weil $x_n \in x_n$ nicht wohlgeformt nach TT ist.

Weil unsere Variablen universell sind, gibt es nun *eine* Allmenge V , *eine* leere Menge und *eine* 2.

⁹⁶Das Extensionalitätsschema von TT bekommt so die Form: $\forall x \forall y (T_{n+1}(x) \wedge T_{n+1}(y) \rightarrow (\forall z (T_n(z) \rightarrow ((z \in x \leftrightarrow z \in y) \rightarrow x = y))))$ Vgl. dazu auch Quine [1973]

⁹⁷Russell selbst versuchte diesen Problemen - wenigstens zum Teil - durch sein *Reduzibilitätsaxiom* Herr zu werden. Vgl. dazu Ramharter/Rieck [2006] S. 85 ff.

Wir wollen noch einmal zusammenfassen, aus welchen Motiven sich der Mengenbegriff von NF (u. damit der von NFU) speist.

Da NF *de facto* aus den Axiomen der naiven Mengenlehre (mit der Einschränkung, dass nur stratifizierte Formeln Mengen bestimmen) besteht, scheint ein Hauptmotiv darin zu bestehen, den Mengenbegriff der naiven Mengenlehre in einem bestimmten Sinne *so unverfälscht wie möglich* zu erhalten.

Der Mengenbegriff der naiven Mengenlehre besteht aber gerade darin, die Elementrelation im Sinne der Prädikation zu deuten, oder wie Gödel es formuliert: Wir erhalten diesen Begriff “by dividing the totality of all existing things into two categories”, nämlich jene, auf die ein Prädikat zutrifft und jene, auf die es nicht zutrifft.⁹⁸

Es ist zu beachten, dass dieser ‘naive’ Mengenbegriff zunächst absolut nichts ‘Anrühiges’ an sich hat und in der Mathematik eigentlich ständig gebraucht wird; wenn wir z.B. etwas beweisen wollen über die Menge aller reellen Zahlen aus dem Intervall $[2, 3]$, so beweisen wir etwas für eine Gesamtheit $\{x : \varphi(x)\}$, wobei $\varphi(x) := 2 \leq x \leq 3$. In der elementaren Zahlentheorie beweisen wir Sätze für die Menge aller Primzahlen P , die Menge aller ganzen Zahlen x , für die $x \equiv_{\text{mod } 3} 0$ oder die Menge aller perfekten Zahlen; wir betrachten in der Analysis die Menge $C^n(X, E)$ aller n -mal stetig differenzierbaren Funktionen von X nach E usw.

All diese Mengen bilden wir mithilfe des naiven Komprehensionsgesetzes “by dividing the totality of all existing things into two categories” - und es ergeben sich keine Widersprüche. Dabei lesen wir $x \in A$ immer als ‘ x ist ein A ’; $f \in C^n(X, E)$ bedeutet, dass f eine n -mal stetig differenzierbare Funktion von X nach E ist, $x \in P$ bedeutet, dass x eine Primzahl ist usw.

Aus der Sicht des ‘working mathematician’ (d.h. des Mathematikers, der sich nicht um die ‘mengentheoretischen Grundlagen’ seiner Disziplin kümmert; der Funktionalanalysis, Lineare Algebra oder Geometrie betreibt), der es immer mit klar abgesteckten Quantifikationsbereichen zu tun hat (Der Analytiker mit seinen reellen Zahlen und Funktionen; der Zahlentheoretiker mit seinen natürlichen Zahlen usw.) birgt dieser ‘logische’ Begriff der Menge keine Risiken;⁹⁹ ‘Riskant’ wird es erst, wenn wir Mengen von *allen möglichen mathematischen Objekten* bilden wollen und diese durch Eigenschaften festgelegten Gesamtheiten wiederum als Werte der Variablen zulassen, über die wir quantifizieren - kurz: wenn wir beginnen, *Mengenlehre* zu betreiben (‘Mengenlehre’ in dem schwachen Sinn, dass Mengen (unter anderem) Werte der Variablen sind). Dennoch zeigen die mannigfachen Beispiele, dass es durchaus von Wert ist, den logischen Begriff der Menge zu *erhalten*.

Dieser ‘logische’ Begriff der Menge liegt nun in zweifacher Hinsicht auch NF(U) zugrunde: Erstens vermöge der typentheoretischen Intuitionen, die dahinter stehen, die besagen, dass man Prädikate nicht von sich selbst präzisieren

⁹⁸Vgl. Gödel [1947] in Benacerraf/Putnam [1983] S. 475

⁹⁹Rosser dazu: “Belief in it [das naive Komprehensionsprinzip, Anm.] is so strong that proofs of its falsity are called paradoxes and are widely ignored.” in Rosser [1953] S. 199

kann; dass Prädikate ‘logisch höhere’ Objekte sind als die Dinge, über die sie aussagen (und nicht, dass eine *Menge* ein höheres Objekt ist als deren *Elemente* - was daraus zwar folgt, wenn man die Elementrelation wie etwa Russell kontextmäßig definiert, aber nicht *primär* ist). Zweitens aber (und *vor allem*) via des Komprehensionsprinzips, das für NF adaptiert wurde und das *de facto* besagt, dass die Elementrelation der Prädikation ‘parallel’ ist - genau das besagt ja die Äquivalenz $x \in y \leftrightarrow F(x)$. Wie Parsons richtig bemerkt ist “die Unterscheidung zwischen F und y [für eine Menge $y := \{x : F(x)\}$ Anm.] einfach die Unterscheidung zwischen $() \in y$ und y ”¹⁰⁰.

Mit gutem Recht könnte man sagen, der durch NF formalisierte Mengenbegriff ist eine *Korrektur* des naiven logischen Begriffs. Es ist also zu differenzieren zwischen dem *naiven* logischen (und inkonsistenten), und dem *modifizierten* logischen, aber dennoch *logischen* (und mutmaßlich konsistenten) Mengenbegriff von NF.¹⁰¹

Vor allem ist der verbreiteten Kritik an NF (zumindest teilweise) zu widersprechen, wonach es *gar keinen* intuitiven Begriff gibt, den NF formalisiert, und derzufolge NF “essentially negative in character”¹⁰² ist. Gemeint ist damit, dass NF keine der Theorien ist, die es gegeben hätte, auch wenn es die Paradoxien niemals gegeben hätte - dass NF aus den Paradoxien und deren Analyse ‘geboren’ wurde und intuitiv eigentlich unverständlich wäre, wenn man noch nie etwas vom Russell-Paradox und seinen Verwandten gehört hätte.

Diese Kritik ist bis zu einem bestimmten Grad berechtigt, übergeht aber die Tatsache, dass der *naive Mengenbegriff* in der Tat hochgradig ‘einsichtig’ ist und nichts an Anschaulichkeit vermissen lässt und wert, so ‘rein’ als möglich erhalten zu werden (auch wenn einem die Stratifizierungsbedingung nicht dazu geeignet und künstlich erscheinen mag).

Sehen wir uns zum Vergleich noch einmal den Mengenbegriff von ZFC an:

8.3 ZFC und der iterative Begriff der Menge

Eine Einführung in die ZFC-Mengenlehre beginnt üblicherweise mit der Präsentation der naiven Mengenlehre und der Antinomie von Russell, die daraus folgt. In weiterer Folge wird das Aussonderungsaxiom dann als *Beschränkung* des naiven Komprehensionsprinzips vorgestellt und die übrigen Axiome als notwendige Instanzen, die einem Mengen liefern sollen, auf die wir das Aussonderungsprinzip anwenden können.

Diese Geschichte suggeriert, dass ZFC (wie NF) bloß eine Korrektur des *naiven logischen Mengenbegriffs* darstellt. Andererseits wurde schon in Kapitel

¹⁰⁰Parsons [1975], wiederabgedruckt in Benacerraf/Putnam [1983] S. 517

¹⁰¹Es sollte darauf hingewiesen werden, dass Russells ‘Funktionen in einem gewissen Gebrauch’-Begriff der Menge ebenfalls eine Korrektur des naiven logischen Mengenbegriffs darstellt, aber eine Korrektur, die Mengen als Grundobjekte überhaupt verwirft und deshalb für uns nicht von Interesse ist. Wir sind hier am logischen Begriff der Menge nur insofern interessiert, als er Motivation für eine *Mengenlehre* (in einem starken Sinn) ist.

¹⁰²So Michael Potter in Potter [1993]

4 darauf hingewiesen, dass der Mengenbegriff, den ZFC formalisiert, der *iterative Begriff* der Menge ist. Wenn also logischer und iterativer Begriff unvereinbar sind - und das sind sie - stellt sich die Frage, welches die richtige Sichtweise ist.

Eine vorläufige Antwort lautet: Der *historische* Ausgangspunkt von Zermelo war in der Tat der naive Begriff der Menge - auch Zermelo selbst scheint sein System als Korrektiv für den naiven Begriff verstanden zu haben.¹⁰³

Inhaltlich oder *begrifflich* aber ist ZFC - auch wenn Zermelo selbst das nicht zur Gänze klar war¹⁰⁴ - eine Formalisierung eines *anderen* - nämlich des *iterativen* - Begriffs der Menge.

Um einzusehen, wieso das so ist, wollen wir noch einmal skizzieren, was den iterativen Begriff ausmacht und warum ZFC diesen Begriff formalisiert.

Wir gehen in unserer Skizze davon aus, dass uns zu Beginn irgendein Bündel von Individuen gegeben ist. Wir nennen dieses Bündel die *0. Stufe*. Individuen stellen wir uns dabei als nicht weiter zerlegbare Einheiten vor. Diese Individuen können wir nun beliebig zusammenfassen, oder wie Kripke es formuliert: “Wir binden ein Lasso um sie.”¹⁰⁵ Das einzige, das wir verlangen ist, dass wir dieselbe Menge erhalten, wenn das Lasso um dieselben Individuen gebunden ist. Wir machen das nun in allen möglichen Konstellationen und nennen die Gesamtheit all dieser Zusammenfassungen die *1. Stufe*. Sind uns etwa $\diamond$ und $\emptyset$ als Individuen gegeben, so besteht die erste Stufe aus den Mengen $\{\diamond\}$, $\{\emptyset\}$, $\{\diamond, \emptyset\}$ und der leeren Menge $\emptyset$. (Falls uns keine Individuen gegeben sind, besteht die erste Stufe nur aus der leeren Menge.)

Auf der *2. Stufe* bilden wir alle Zusammenfassungen von Individuen und Zusammenfassungen der *1. Stufe*. Diese Stufe enthält z.B. als Elemente die Mengen $\{\{\diamond\}\}$, $\{\emptyset, \emptyset\}$ oder $\{\{\diamond\}, \emptyset, \emptyset\}$.

Wir wiederholen diesen Prozess immer und immer wieder. Nachdem wir alle Stufen 0,1,2,3... durchgegangen sind, und alle Zusammenfassungen von Zusammenfassungen von... der *0., 1., 2., ... Stufe* gebildet haben, bilden wir die Gesamtheit aller Zusammenfassungen von Zusammenfassungen von... der *0., 1., 2., ... Stufe*. Wir nennen diese Gesamtheit die *ω . Stufe*.

Danach bilden wir alle möglichen Zusammenfassungen von Zusammenfassungen von ... der *0., 1., 2., ...* und der *ω . Stufe* und nennen diese Stufe die *$\omega + 1$. Stufe*. Danach bilden wir alle möglichen Zusammenfassungen von ...

Durch Beobachtung dieses iterativen Prozesses können wir nun einige Fakten zum iterativen Begriff ableiten.¹⁰⁶ Z.B. sehen wir, dass jede Menge, die

¹⁰³Wang z.B. betont, dass Zermelo als ‘bankruptcist’ *spricht* (d.h. so tut als würde durch die Paradoxien gezeigt worden sein, dass unsere logischen Intuitionen irreführend sind), tatsächlich aber wie ein Vertreter der “misunderstanding”-Theorie *handelt* (d.h. seinen “guten Intuitionen” zum (iterativen) Mengenbegriff folgend, seine Axiome aufstellt). Vgl. Wang [1974], wiederabgedruckt in Benacerraf/Putnam [1983] S. 543

¹⁰⁴Und auch nicht vollständig klar sein *kann*, weil der Begriff des *iterativen Begriffs* eine Erfindung der zweiten Hälfte des 20. Jahrhunderts ist.

¹⁰⁵Boolos [1971] in Benacerraf/Putnam [1983] S.491

¹⁰⁶Folgendes stammt im Wesentlichen von George Boolos in Boolos [1971]. Eine überarbeitete Variante davon findet man in Boolos [1989].

überhaupt gebildet werden kann, auf irgendeiner bestimmten Stufe *zuerst* gebildet worden sein muss¹⁰⁷ und dass in der Folge der Stufen keine Stufe vor sich selbst kommt. Daraus folgt z.B., dass es keine Menge geben kann, die sich selbst als Element enthält, denn dazu müsste diese Menge auf einer bestimmten Stufe zuerst gebildet worden sein und diese Stufe müsste vor sich selbst kommen. Es kann deshalb auch keine Menge geben, die *alle* Mengen enthält, weil diese Menge sich selbst enthalten müsste. (Aus dieser Perspektive zeigt sich auch, dass ZFC insofern keine *limitation of size* Theorie ist, als sie große Gesamtheiten *per se* ausschließt - ZFC schließt große Gesamtheiten (zumindest nachweislich V) *deshalb* aus, weil sie unvereinbar mit dem *iterativen Begriff* sind.)

Weitere Beobachtungen sind:

1. Es gibt eine erste Stufe
 2. Jede Stufe hat eine unmittelbare Nachfolgerstufe
 3. Jedes Element einer Menge wird vor dieser Menge (zum ersten mal) gebildet
 4. Jede Menge wird auf einer eindeutig bestimmten Stufe (zum ersten mal) gebildet
- usw.

Aus Beobachtungen dieser Art können wir uns ein Axiomensystem basteln, dessen Grundobjekte Mengen und Stufen sind und wir können von diesem Axiomensystem, das *allein* auf (intuitiv einsichtigen) Beobachtungen zu diesem iterativen Prozess besteht (das also *auf jeden Fall* den iterativen Begriff formalisiert), alle Axiome von ZF ableiten.¹⁰⁸

Wir wollen beispielhaft zeigen, wie das für das Aussonderungsaxiom funktioniert und erhalten dadurch eine inhaltliche Rechtfertigung des Aussonde-

¹⁰⁷Die seltsame Ausdrucksweise in diesem Satz verweist auf das Problem, was man unter *Bildung* verstehen sollte. Kurz gesagt: Werden die Mengen auf jeder Stufe *tatsächlich* gebildet, oder wird nur gefordert, dass man sie bilden *kann*? Wang schlägt vor, dass man den iterativen Begriff 'maximal' interpretiert, d.h. wir betrachten jede Menge auf einer bestimmten Stufe als *verfügbar*, wenn wir sie *hätten bilden können*. *Existenz* einer Menge auf einer bestimmten Stufe bedeutet demnach, dass sie auf dieser Stufe *hätte gebildet werden können*. Vgl. dazu Wang [1974], wiederabgedruckt in Benacerraf/Putnam [1983] S. 530 ff. und dazu Parsons ebd. S. 502 ff.

¹⁰⁸Boolos selbst meint, das Schema der Ersetzungsaxiome folge nicht aus dem iterativen Begriff. Es gibt aber Rechtfertigungen aus dem iterativen Begriff z.B. von Wang. Von einer neueren Axiomatisierung, dem *Scott'schen Axiomensystem*, das ebenfalls den iterativen Begriff formalisiert (oder zumindest den Anspruch hat, ihn zu formalisieren), und aus der *stage theory* von Boolos hervorging, kann man ebenfalls zeigen, dass sie äquivalent zu ZF ist (Vgl. dazu Ebbinghaus [2003]) Dieses Axiomensystem basiert aber auf einem Reflexionsaxiom, das - grob gesprochen - besagt, dass jede Aussage, die im Universum gilt, auch schon in einer bestimmten Stufe gilt und umgekehrt. In ZF benötigt dieses Reflexionstheorem aber wesentlich die Ersetzungsaxiome, was die Argumentation zirkulär macht, falls man der Meinung ist, die Ersetzungsaxiome würden nicht in der ein oder anderen Weise aus dem intuitiven iterativen Begriff folgen.

rungsschemas (beinahe) *unmittelbar* aus dem iterativen Begriff der Menge.¹⁰⁹ Für das Folgende stelle man sich unter Variablen s, t *Stufen* vor (intuitiv: *Von-Neumannsche Stufen*) und unter Variablen x, y *Mengen*, E stehe für die Relation zwischen zwei Stufen ‘kommt in der Reihenfolge früher’ und F bedeute die Relation ‘wird gebildet auf der Stufe’ zwischen einer Menge und einer Stufe.

Die entscheidende Beobachtung, die wir benötigen, ist folgender Sachverhalt:

Weil jede Stufe aus allen möglichen Zusammenfassungen von Zusammenfassungen niedrigerer Stufen (und Individuen) besteht, sollte es möglich sein, beliebige Mengen von Mengen, die auf einer bestimmten Stufe ‘verfügbar’ sind, wiederum zusammenzufassen. Es sollte demnach möglich sein, *alle Objekte einer Stufe mit einer bestimmten Eigenschaft* zusammenzufassen (einfach weil *alle Zusammenfassungen* von Objekten einer bestimmten Stufe erlaubt sind). In der ‘*stage theory*’ (so nennt Boolos die gerade betrachtete Theorie) wird das ausgedrückt durch das Schema (Boolos nennt sie die *Axiome der Spezifikation*)

$$\forall s \exists y \forall x (x \in y \leftrightarrow (\phi(x) \wedge \exists t (tEs \wedge xFt)))$$

In Worten: Für jede Stufe existiert die Menge aller Objekte x , die auf einer niedrigeren Stufe gebildet wurden und der Bedingung $\phi(x)$ genügen.

Man sieht leicht, dass daraus das Aussonderungsschema folgt, wenn man bedenkt, dass jedes z in $\forall z \exists y \forall x (x \in y \leftrightarrow x \in z \wedge \phi(x))$ auf einer bestimmten Stufe gebildet worden sein muss. Genauer: wir wählen für $\phi(x)$ die Formel $x \in z \wedge \phi(x)$ und nehmen an, z sei auf der Stufe s gebildet worden. Aus der dritten und vierten Beobachtung von oben ergibt sich nun, dass alle $x \in z$ auf früheren Stufen als s gebildet worden sein müssen, d.h.: y ist die Menge aller $x \in z$ die $\phi(x)$ erfüllen.

Betrachtet man das Aussonderungsschema in dieser Weise als *vom iterativen Begriff abgeleitet*, so verliert sich auch der (historisch korrekte, aber inhaltlich falsche) Eindruck, es wäre wie das S-Komprehensionsschema nur eine Art der Beschränkung des naiven Komprehensionsprinzips, nur dazu da, den naiven Begriff der Menge zu ‘retten’.

In diesem Sinne ist auch die Bemerkung aus dem 4. Abschnitt zu verstehen, dass das S-Komprehensionsschema und das Aussonderungsschema ‘nicht zu vergleichen’ sind, weil sie ‘verschiedene Dinge wollen’ (und überdies die Bemerkung im Vorwort, wo von einer *bestimmten Geschichte der Mengenlehre* die Rede war).

Entscheidend für den iterativen Begriff der Menge ist folgendes Bild der Mengenbildung: Wir gehen von - in einem bestimmten Sinne - *vorher gegebenen* Objekten aus und bilden aus diesen *durch Zusammenfassung* Mengen. Es soll hier nicht behauptet werden, dass dieser Begriff ‘philosophisch’ restlos klar ist. Man kann sich etwa fragen, was hier überhaupt ‘zusammenfassen’ oder ‘vorher gegeben’ heißen soll; ist mit ‘zusammenfassen’ eine *Tätigkeit des Geistes* gemeint?

¹⁰⁹Nur *beinahe* unmittelbar, weil sich natürlich alle ‘Beobachtungen’ des iterativen Prozesses selbst wiederum auf ‘intuitive Einsichtigkeit’ berufen müssen.

‘Erschaffen’ wir wir beim Zusammenfassen buchstäblich ein *Ding* aus unserem Geist, das wir *Menge* nennen? Ist das ‘vorher gegeben’ tatsächlich *zeitlich* zu verstehen¹¹⁰ oder ist damit “priority in order of construction” gemeint?¹¹¹

All das hat uns nicht zu bekümmern - der iterative Begriff ist ein *intuitiver Begriff* und Ziel dieser Untersuchung ist nicht, diesen Begriff bis ins Letzte auszubuchstabieren, sondern ihn vom intuitiven *logischen* Begriff *abzugrenzen*. Der entscheidende Punkt ist: Aus der Sicht des *iterativen Begriffs* bilden wir Mengen durch *Zusammenfassung*, während wir aus der Sicht des *logischen Begriffs* Mengen über *Eigenschaften* generieren.¹¹²

8.4 Intuitive Begriffe, Formalisierungen und Vergleichbarkeit von Mengenlehren

Wir haben bisher gesehen, dass es sich bei NF(U) und ZFC tatsächlich um wesentlich verschiedene Systeme handelt, und zwar sowohl ‘mathematisch’ (wir konnten z.B. in ZFC, nicht aber in NFU, beweisen, dass die Potenzmenge jeder Menge echt größer als diese selbst ist), als auch bzgl. der Intuitionen, die die Systeme motivieren. Dennoch würde es zu weit gehen, wenn man meinte, die beiden Intuitionen sprechen über etwas *völlig Anderes*, wenn sie das Wort ‘Menge’ benutzen. Beide scheinen auf jeweils verschiedene Weisen *etwas* vom Begriff ‘Menge’ zu erfassen (abgesehen vom rein quantitativen Aspekt der Extensionalität). Es handelt sich bei beiden Begriffen also nicht um *absolut unvergleichbare* Begriffe: Es scheint so zu sein, als hätten beide jeweils Anteil an einem ‘ursprünglicheren’ Begriff, der sich nicht eindeutig in logischen oder iterativen Begriff kategorisieren lässt.

Bevor wir uns diesen Punkt genauer ansehen, werfen wir noch einen kurzen Blick auf die Rolle von Axiomatisierungen oder Formalisierungen von intuitiven mathematischen Begriffen im Allgemeinen.

Das Verhältnis von intuitiven Begriffen zu deren formalisierten Versionen kann von Fall zu Fall sehr verschieden sein. Während es manchmal so ist, dass ein jahrhundertlang gebrauchter Begriff sich in seiner formalisierten Version als *zu weit* erweist, um den intuitiven Begriff genau zu treffen, gibt es andere Fälle, wo die Formalisierung sich als *zu eng* erweist, um alle intuitiven Fälle abzudecken, die man abgedeckt haben *will*. Ein Beispiel für den ersten Fall wäre der Begriff der Stetigkeit, von dem man sich gut vorstellen kann, dass er einmal ‘ohne abzusetzen zeichenbar’ bedeutet hat (so wie man das in der Schule lernt). Als Weierstrass dann seine ϵ - δ -Definition vorgelegt hat, hat sich dann schnell herausgestellt, dass *diese* Definition offenbar nicht mit dem intuitiven Begriff

¹¹⁰Michael Potter treibt mit seiner blumigen Ausdrucksweise vom ‘Geburtstag’ einer Menge (für den Rang), ‘Tagen’ (für die Stufen) usw. die zeitliche Metapher auf die Spitze.

¹¹¹Vgl. dazu Parsons [1975], wiederabgedruckt in Putnam/Benacerraf [1983] S. 503 ff.

¹¹²Das relativiert auch die zum Eingang dieses Kapitels erwähnten ‘Gemeinsamkeiten’ von ZFC und NF(U): Dass wir z.B. für a, b in NF(U) die Paarmenge $\{a, b\}$ bilden können, liegt daran, dass es eine (stratifizierte) *Eigenschaft* $\varphi(x) : \leftrightarrow x = a \vee x = b$ gibt, die diese Menge bestimmt - in ZFC dagegen daran, dass wir beliebige vorher gegebene Objekte a, b zu einer neuen Menge *zusammenfassen* können.

übereinstimmt. In diesem Fall hat sich die Bezeichnung des intuitiven Begriffs für den zu allgemeinen formalisierten Begriff durchgesetzt und es wurden weitere Differenzierungen getroffen (gleichmäßige Stetigkeit, Lipschitz-Stetigkeit, usw.). In anderen Fällen ist es so, dass wir verallgemeinern müssen, um zu dem Begriff zu gelangen, den wir eigentlich einfangen *wollen*.

Aber welche *Bezeichnung* wir für eine formalisierte Version eines intuitiven Begriffs verwenden, ist im wesentlichen *konventionell*. Jahrhundertlang wurde z.B. der Begriff ‘Geometrie’ ausschließlich für das verwendet, was wir *heute* ‘Euklidische Geometrie’ nennen. D.h., wie wir das Wort ‘Geometrie’ verwenden, hat sich mit der Entdeckung nicht-euklidischer Geometrien verändert. Von diesem Zeitpunkt an ist ‘Euklidische Geometrie’ ein Spezialfall des allgemeineren Begriffs ‘Geometrie’.

All das ist, wie gesagt, konventionell, und hängt von bestimmten Umständen, Zeiten und besonderen Autoritäten ab.

Formalisierungen können aber auch in einer anderen Hinsicht differenziert werden. Nehmen wir z.B. die Gruppenaxiome: Die Gruppenaxiome wurden aufgestellt, um einer bestimmten algebraischen Struktur, die in der Mathematik immer wieder auftaucht einen *Namen* zu geben. Wesentlich dabei ist der *Klassifikationsaspekt*. (Das Erlernen der Axiome für die algebraischen Strukturen (Monoid, Gruppe, abelsche Gruppe, Ring, Körper usw.) ähnelt eher dem Erlernen des Linné’schen Systems als dem Betreiben von Mathematik.) Man bemerkt, dass bei bestimmten Untersuchungen bestimmte Dinge immer wieder auftreten und *benennt* sie - ähnlich wie man bestimmten Zahlen wie der Euler’schen Zahl e oder π *Namen* gibt.

Für die Peano-Axiome für die natürlichen Zahlen gilt das scheinbar nicht. Man hat nicht das Gefühl, dass man damit etwas *definiert*, sondern etwas *expliziert*, das schon ‘*da*’ ist. Jeder hat ein intuitives Vorverständnis von den natürlichen Zahlen, und dieses Vorverständnis wollen wir durch die Peano-Axiome *in Formeln gießen*.

Haben wir einen sehr klaren intuitiven Begriff von einem zu formalisierenden Objekt oder einer Theorie, so wird eine formalisierte Version offenbar weniger Spielraum lassen - sie wird weder allzu *weit* noch allzu *eng* sein. Mit den (second-order) Peano-Axiomen etwa können wir die Struktur der natürlichen Zahlen *kategorisch* (d.h. *bis auf Isomorphie eindeutig*) festlegen.

Wie steht es mit einer Axiomatisierung des intuitiven Mengenbegriffs? Hoffnung auf eine (auch nur annähernd) *kategorische* Charakterisierung des Mengenuniversums wird kaum bestehen.¹¹³ Abgesehen davon gibt es *den* intuitiven Mengenbegriff nicht und auch für einen *bestimmten* intuitiven Begriff gibt es eine Reihe (zum Teil unvereinbarer) Axiomatisierungen. Während man von den natürlichen Zahlen noch mit einem gewissen Recht an ihre ‘unabhängig Exis-

¹¹³Das verbietet schon der Satz von Löwenheim-Skolem, der besagt, dass jede first-order Theorie mit unendlichen Modellen, unendliche Modelle *jeder Mächtigkeit* hat.

tenz' in einem 'platonischen' Sinne glauben könnte, wäre ein Glaube an ein 'platonisches Mengenuniversum' ziemlich vermessen angesichts des bisher Vor-gebrachten.

Die Güte einer Axiomatisierung muss sich daher (zunächst) daran bemessen, inwieweit sie einen *bestimmten* intuitiven Begriff 'korrekt' formalisiert.

'Güte-Fragen' dieser Art wären z.B.:

Welches System (Scott'sches Axiomensystem, $ZF(C)$, stage theory,...) formalisiert den iterativen Begriff angemessener?

In welcher Weise modifiziert man die naive Mengenlehre am geeignetsten, um so viel wie möglich von den Intuitionen dahinter zu retten? D.h. wie bewerkstellige ich es, dass so viele Eigenschaften wie möglich Mengen bestimmen, ohne dass es zu Widersprüchen kommt?

Fragen dieser Art betreffen aber wie gesagt nur das Verhältnis eines bestimmten Mengenbegriffs zu seinen Formalisierungen. Um in weiterer Folge dann 'globalere' Perspektiven zu eröffnen, d.h. um Axiomatisierungen verschiedener Mengenbegriffe vergleichen zu können, bedarf es anderer Maßstäbe. Es stellt sich aber die Frage, was so ein Maßstab sein kann. Sicherlich wollen wir den (intuitiv-) mathematischen Mengenbegriff nicht auf historisch-sprachwissenschaftliche Studien gegründet wissen (z.B. darauf wie das Wort 'Menge' im Mittelhochdeutschen verwendet wurde). Wir werden auch Begriffe wie Dedekinds 'System' als etwas auffassen wollen, das auf den Begriff 'Menge' verweist.

Es bieten sich verschiedene Vergleichsmöglichkeiten an:

Wir könnten die konkurrierenden Begriffe z.B. in Hinblick auf bestimmte *anerkannte Autoritäten* beurteilen. Cantor bietet sich dafür besonders an, weil er von den meisten als unumstrittener Begründer der Mengenlehre angesehen wird.

Das hat den Vorteil, dass wir nun (wenn auch möglicherweise schwammige) Kriterien zur Beurteilung der verschiedenen Mengenbegriffe zur Verfügung haben. Andererseits sieht das sehr nach Willkür aus; wieso sollte man ausgerechnet Cantors Mengenbegriff für den Mengenbegriff schlechthin halten? Was ein bestimmter Mathematiker für entscheidend hält, muss nicht unbedingt das sein, womit wir es im folgenden zu tun haben *wollen*, und auch nicht das, was der betreffende Mathematiker tatsächlich *gemacht* hat. Burali-Forti hat in seinem 'Beweis', dass die Gesamtheit der Ordinalzahlen mit der üblichen Ordnung keine Wohlordnung ist, den Begriff der *perfekt geordneten Menge* benutzt, den er für äquivalent zur *wohlgeordneten Menge* Cantors gehalten hat. Er hat also einen Begriff der 'Wohlordnung' verwendet, über den wir eigentlich gar nichts wissen wollten.

Diese Situation ist in der Mathematik nicht selten - immer wieder kommt es vor, dass Definitionen *zu eng* oder *zu weit* sind und dass Mathematiker Dinge 'beweisen', die eigentlich uninteressant sind, weil abhängig von 'schlechten' Definitionen. Wieso also den Mengenbegriff Cantors (oder den irgendeiner anderen 'Autorität') für bare Münze nehmen?

Weitere Vergleichskriterien finden sich pragmatisch-technisch im *Erfolg* und in der *Bequemlichkeit*. Beide sind aber für uns, die wir Mengenlehren mit verschiedenen dahinter liegenden intuitiven Mengenbegriffen vergleichen wollen, unbrauchbar.

Sehen wir von der Minimalforderung nach Repräsentierbarkeit aller möglichen mathematischen Begriffe und Theorien ab (d.h. von der Forderung, ‘klassische Mathematik’ (Analysis, Zahlentheorie, Topologie usw.) betreiben zu können), so bemisst sich der *Erfolg* einer Mengenlehre ja (unter anderem) daran, inwieweit diese Mengenlehre fähig ist intuitiv einsichtige oder ‘klassische’ mengentheoretische Resultate zu generieren. Das verweist aber geradewegs zurück auf die Frage nach den *Intuitionen* über den Begriff ‘Menge’. Wir wollen ja gerade wissen, welcher Mengenbegriff zweier rivalisierender Mengenlehren der ‘bessere’, ‘einsichtiger’ ist.

Auch die *Bequemlichkeit* taugt nur eingeschränkt als Vergleichskriterium, nämlich für *verschiedene* Formalisierungen *desselben* intuitiven Mengenbegriffs. *Bequem* kann ja nur heißen, *bequemer zum Erreichen eines bestimmten Ziels*, d.h. die Resultate, die wir erreichen wollen, müssen schon vorgegeben sein.

Wir könnten aber auch sagen, dass solche Systeme, die verschiedenen Intuitionen entspringen, im Prinzip *unvergleichbar bzgl. ihrer Güte* sind. D.h. wir können ihre Verschiedenheit zwar aufzeigen, und auch vergleichen, in dem Sinne, dass wir klar machen können, worauf sie ihre Akzente legen (so wie wir das auch bisher getan haben) - aber wir können, wie Quine es formuliert, keine Mengenlehre “in offensichtlicher Weise als Standardsystem auszeichnen”, denn dazu bräuchten wir ‘absolute’ Vergleichskriterien.

Diese Haltung scheint in Ermangelung eindeutiger Vergleichskriterien tatsächlich die vernünftigste zu sein und man sollte mißtrauisch gegenüber Quines Optimismus sein, wonach “die Forschung vielleicht doch eines Tages eine Mengenlehre entwickelt, die ganz klar die beste ist”.¹¹⁴

Man sollte intuitive Mengenbegriffe *normativ* verstehen in einem ähnlichen Sinne, wie Shapiro den Begriff der ‘Berechenbarkeit’ *normativ* verstanden wissen will.¹¹⁵ Jeder intuitive Mengenbegriff schafft sich in gewisser Weise ‘seine eigene Realität’.

Verschiedene Mengenbegriffe, wie der *iterative* oder der *modifizierte logische*, wie wir sie besprochen haben, machen *Angebote*, was man unter einer Menge verstehen *sollte* - und verschiedene Formalisierungen dieser verschiedenen Mengenbegriffe machen wiederum *Angebote*, wie man diese Mengenbegriffe adäquat und fruchtbar fixieren *sollte* - die Formalisierungen selbst laufen dann, so wie es in der Mathematik sein muss - wie eine Maschine - unabhängig von den davor ‘hineingesteckten’ Intuitionen. Falls sich ungewollte Konsequenzen (oder sogar Widersprüche) aus den Axiomen ergeben, so kann das verschiedene Ursachen haben. Es könnte bei der Fixierung (beim Übergang vom intuitiven zum for-

¹¹⁴Vgl. Quine [1973] Vorwort S. VII

¹¹⁵Shapiro [2005]

malisierten Begriff) etwas schief gelaufen sein, oder die Intuitionen selbst waren unzureichend, zu undeutlich oder schlicht *falsch* (auch diese Möglichkeit muss natürlich gegeben sein). Das kann Korrekturen auf der intuitiven oder der formalen Seite mit sich ziehen. Das übliche mathematische Prozedere kommt in Gange: Vorschläge zur Berichtigung werden gemacht - manche werden akzeptiert, manche werden verworfen usw. kurz: es werden neue '*Angebote*' gemacht, wie man Begriffe verstehen *sollte*.

Diese 'relativistische' Sichtweise hat den Vorteil, dass wir nicht auf einen 'platonistischen' Mengenbegriff Bezug nehmen müssen, an dem sich jede Mengenlehre messen müsste; sie entbindet uns von der Bürde, 'endgültig' entscheiden zu müssen, welche Sätze über Mengen 'wahr' und welche 'falsch' sind - sie hindert uns aber auch nicht daran, uns in einer 'informellen' Weise Gedanken über die Wahrheit eines Satzes *für einen bestimmten Begriff (in einem intuitiven Modell)* zu machen.¹¹⁶

Sie kann uns nicht daran hindern einen bestimmten Mengenbegriff für 'einsichtiger' als einen anderen zu halten - sie kann uns aber vor allem nicht dazu zwingen, einen bestimmten Begriff der Menge und eine bestimmte dazugehörige Formalisierung für den/die 'richtige(n)' zu halten.

¹¹⁶Die Suche nach intuitiv einsichtigen Axiomen, die die Kontinuumshypothese für ZFC entscheiden, gehört in diesen Bereich

9 Anhang

Abstract

Von Beginn an ist dasjenige Gebiet der mathematischen Grundlagenforschung, das man ‘Mengenlehre’ nennt, heiß umfehdet – und zwar nicht bloß in ‘mathematischer’, sondern auch in philosophischer Hinsicht.

Formal ist eine (axiomatische) Mengenlehre eine mathematische Theorie wie jede andere, gleichgestellt der Gruppentheorie, der Peano-Arithmetik erster Stufe oder der Theorie der partiellen Ordnungen. *De facto* ist sie *primus inter pares*, weil es möglich ist alle anderen mathematischen Theorien innderhalb der Mengenlehre zu ‘verhandeln’.

Unmengen an verschiedenen Systemen der axiomatischen Mengenlehre gibt es mittlerweile, beinahe jeder große Logiker des zwanzigsten Jahrhunderts hat ein eigenes. Einige unterscheiden sich nur graduell, andere entstammen grundverschiedenen Denktraditionen – und dennoch beanspruchen alle, den Begriff der ‘Menge’ angemessen zu formalisieren.

Zwei solche Systeme (NFU als Repräsentant einer eher ‘logisch’ orientierten Denktradition und ZFC als Repräsentant einer genuin ‘mathematischen’ Tradition) sollen in dieser Arbeit gegenübergestellt werden. Dabei ist ein erstes Ziel, herauszustellen, welche Sätze in dem einen System beweisbar sind, nicht jedoch im anderen (und umgekehrt). Vor allem in der transfiniten Arithmetik wird sich zeigen, dass die Unterschiede vielfältig sind.

Das zweite (und vorrangige) Ziel besteht darin, aufzuzeigen, welche konkurrierenden Intuitionen zum Begriff ‘Menge’ die beiden Systeme motivieren, und in welchem Verhältnis diese Intuitionen zu deren formalisierten Versionen und zueinander stehen.

Lebenslauf

Ich wurde am 9. Juni 1982 in Bruck/Mur geboren. Nach der Volks- und Hauptschule in Allerheiligen (bzw. Kindberg) absolvierte ich die Handelsakademie in Bruck/Mur. Ich studiere seit 2003 Philosophie in Wien, wo ich mich hauptsächlich mit analytischer Philosophie, Logik und Philosophie der Mathematik beschäftigt habe. Seit 2005 studiere ich auch Mathematik an der Universität Wien.

10 Literatur

Aczel [1988]; Peter Aczel, Non-well-founded sets, Center for the Study of Language and Information, Leland Stanford Junior University, Printed in the US, 1988

Benacerraf [1965]; Paul Benacerraf, What numbers could not be, wiederabgedruckt in Benacerraf/Putnam [1983], S. 272 - 294

Benacerraf/Putnam [1983]; Paul Benacerraf, Hilary Putnam (Hrsg.); Philosophy of Mathematics, Cambridge University Press 1983

Boolos [1971]; George Boolos, The iterative conception of set, Journal of Philosophy, vol. 68, wiederabgedruckt in Benacerraf/Putnam [1983] S. 486 - 502

Boolos [1989]; George Boolos, Iteration Again, Philosophical Topics 42, wiederabgedruckt in Boolos [1998] S. 88 - 104

Boolos [1997]; George Boolos, Must we believe in set theory?, wiederabgedruckt in Boolos [1998] S. 120 - 142

Boolos [1998]; George Boolos, Logic, Logic, and Logic, Harvard University Press 1998, herausgegeben von Richard C. Jeffrey

Cantor [1883]; Georg Cantor, Grundlagen einer allgemeinen Mannigfaltigkeitslehre, wiederabgedruckt in Cantor [1932] S. 165 - 207

Cantor [1895]; Georg Cantor, Beiträge zur Begründung der transfiniten Mengenlehre, wiederabgedruckt in Cantor [1932] S. 282 - 356

Cantor [1899]; Georg Cantor, Brief an Dedekind, wiederabgedruckt in [1932] S. 442 - 451

Cantor [1932]; Georg Cantor, Gesammelte Abhandlungen mathematischen und philosophischen Inhalts, herausgegeben von E. Zermelo, Hildesheim 1932

Deiser [2002]; Oliver Deiser, Einführung in die Mengenlehre, Springer Verlag, Berlin - Heidelberg - New York 2002

Ebbinghaus [2003]; Heinz-Dieter Ebbinghaus, Einführung in die Mengenlehre, Spektrum Akademischer Verlag 2003

Frege [1891]; Gottlob Frege, Funktion und Begriff, wiederabgedruckt in Frege [2002] S. 2 - 22

Frege [2002]; Mark Textor (Hrsg.), Gottlob Frege, Funktion - Begriff - Bedeutung, Vandenhoeck und Ruprecht 2002

Frege [1884]; Gottlob Frege, Die Grundlagen der Arithmetik, Reclam 1987

Frege [1893]; Gottlob Frege, Grundgesetze der Arithmetik, Hildesheim 1966

Forster [1992]; T.E. Forster, Set theory with a Universal Set, Clarendon Press, Oxford 1992

Forster [2008]; T.E. Forster, The iterative Conception of Set, in The Review of Symbolic Logic Vol. 1, No. 1, 2008

Gödel [1947]; Kurt Gödel, What is Cantor's continuum problem?, wiederabgedruckt in Benacerraf/Putnam [1983]

Gödel [1944]; Kurt Gödel, Russells Mathematical Logic, wiederabgedruckt in Benacerraf/Putnam [1983]

Hailperin [1944]; T. Hailperin, A set of axioms for logic, in Journal of symbolic logic 9, S. 1 - 19

Henson [1973]; C. Ward Henson, Type-Raising Operations on Cardinal and Ordinal Numbers in Quine's "New Foundations", in The journal of Symbolic Logic, Vol. 38, No. 1, März 1973

Holmes [1998]; M. Randall Holmes, Elementary Set Theory with a universal set, in Cahier du centre logique, Vol. 10

Holmes [1994]; M. Randall Holmes, The set-theoretical program of Quine succeeded, but nobody noticed, in Modern Logic, Vol. 4, No. 1 (Jänner 1994)

Van Heijenoort [1967]; Jean van Heijenoort (Hrsg.), From Frege to Gödel, A source Book in Mathematical Logic, 1879 - 1931, 1967

Jensen [1969]; R.B. Jensen, On the consistency of a slight(?) modification of Quines NF. in Synthese 19, S 250 - 263

Parsons [1975]; Charles Parsons, What is the iterative Conception of set?, wiederabgedruckt in Benacerraf/Putnam [1983]

Peano [1889]; Guiseppe Peano, The principles of arithmetic, presented by a new method, wiederabgedruckt in Van Heijenoort [1967] S. 83 - 97

Potter [1993]; Michael Potter, Iterative set theory, in The Philosophical Quarterly Vol. 43, No. 171, 1993

Quine [1937]; W.V. Quine, New Foundations for mathematical logic, wiederabgedruckt in Quine [1980]

Quine [1937a]; W.V. Quine, On Cantors Theorem, in The Journal of symbolic logic, Vol. 2, No. 3, September 1937

Quine [1945]; W.V. Quine, On ordered pairs, The Journal of symbolic logic, Vol. 10, No. 3, September 1945

Quine [1964]; W.V. Quine, Grundzüge der logic, Suhrkamp Verlag Frankfurt am Main 1969

Quine [1970]; W.V. Quine, Philosophie der Logik, C.C.Buchners Verlag, Bamberg 2005

Quine [1973]; W.V. Quine, Mengenlehre und ihre Logik, Vieweg, Braunschweig 1973

Quine [1980]; W.V. Quine, From a logical point of view, Harvard University Press, Cambridge, Massachusetts and London, England 1980

Ramharter/Rieckh [2006]; Esther Ramharter, Georg Rieckh, Die Principia Mathematica auf den Punkt gebracht, öbvht VerlagsgmbH und Co. KG, Wien 2006

Rosser [1939]; J. Barkley Rosser, Definition by Induction in Quine's New Foundations for mathematical logic, The journal of symbolic logic Vol. 4, No. 2, Juni 1939

Rosser [1942]; J. Barkley Rosser, The Burali-Forti Paradox, in The Journal of symbolic logic Vol. 7, No. 1, März 1942

Rosser [1952]; J. Barkley Rosser, The axiom of infinity in Quines New Foundations, in The Journal of symbolic logic Vol. 17, No. 9, Dezember 1952

Rosser [1953]; J. Barkley Rosser, Logic for mathematicians, McGraw Hill Book Company Inc. 1953

Russell/Whitehead [1910]; A. N. Whitehead, B. Russell, Principia Mathematica, second edition, Cambridge University Press 1968

Russell [1903]; B. Russell, The Principles of Mathematics, W.W. Norton and Company 1996

Russell [1908]; B. Russell, Mathematical Logic as based on the theory of types, wiederabgedruckt in Van Heijenoort [1967] S. 150 - 182

Shapiro [2006]; Stewart Shapiro, Computability, Proof, and Open-Texture, in Olszewski, Wolenski, Janusz (Hrsg.), Churchs Thesis after 70 years, Ontos Verlag 2006, S. 421 - 455

Specker [1953]; Ernst Specker, The axiom of choice in Quines new foundations for mathematical logic, in Proceedings of the National Academy of Sciences of the USA 39, S. 972-975

Wang [1974]; Hao Wang, The concept of set, wiederabgedruckt in Benacerraf/Putnam [1983]

Walczak-Typke [2007]; Agatha Walczak-Typke, Axiomatic Set Theory I, Skriptum zur Vorlesung 'Axiomatic Set Theory I' WS 2007

Wright [2007]; Crispin Wright, On Quantifying into Predicate Position: Steps towards a New(tralist) Perspective, in Michael Potter (Hrsg.), Mathematical Knowledge, Oxford University Press 2007, S. 150 - 174

Zermelo [1908]; Investigations in the foundations of set theory I, wiederabgedruckt in Van Heijenoort [1967] S. 199 - 215