On the Relationship between Choice Schemes and Iterated Class Comprehension in Set Theory

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Jürg Krähenbühl

von Trub

Leiter der Arbeit:
Prof. Dr. G. Jäger
Institut für Informatik und angewandte Mathematik

Von der Philosophisch-naturwissenschaftlichen Fakultät angenommen.

Bern, 19. 12. 2011

Der Dekan:
Prof. Dr. S. Decurtins

Für Christine

Abstract

The aim of this thesis is to show a few specific results about extensions of Von Neumann–Bernays–Gödel set theory NBG, by applying proof-theoretic techniques. We get the main results in a uniform way, by using cut-elimination and asymmetric interpretations. The same technique was applied a few decades ago, to analogous systems of second order arithmetic, by Cantini [1].

We consider natural extensions of NBG by a few axiom schemes, i.e., choice $AC[\Sigma^1_1]$, dependent choice $DC[\Sigma^1_1]$, full induction $TI_\in[\mathcal{L}^1]$, and iterated elementary comprehension $(CA[\Pi^1_0])_{<\mathsf{c}}$. And we are going to establish proof-theoretic equivalences between these schemes, similar to the results for analogous systems of arithmetic. The equivalences proven in this thesis are

$$\begin{array}{rcl} \mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] & \equiv & \mathcal{T} \cup \mathsf{NBG}, \\ \mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<\Omega^\omega}, \\ \mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \cup TI_\in[\mathcal{L}^1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}, \\ \mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \cup TI_\in[\mathcal{L}^1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}, \\ \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0} \cup TI_\in[\mathcal{L}^1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}, \end{array}$$

where $\mathcal{T}$ is any set of axioms with logical complexity essentially Σ^1_2, and the equivalence, $\equiv$, means that any sentence essentially Π^1_2 is either provable in both theories or in none of them.

The first equivalence has already been stated (without proof) by Feferman and Sieg [4]. The last two equivalences are easy consequences, by using the third one. The second last equivalence has already been shown in a slightly weaker form by Jäger and Krähenbühl [10].

Acknowledgements

First of all, I want to thank Professor Gerhard Jäger for his continuous support and his valuable advice on my thesis project.

Further, I am greatful to Thomas Strahm and Thomas Studer who always are great sources of inspiration on technical topics as well as on informal ones.

I also want to thank all the members of the TIL group for making my life much more comfortable during research; many thanks go to Luca Alberucci, Dieter Probst, David Steiner, Daria Spescha, Peppo Brambilla, Roman Kuznets, Dandolo Flumini, Remo Goetschi, Bettina Choffat, Michael Wachter, Jacek Jonczy, Rico Zumbrunnen, Sebastian Eberhard, and Ricardo Wehbe.

Many thanks also go to Professor Andrea Cantini for his careful reading of this thesis.

And last but not least, I want to thank my beloved family and my close friends for their steady support and the many good times we had.

Contents

Introduction

Von Neumann–Bernays–Gödel set theory (NBG) is a conservative extension of Zermelo–Fraenkel set theory (ZFC), see e.g. Levy [13]. NBG extends ZFC such that, in addition to sets, we also have *classes* as individual objects. The two sorts of variables $x, y, z, ...$ and $X, Y, Z, ...$ for sets and classes, respectively, make the language much more expressive, e.g., NBG is finitely axiomatizable. NBG can be formulated in many different ways; in this thesis we define NBG to consist of ZFC and the following (infinitely many) axioms

$$\begin{array}{ll} \text{(Comprehension)} & \exists X \forall y (y \in X \leftrightarrow A[y]) \text{ for any } A \in \Pi^1_0, \\ \text{(Replacement)} & \forall F (Fun[F] \to \forall x \exists y (y = F[x])), \\ \text{(Global Choice)} & \exists F (Fun[F] \wedge \forall x (x = \emptyset \vee F(x) \in x)). \end{array}$$

All the theorems about sets in ZFC are exactly the same as in NBG, i.e., “pure set theory” is not affected by the classes in NBG. Of course, the underlying set theory can be easily made stronger, by adding appropriate class axioms to NBG, e.g. this is the case for Morse–Kelly set theory (MK), which is just NBG with comprehension allowed for any formula A. The increased expressiveness and the conservation of common sense set theory ZFC within NBG, make NBG to an interesting choice for logical investigations. Furthermore, on the meta-level of logic, the extension of ZFC to NBG is similar to the shift from Peano Arithmetic (PA) to Arithmetical Comprehension (ACA_0), which is a subsystem of Second Order Arithmetic (Z_2) (MK corresponds to Z_2, in the same way as NBG corresponds to ACA_0). This logical analogy is the starting point of this thesis. A huge amount of research in mathematical logic has been done in the field of arithmetic, and the aim of this thesis is to get a few specific results about set theory, by applying proof-theoretic techniques, i.e. cut-elimination and asymmetric interpretation, analogously to the way these techniques were applied to arithmetic.

We consider natural extensions of NBG by axiom schemes with classes involved, i.e., choice $AC[\Sigma^1_1]$, dependent choice $DC[\Sigma^1_1]$, induction $TI_\in[\mathcal{L}^1]$, and iterated comprehension $(CA[\Pi^1_0])_{<\mathsf{c}}$. And we are going to establish proof-theoretic equivalences between these schemes, similar to some results that were achieved a few decades ago for analogous systems of arithmetic.

We observe that any class Y can be considered to be a collection of classes, $(Y)_x$, where $(Y)_x$ is the class $\{z \mid \langle z, x\rangle \in Y\}$, or it can be considered to be a function $Y : x \mapsto (Y)_x$, mapping sets to classes. In this context it makes sense to have the restriction $(Y)^x$, i.e. the function Y restricted to x, where $(Y)^x$ is the class $\{\langle z, y\rangle \in Y \mid y \in x\}$, or more generally, $(Y)^{<x}$ is the class $\{\langle z, y\rangle \in Y \mid y < x\}$ for any relation $<$. Having this notation, the axiom schemes of interest can now be formulated precisely as the following

- Choice $AC[\Sigma^1_1]$, means that for any $A \in \Sigma^1_1$ we have that

$$\forall x \exists Y A[x, Y] \to \exists Y \forall x A[x, (Y)_x],$$

  that is, if we have $\forall x \exists Y A[x, Y]$ then there also exists a "choice function" Y such that $A[x, (Y)_x]$ for all x.

- Dependent Choice $DC[\Sigma^1_1]$, means that for any $A \in \Sigma^1_1$ we have that

$$\forall x \forall Y \exists Z A[x, Y, Z] \to \exists Z \forall x A[x, (Z)^x, (Z)_x],$$

  i.e., if we have $\forall x \forall Y \exists Z A[x, Y, Z]$ then there is a "choice function" Z such that $A[x, (Z)^x, (Z)_x]$ for all x, hence $(Z)_x$ depends on $(Z)^x$, that is, $(Z)_x$ depends on the choices made "previous" to x.

- Induction $TI_\in[\mathcal{L}^1]$ means that for all formulas A we have

$$\forall x((\forall y \in x)A[y] \to A[x]) \to \forall x A[x],$$

  i.e., from the progressivity of the element relation $\in$ on $A[x]$ we get $A[x]$ for all x.

- Iterated comprehension $(CA[\Pi^1_0])_{<\mathsf{c}}$ means that for any $A \in \Pi^1_0$ we have

$$\exists X(\forall y < \mathsf{c})(X)_y = \{z \mid A[z, y, (X)^{<y}]\},$$

  i.e., there exists a class hierarchy X, such that for all levels y "below" c the class $(X)_y$ consists of the sets z with $A[z, y, (X)^{<y}]$, hence $(X)_y$ depends on all levels of the hierarchy "previous" to y, that is $(X)^{<y}$.

The main results of this thesis can now be stated as the following proof-theoretic equivalences

$$\begin{array}{rcl}
\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] & \equiv & \mathcal{T} \cup \mathsf{NBG}, \\
\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<\Omega^\omega}, \\
\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \cup TI_\in[\mathcal{L}^1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}, \\
\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \cup TI_\in[\mathcal{L}^1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}, \\
\mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0} \cup TI_\in[\mathcal{L}^1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0},
\end{array}$$

where $\mathcal{T}$ is any set of axioms with logical complexity essentially Σ^1_2, and the equivalence, $\equiv$, means that any sentence essentially Π^1_2 is either provable in both theories or in none of them. We can see that the proof-theoretic strength of these theories is essentially captured by the bounds Ω^ω and E_0 for iterated elementary comprehension.

To be able to explain what Ω^ω and E_0 are, we need to say a few words about the uniform representation, i.e. notation, of ordinals in set theory; because of the Cantor normal form for ordinals, we can write down any ordinal α by just using brackets $\langle\!\langle, \rangle\!\rangle$, epsilon numbers $\varepsilon_\beta \leq \alpha$, and $\emptyset$, where ε enumerates the class $\{\gamma \mid \gamma = \omega^\gamma\}$. The notation $\widehat{\alpha}$ representing the ordinal α is built according to the following recursion

$$\widehat{\alpha} \quad := \quad \begin{cases} \alpha & \alpha = \omega^\alpha \text{ or } \alpha = \emptyset, \\ \langle\!\langle \widehat{\gamma_0}, ..., \widehat{\gamma_r} \rangle\!\rangle & \alpha =_{\mathsf{CNF}} \omega^{\gamma_0} + ... + \omega^{\gamma_r} \neq \omega^\alpha. \end{cases}$$

E.g., if $\alpha = \omega^{\varepsilon_\beta + 1} + \omega$ then $\widehat{\alpha} = \langle\!\langle \langle\!\langle \varepsilon_\beta, \emptyset \rangle\!\rangle, \langle\!\langle \emptyset \rangle\!\rangle \rangle\!\rangle$. We can recursively define operations on these notations, reflecting addition, multiplication, exponentiation and the ordering relation on ordinals, such that $\widehat{\alpha} \mathbin{\hat{+}} \widehat{\beta}$, $\widehat{\alpha} \mathbin{\hat{\cdot}} \widehat{\beta}$, and $\widehat{\alpha}^{\widehat{\beta}}$, are the notations of $\alpha + \beta$, $\alpha \cdot \beta$, and α^β, respectively, and $\widehat{\alpha} \lhd \widehat{\beta}$ iff $\alpha \in \beta$. We further extend this notation system by an additional "virtual" epsilon number Ω, with $\varepsilon_\beta \lhd \Omega$ for any ε_β, and we allow notations $\widehat{\alpha}[\Omega]$, where $\widehat{\alpha}[\Omega]$ is just $\widehat{\alpha}$ with *all* occurences of the biggest epsilon number ε_β in $\widehat{\alpha}$ being replaced by Ω. E_0 is the collection of all notations $\widehat{\alpha}$ and $\widehat{\alpha}[\Omega]$. The arithmetical operations and the relation $\lhd$ are easily adapted to include all such notations, by treating Ω just like a real epsilon number bigger than any other epsilon number. In some way, $(E_0, \lhd)$ can be seen as the analogue of $(\varepsilon_0, \in)$, with the set of the natural numbers, i.e. the ordinal ω, replaced by the class of all ordinals, i.e. the notation Ω.

$(CA[\Pi^1_0])_{<\Omega^\omega}$ and $(CA[\Pi^1_0])_{<E_0}$ are the axioms of iterated comprehension along $\lhd$ up to Ω^n and Ω_n for any n, respectively, where $\Omega^{n+1} := \Omega^n \mathbin{\hat{\cdot}} \Omega$ and $\Omega_{n+1} := \widehat{\omega}^{\Omega_n}$ ($\Omega^0 := \widehat{1}$ and $\Omega_0 := \Omega + \widehat{1}$). The ordering $\lhd$ on the initial segments up to any Ω_n, is shown to be a *provable* well-founded class relation in NBG, for any n, hence the class hierarchies defined by iterated comprehension along $\lhd$ up to Ω_n are well-defined in NBG.

One direction of the proof-theoretic equivalences of the main results is shown with little effort, because the choice axioms already imply iterated comprehension, that is

$$\begin{array}{l} \mathsf{NBG} \cup DC[\Pi^1_0] \vdash (CA[\Pi^1_0])_{<\Omega^\omega}, \\ \mathsf{NBG} \cup AC[\Pi^1_0] \cup TI_\in[\mathcal{L}^1] \vdash (CA[\Pi^1_0])_{<E_0}. \end{array}$$

For the other direction we use asymmetric interpretations, similar to the asymmetric interpretations used, e.g., by Cantini [1] for subsystems of second order arithmetic, and by Jäger[7, 8, 9] and Jäger and Strahm [11] for theories of admissible sets, explicit mathematics and operational set theory. That is, for the other direction we reduce each of the systems with choice, $\mathcal{T}_{\mathsf{ch}}$ (i.e., $CA[\Pi^1_0] \cup AC[\Pi^1_0]$, $CA[\Pi^1_0] \cup DC_{\mathcal{O}n}[\Pi^1_0]$, and $CA[\Pi^1_0] \cup DC_{\mathcal{O}n}[\Pi^1_0] \cup TI_{\in}[\mathcal{L}^1]$), to the corresponding system with iterated comprehension, $\mathcal{T}_{\mathsf{it}}$ (i.e., NBG, $\mathsf{NBG} \cup (CA[\Pi^1_0])_{<\Omega^\omega}$, and $\mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}$), by an asymmetric interpretation, that is, we interpret any formula A into a hierarchy of classes U, such that (in a nutshell)

(1) After the asymmetric interpretation of the formula A into the hierarchy U, denoted by $A\langle \mathfrak{a}, \mathfrak{b} \rangle^U$, every quantified class variable of A ranges over some specific level of the class hierarchy, i.e. generally all existential quantifiers range over some level $(U)_{\mathfrak{b}}$ higher than the level of the universal quantifiers $(U)_{\mathfrak{a}}$ (hence the name *asymmetric* interpretation).

(2) All formulas *provable* in the system $\mathcal{T}_{\mathsf{ch}}$ we want to reduce, hold true after asymmetric interpretation into the class hierarchy U, i.e.,

$$\mathcal{T}_{\mathsf{ch}} \vdash A \quad \Rightarrow \quad \mathcal{T}_{\mathsf{it}} \vdash Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \rightarrow \exists \mathfrak{b} (A\langle \emptyset, \mathfrak{b} \rangle^U).$$

This statement corresponds to the statements proved in Lemma 109, where we can see that the situation is actually a bit more complicated. We write $Cl_{\mathcal{T}_{\mathsf{ch}}}[U]$ to emphasize the dependence of the class hierarchy U on the specific theory $\mathcal{T}_{\mathsf{ch}}$. Actually, the hierarchy U also depends on the formula A, i.e. on the free class variables in A, and on the length of the derivation of A in a particular proof system for $\mathcal{T}_{\mathsf{ch}}$ (see the proof of Lemma 110).

(3) For any formula up to some logical complexity, i.e. essentially Σ^1_1, the asymmetric interpretation of the formula reflects the truth of the original formula, that is,

$$A \text{ essentially } \Sigma^1_1 \quad \Rightarrow \quad \mathcal{T}_{\mathsf{it}} \vdash Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \rightarrow (A\langle \mathfrak{a}, \mathfrak{b} \rangle^U \rightarrow A).$$

(4) The class hierarchy exists in the system $\mathcal{T}_{\mathsf{it}}$ we want to reduce to, i.e.,

$$\mathcal{T}_{\mathsf{it}} \vdash \exists U \, Cl_{\mathcal{T}_{\mathsf{ch}}}[U].$$

By putting (2) to (4) together, we have that if a formula A is essentially Σ^1_1 and $\mathcal{T}_{\mathsf{ch}} \vdash A$ then we also have $\mathcal{T}_{\mathsf{it}} \vdash A$, and hence the proof-theoretic equivalences are fully established. Technically, the implication in (2) actually consists of two steps; the implication is proved by induction on the length of the cut-reduced proof of the formula A, that is, we need proofs of the formulas A where only cut formulas of complexity at most Σ^1_1 are used, i.e., we need partial cut-elimination. Hence we first show

$$\mathcal{T}_{\mathsf{ch}} \vdash A \quad \Rightarrow \quad \mathcal{T}_{\mathsf{ch}} \vdash_0 A,$$

and having this simpler (but usually much longer) proofs, we are able to get the desired implication by just showing

$$\mathcal{T}_{\mathsf{ch}} \vdash_0 A \quad \Rightarrow \quad \mathcal{T}_{\mathsf{it}} \vdash Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \to \exists \mathfrak{b}(A\langle \emptyset, \mathfrak{b}\rangle^U),$$

which is now proved by induction on the length of the cut-reduced proofs.

The reduction of the system $\mathcal{T}_{\mathsf{ch}}$ with full induction $TI_{\in}[\mathcal{L}^1]$ is actually a bit more involved than described above. The reduction is more complicated because of the intermediate step in (2), just described. The logical complexity of the formulas in $TI_{\in}[\mathcal{L}^1]$ is unbounded, hence the complexity of cut formulas gets far beyond Σ^1_1. We are dealing with this situation in a standard way, e.g., in the same way as in Jäger and Krähenbühl [10], that is, we change over to an *infinitary* proof system with $TI_{\in}[\mathcal{L}^1]$ already built in, such that instances from $TI_{\in}[\mathcal{L}^1]$ are derivable even without using any cut at all, and such that we can still prove partial cut elimination for this system, too. The proof system makes use of an infinitary rule for universal quantification over sets, that is, the rule applies to infinitely many premises (one for each set).

For the asymmetric interpretation including $TI_{\in}[\mathcal{L}^1]$, the infinitary proof system $Pr_{\mathcal{T}_{\mathsf{ch}}}$ for $\mathcal{T}_{\mathsf{ch}}$ is formalized within the system $\mathcal{T}_{\mathsf{it}}$. Here, we simply write $Pr_{\mathcal{T}_{\mathsf{ch}}}[\ulcorner A \urcorner]$ for the complex formula $\exists Z(Pr^{DC+}_{\Omega_{n+3}}[Z] \wedge \{\ulcorner A \urcorner\} \in (Z)_{\Omega \hat{+} \widehat{\omega}, \overline{n}})$, where $\{\ulcorner A \urcorner\} \in (Z)_{\Omega \hat{+} \widehat{\omega}, \overline{n}}$ means that $\ulcorner A \urcorner$ is derivable in at most $\Omega \hat{+} \widehat{\omega}$ steps with cut formulas of rank at most n, and where n, and hence $Pr_{\mathcal{T}_{\mathsf{ch}}}$, actually depend on the derivation of A in a particular (finitary) proof system for $\mathcal{T}_{\mathsf{ch}}$ (see the proof of Lemma 110). To accomplish the asymmetric interpretation, we use Gödelization of formulas and an appropriate definition of truth for the codes of formulas, $\ulcorner A \urcorner$, where truth is such that the class quantifiers in A range over some specified class universe U. Truth of the code of A is denoted by $\ulcorner A \urcorner[f, g]^U_\infty$ (where f and g take account of the free set and class variables in A, respectively). Property (2) and (3) now essentially become

(2')

$$\mathcal{T}_{\mathsf{it}} \vdash Pr_{\mathcal{T}_{\mathsf{ch}}}[\ulcorner A \urcorner] \to (Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \to \exists \mathfrak{b}(\ulcorner A \urcorner \langle \dot{\mathtt{C}}_\emptyset, \dot{\mathtt{C}}_\mathfrak{b}\rangle [f, g]^U_\infty)),$$

(3') If A is essentially Σ^1_1 then

$$\mathcal{T}_{\mathsf{it}} \vdash Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \to (\ulcorner A \urcorner \langle \dot{\mathtt{C}}_\mathfrak{a}, \dot{\mathtt{C}}_\mathfrak{b}\rangle [f, g]^U_\infty \to \ulcorner A \urcorner [f, g]^U_\infty).$$

Furthermore, the proof system $Pr_{\mathcal{T}_{\mathsf{ch}}}$ and the truth definition are such that the following properties hold

$$\mathcal{T}_{\mathsf{ch}} \vdash A \quad \Rightarrow \quad \mathcal{T}_{\mathsf{it}} \vdash Pr_{\mathcal{T}_{\mathsf{ch}}}[\ulcorner A \urcorner],$$

$$A \text{ essentially } \Sigma^1_1 \quad \Rightarrow \quad \mathcal{T}_{\mathsf{it}} \vdash \sharp_A[f, g, U] \to (\ulcorner A \urcorner [f, g]^U_\infty \to A),$$

where the formula $\sharp_A$ stands for the proper assignment of sets and classes to the free variables in A (through f and g). Hence, together with (4) and because of $\mathcal{T}_{\mathsf{it}} \vdash Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \rightarrow \exists f \exists g(\sharp_A[f, g, U])$, we have that if a formula A is essentially Σ^1_1 and $\mathcal{T}_{\mathsf{ch}} \vdash A$ then we also have $\mathcal{T}_{\mathsf{it}} \vdash A$, and finally we have established the proof-theoretic equivalence in the more complicated cases, too.

1. Logic for Theories of Sets and Classes

In this section we define the language and the logic for theories of sets and classes, as we use it throughout this text. We work in classical logic with two sorts of variables $x, y, z...$, and $X, Y, Z...$, for sets and classes, respectively. We follow the style of Tait, that is, the logic is defined analogous to Tait-language and Tait-calculus, e.g. as the language and logic for second order arithmetic is defined by Pohlers [15]. Our language of set theory is very simple, because it only has the element relation symbol $\in$, and no other relation or function symbols. Actually, for technical reasons, there are two relation symbols $\in^0$ and $\in^1$, i.e. $x{\in^0}y$ and $x{\in^1}Y$, because we syntactically distinguish between sets and classes. Equality will be defined in terms of $\in$, hence we have logic without equality. See, e.g., Mendelson [14] for axiomatic set theory in pure first-order logic with just one sort of variables and one single relation symbol, and consider Appendix B for the exact relationship between these different formalizations of axiomatic set theory.

The very heart of this section is the definition of the notion of formal proof in the form of a provability relation, and some theorems about important structural properties thereof, i.e. like partial cut elimination.

We write $\mathbb{N}$ for the collection of all (standard) *natural numbers* $0, 1, 2, ...$ and we use the letters i, j, k, l, m, n (with subscripts) to denote natural numbers.

Definition 1. (Language $\mathcal{L}^0$, $\mathcal{L}^1$)

The *language* $\mathcal{L}^1$ of *Von Neumann–Bernays–Gödel* set theory consists of the following

(1) The *logical symbols* of $\mathcal{L}^1$ are

a) the *free set variables* $\mathtt{v}_i$, for all $i \in \mathbb{N}$,

b) the *bound set variables* $\mathtt{u}_i$, for all $i \in \mathbb{N}$,

c) the *free class variables* $\mathtt{V}_i$ for all $i \in \mathbb{N}$,

d) the *bound class variables* $\mathtt{U}_i$ for all $i \in \mathbb{N}$,

e) the *propositional connectives* $\wedge$, $\vee$, $\sim$,

f) the *quantifiers* $\forall$, $\exists$,

g) and the *auxiliary symbols* (,).

We use the letters f, g, h, u, v, w, x, y, z (with subscripts) to denote *free or bound* set variables, and we use the letters F, G, H, U, V, W, X, Y, Z (with subscripts) to denote *free or bound* class variables. The mapping $\sharp$ on variables is defined such that $\sharp\mathtt{v}_i := \sharp\mathtt{V}_i := \sharp\mathtt{u}_i := \sharp\mathtt{U}_i := i$ for $i \in \mathbb{N}$.

(2) The *non logical symbols* of $\mathcal{L}^1$ are the *element relation* symbols $\in^0$ and $\in^1$.

(3) The *atomic $\mathcal{L}^1$-formulas* are $\mathtt{v}_i{\in^0}\mathtt{v}_j$, $\sim\mathtt{v}_i{\in^0}\mathtt{v}_j$, $\mathtt{v}_i{\in^1}\mathtt{V}_j$, $\sim\mathtt{v}_i{\in^1}\mathtt{V}_j$, for all $i, j \in \mathbb{N}$. We write $x \in y$, $x \notin y$, $x \in Y$, $x \notin Y$, for $x{\in^0}y$, $\sim x{\in^0}y$, $x{\in^1}Y$, $\sim x{\in^1}Y$, respectively.

(4) The *$\mathcal{L}^1$-formulas* are inductively defined, such that,

a) all atomic $\mathcal{L}^1$-formulas are $\mathcal{L}^1$-formulas,

b) if A and B are $\mathcal{L}^1$-formulas then $(A \wedge B)$ and $(A \vee B)$ are $\mathcal{L}^1$-formulas,

c) and if A is a $\mathcal{L}^1$-formula, and u, V are free variables, and the bound variables x, Y do not occur in A (and hence x, Y are free for u, V in A), then $\forall x A[x/u]$, $\exists x A[x/u]$, $\forall Y A[Y/V]$, $\exists Y A[Y/V]$ are $\mathcal{L}^1$-formulas. $A[x/u]$ and $A[Y/V]$ stand for the expressions that are obtained from A by replacing all occurrences of u and V by x and Y, respectively.

We use the letters A, B, C, D, E (with subscripts), to denote $\mathcal{L}^1$-formulas. The mapping $\sharp$ on $\mathcal{L}^1$-formulas is some fixed injective mapping to $\mathbb{N}$, that is, if $\sharp A = \sharp B$ then $A = B$.

The *language* $\mathcal{L}^0$ of *Zermelo–Fraenkel* set theory is defined analogous to $\mathcal{L}^1$, by *omitting* the relation symbol $\in^1$, and all class variables $\mathtt{V}_i, \mathtt{U}_j$, for all $i, j \in \mathbb{N}$, in the definition of $\mathcal{L}^1$.

If $\underline{x}$ and $\underline{u}$ are sequences of variables $x_0, ..., x_k$, and $u_0, ..., u_k$, respectively, and if the variables in $\underline{u}$ are pairwise distinct, then $A[\underline{x}/\underline{u}]$ stands for the expression that is obtained from A by replacing all occurrences of u_i by x_i, simultaneously for all $i \leq k$. We may write $A[\underline{x}]$ instead of $A[\underline{x}/\underline{u}]$ whenever the variables in $\underline{u}$ are unimportant or $\underline{u}$ is determined by context. Analogously $A[\underline{X}]$ for $A[\underline{X}/\underline{U}]$, and the same for sequences of variables of mixed type, e.g. $A[x, y, Z]$ or $A[X, y, z]$.

For any language $\mathcal{L}$, we write $A \in \mathcal{L}$, if A is a $\mathcal{L}$-formula, and we write $\mathcal{C} \subseteq \mathcal{L}$, if $\mathcal{C}$ is a set of $\mathcal{L}$-formulas. We write $x \in A$ if the *free or bound* variable x occurs in A, otherwise we write $x \notin A$. Analogously $X \in A$ or $X \notin A$. Formulas with *no free variables* are called *sentences* or *closed* formulas. A set of formulas $\mathcal{C} \subseteq \mathcal{L}^1$ is *closed under substitution* if for all $A \in \mathcal{C}$ and all free variables u, v, U, V, we have that $A[v/u] \in \mathcal{C}$ and $A[V/U] \in \mathcal{C}$.

Because of the simplicity of the language $\mathcal{L}^1$ we need a lot of syntactic abbreviations. That is, we define short notations for (sometimes very long) formulas, and we extensively use such expressions.

Definition 2. (Abbreviations)

For $A \in \mathcal{L}^1$ we define $\neg A \in \mathcal{L}^1$ such that

$$\neg A \quad := \quad \begin{cases} \sim A & A \text{ atomic}, A \neq \sim B, \\ B & A = \sim B, \\ (\neg B \vee \neg C) & A = (B \wedge C), \\ (\neg B \wedge \neg C) & A = (B \vee C), \\ \forall x \neg B[x] & A = \exists x B[x], \\ \exists x \neg B[x] & A = \forall x B[x], \\ \forall X \neg B[X] & A = \exists X B[X], \\ \exists X \neg B[X] & A = \forall X B[X]. \end{cases}$$

For $A, B, C \in \mathcal{L}^1$ (with $x, y \notin C$, $x \neq y$) and free variables u, v we define abbreviations for the following formulas,

$$\begin{aligned} (A \to B) \quad &:= \quad (\neg A \vee B), \\ (A \leftrightarrow B) \quad &:= \quad ((A \to B) \wedge (B \to A)), \\ (\forall x \in u)C[x] \quad &:= \quad \forall x(x \in u \to C[x]), \\ (\exists x \in u)C[x] \quad &:= \quad \exists x(x \in u \wedge C[x]), \\ u \subseteq v \quad &:= \quad (\forall x \in u)x \in v, \\ (u = v) \quad &:= \quad (u \subseteq v \wedge v \subseteq u), \\ \exists! x C[x] \quad &:= \quad \exists x C[x] \wedge \forall x \forall y(C[x] \wedge C[y] \to x = y), \\ \bot \quad &:= \quad \forall x(x \in x \wedge x \notin x), \\ \top \quad &:= \quad \neg\bot. \end{aligned}$$

$(\forall x \in U)C[x]$, $(\exists x \in U)C[x]$, and $x \subseteq Y$, $X \subseteq Y$, $X \subseteq y$, and $x = Y$, $X = Y$, $X = y$ are defined analogously.

If $\mathcal{C} \subseteq \mathcal{L}^1$ then we write $\mathcal{C}^{\to}$ for the set $\{B \mid \exists A\, (A \to B) \in \mathcal{C}\}$.

The logical complexity of formulas is an important measure, because a lot of theorems in logic, e.g. the main Theorem 113 of this thesis, only hold for a specific level of formula complexity. In the following we define classes of formulas that correspond to the number of the nesting of alternating class quantifiers.

Definition 3. (Σ^1_+, Π^1_+, Σ^1, Π^1)

For sets of formulas $\mathcal{C} \subseteq \mathcal{L}^1$ we define

(1) $\Sigma^1_+(\mathcal{C}) := \mathcal{C} \cup \{\exists X A[X] \in \mathcal{L}^1 \mid A \in \mathcal{C}\}$,
$\Pi^1_+(\mathcal{C}) := \mathcal{C} \cup \{\forall X A[X] \in \mathcal{L}^1 \mid A \in \mathcal{C}\}$.

(2) $\Sigma^1(\mathcal{C})$ and $\Pi^1(\mathcal{C})$ are inductively defined, such that both contain *all formulas in* $\mathcal{C}$, and if $A, B \in \Sigma^1(\mathcal{C})$ then the $\mathcal{L}^1$-formulas $\forall x A[x]$, $\exists x A[x]$, $\exists X A[X]$, $(A \wedge B)$, $(A \vee B)$ are in $\Sigma^1(\mathcal{C})$, and if $A, B \in \Pi^1(\mathcal{C})$ then the $\mathcal{L}^1$-formulas $\forall x A[x]$, $\exists x A[x]$, $\forall X A[X]$, $(A \wedge B)$, $(A \vee B)$ are in $\Pi^1(\mathcal{C})$.

(3) $\neg\mathcal{C} := \{\neg A \mid A \in \mathcal{C}\}$.

Lemma 4.

For any $A \in \mathcal{L}^1$ and $\mathcal{C} \subseteq \mathcal{L}^1$ we have that

(1) $\neg\neg A = A$, i.e. $\neg\neg\mathcal{C} = \mathcal{C}$.

(2) $\neg\Sigma^1_+(\mathcal{C}) = \Pi^1_+(\neg\mathcal{C})$.

(3) $\neg\Sigma^1(\mathcal{C}) = \Pi^1(\neg\mathcal{C})$.

Proof.

(1) By induction on the structure of the formula A, we show $\neg\neg A = A$.

(2)

$$\begin{aligned} \neg\Sigma^1_+(\mathcal{C}) &= \neg\mathcal{C} \cup \{\forall X \neg A[X] \in \mathcal{L}^1 \mid A \in \mathcal{C}\} \\ &= \neg\mathcal{C} \cup \{\forall X B[X] \in \mathcal{L}^1 \mid B \in \neg\mathcal{C}\} = \Pi^1_+(\neg\mathcal{C}). \end{aligned}$$

(3) We show $\neg\Sigma^1(\mathcal{C}) \subseteq \Pi^1(\neg\mathcal{C})$, by induction on the structure of the formulas in $\Sigma^1(\mathcal{C})$. E.g. if $B = \exists X A[X] \in \mathcal{L}^1$ with $A \in \Sigma^1(\mathcal{C})$ then $\neg B = \forall X \neg A[X]$ with $\neg A \in \neg\Sigma^1(\mathcal{C})$, and $\neg A \in \Pi^1(\neg\mathcal{C})$ by i.h., hence $\neg B \in \Pi^1(\neg\mathcal{C})$. We prove $\Pi^1(\neg\mathcal{C}) \subseteq \neg\Sigma^1(\mathcal{C})$, by induction on the structure of the formulas in $\Pi^1(\neg\mathcal{C})$. E.g. if $B = \forall X A[X] \in \Pi^1(\neg\mathcal{C})$ with $A \in \Pi^1(\neg\mathcal{C})$, hence $A \in \neg\Sigma^1(\mathcal{C})$ by i.h., then $\neg\exists X \neg A[X] = \forall X \neg\neg A[X] = B \in \neg\Sigma^1(\mathcal{C})$ by Part 1. □

Definition 5. (Formula Classes Σ^1_n, $\Pi^1_n \subseteq \mathcal{L}^1$)

(1) $\Sigma^1_0 := \Pi^1_0$, where Π^1_0 is inductively defined, such that Π^1_0 contains *all atomic* $\mathcal{L}^1$*-formulas*, and if $A, B \in \Pi^1_0$ then the $\mathcal{L}^1$-formulas $\forall x A[x]$, $\exists x A[x]$, $(A \wedge B)$, $(A \vee B)$ are in Π^1_0. The formulas in Π^1_0 are called *elementary* formulas.

(2) $\Sigma^1_{n+1} := \Sigma^1_+(\Pi^1_n)$,
$\Pi^1_{n+1} := \Pi^1_+(\Sigma^1_n)$.

Lemma 6.

(1) $\neg\Sigma^1_n = \Pi^1_n$.

(2) $\Sigma^1_n \cup \Pi^1_n = \Sigma^1_{n+1} \cap \Pi^1_{n+1}$.

Proof.

(1) We prove $\neg\Pi^1_0 \subseteq \Pi^1_0$, by induction on the structure of the formulas in Π^1_0, hence $\Pi^1_0 = \neg\neg\Pi^1_0 \subseteq \neg\Pi^1_0 = \neg\Sigma^1_0 \subseteq \Pi^1_0$. By induction on n and by Lemma 4 we get $\neg\Sigma^1_{n+1} = \neg\Sigma^1_+(\Pi^1_n) = \Pi^1_+(\neg\Pi^1_n) = \Pi^1_+(\Sigma^1_n) = \Pi^1_{n+1}$.

(2) By definition we have $\Sigma^1_0 \cup \Pi^1_0 \subseteq \Sigma^1_1 \cap \Pi^1_1$. We prove $\Sigma^1_n \cup \Pi^1_n \subseteq \Sigma^1_{n+1} \cap \Pi^1_{n+1}$ by induction on n. By definition we have $\Pi^1_n \subseteq \Sigma^1_{n+1}$, and by i.h. we get $\Sigma^1_{n-1} \subseteq \Sigma^1_n$, hence $\Pi^1_n = \Pi^1_+(\Sigma^1_{n-1}) \subseteq \Pi^1_+(\Sigma^1_n) = \Pi^1_{n+1}$, analogously $\Sigma^1_n \subseteq \Sigma^1_{n+1} \cap \Pi^1_{n+1}$. To show $\Sigma^1_{n+1} \cap \Pi^1_{n+1} \subseteq \Sigma^1_n \cup \Pi^1_n$, we assume $A \in \Sigma^1_{n+1} \cap \Pi^1_{n+1}$ and $A \notin \Sigma^1_n$, and we show $A \in \Pi^1_n$. Because of $A \in \Pi^1_{n+1}$ and $A \notin \Sigma^1_n$, we have $A = \forall X B[X] \in \mathcal{L}^1$ for some $B \in \Sigma^1_n$, and because of $A \in \Sigma^1_{n+1}$ and $A = \forall X B[X] \neq \exists Y C[Y] \in \mathcal{L}^1$ for all $C \in \Pi^1_n$, we get $A \in \Pi^1_n$. □

The rank $\mathsf{rk}_\mathcal{C}(A)$ with respect to $\mathcal{C} \subseteq \mathcal{L}^1$ of a formula A is defined such that $\mathsf{rk}_\mathcal{C}(A) = 0$ iff $A \in \mathcal{C} \cup \neg\mathcal{C}$, and $\mathsf{rk}_\mathcal{C}(A) = \mathsf{rk}_\mathcal{C}(\neg A)$ for every A.

Definition 7. (Formula Rank $\mathsf{rk}_\mathcal{C}$)

For sets of formulas $\mathcal{C} \subseteq \mathcal{L}^1$ and formulas $A \in \mathcal{L}^1$ we define

$$\mathsf{rk}_\mathcal{C}(A) \quad := \quad 0 \text{ if } A \in \mathcal{C} \cup \neg\mathcal{C}, \text{ and otherwise}$$

$$\mathsf{rk}_\mathcal{C}(A) \quad := \quad \begin{cases} 1 & A \text{ atomic,} \\ \max\{\mathsf{rk}_\mathcal{C}(B), \mathsf{rk}_\mathcal{C}(C)\} + 1 & A = (B \wedge C), (B \vee C), \\ \mathsf{rk}_\mathcal{C}(B) + 1 & A = \begin{cases} \exists x B[x], \forall x B[x], \\ \exists X B[X], \forall X B[X]. \end{cases} \end{cases}$$

We write rk for $\mathsf{rk}_\emptyset$.

We observe that by definition we have $\mathsf{rk}_\mathcal{C} = \mathsf{rk}_{\mathcal{C} \cup \neg\mathcal{C}}$.

Lemma 8.

(1) If $\mathcal{C} \subseteq \mathcal{D}$ then $\mathsf{rk}_\mathcal{D}(A) \leq \mathsf{rk}_\mathcal{C}(A)$.

(2) $\mathsf{rk}_\mathcal{C}(A) = \mathsf{rk}_\mathcal{C}(\neg A)$.

Proof. By induction on the structure of the formula A. □

For the sake of completeness, and because of some logical considerations in Appendix A and B, we give a formal definition of the semantics of $\mathcal{L}^1$. As we work in proof theory, which by its nature mostly deals with pure syntax (except for completeness results), we are not going to use semantics any further in this thesis. Clearly, we inherently use semantics whenever we give just an *informal* proof for some statement within a specific set theory. But, of course, any of this informal proofs could also be replaced by a formal proof, as we know by invoking the completeness theorem, cf. Theorem 13.

Definition 9. (Semantics)

(1) A $\mathcal{L}^1$-*structure* is a tuple $\mathcal{M} := (|\mathcal{M}|, \|\mathcal{M}\|, \in^0_{\mathcal{M}}, \in^1_{\mathcal{M}})$ with *non-empty* domains $|\mathcal{M}|$, $\|\mathcal{M}\|$, and relations $\in^0_{\mathcal{M}}, \in^1_{\mathcal{M}}$ such that $\in^0_{\mathcal{M}} \subseteq |\mathcal{M}| \times |\mathcal{M}|$, and $\in^1_{\mathcal{M}} \subseteq |\mathcal{M}| \times \|\mathcal{M}\|$.

(2) A $\mathcal{L}^1$-*valuation* is a tuple $\mathcal{V} := (\langle\mathcal{V}\rangle, f, g)$ such that $\langle\mathcal{V}\rangle$ is a $\mathcal{L}^1$-structure, and $f : \mathbb{N} \to |\langle\mathcal{V}\rangle|$, and $g : \mathbb{N} \to \|\langle\mathcal{V}\rangle\|$.

(3) Given a function f with domain $\mathbb{N}$, the mapping $f[a/n]$ is such that $f[a/n](n) := a$ and $f[a/n](i) := f(i)$ for $i \neq n$. For $\mathcal{L}^1$-valuations $\mathcal{V} = (\mathcal{M}, f, g)$ and free variables x, Y we define $\mathcal{V}[a/x] := (\mathcal{M}, f[a/\sharp x], g)$ and $\mathcal{V}[b/Y] := (\mathcal{M}, f, g[b/\sharp Y])$. $\mathcal{V}[\,]$ is $\mathcal{V}$. If $\underline{a}$ and $\underline{n}$ are the sequences $a_0, ..., a_k$ and $n_0, ..., n_k$, resp., then $f[\underline{a}/\underline{n}]$ stands for $f[a_0/n_0]...[a_k/n_k]$, and analogously $\mathcal{V}[\underline{a}/\underline{x}]$, $\mathcal{V}[\underline{b}/\underline{Y}]$.

(4) $\mathcal{V} \models A$ holds iff $\mathcal{V} = (\mathcal{M}, f, g)$ is a $\mathcal{L}^1$-valuation, $A \in \mathcal{L}^1$ and one of the following holds:

$$\begin{array}{lll}
A = x \in y & \text{and} & f(\sharp x) \in^0_{\mathcal{M}} f(\sharp y), \\
A = x \in Y & \text{and} & f(\sharp x) \in^1_{\mathcal{M}} g(\sharp Y), \\
A = {\sim} B & \text{and} & \mathcal{V} \not\models B, (\text{not } \mathcal{V} \models B) \\
A = B \wedge C & \text{and} & (\mathcal{V} \models B \text{ and } \mathcal{V} \models C), \\
A = B \vee C & \text{and} & (\mathcal{V} \models B \text{ or } \mathcal{V} \models C), \\
A = \forall x B[x/u] & \text{and} & (\forall a \in |\mathcal{M}|)\ \mathcal{V}[a/u] \models B, \\
A = \exists x B[x/u] & \text{and} & (\exists a \in |\mathcal{M}|)\ \mathcal{V}[a/u] \models B, \\
A = \forall X B[X/U] & \text{and} & (\forall b \in \|\mathcal{M}\|)\ \mathcal{V}[b/U] \models B, \\
A = \exists X B[X/U] & \text{and} & (\exists b \in \|\mathcal{M}\|)\ \mathcal{V}[b/U] \models B.
\end{array}$$

$\mathcal{V} \models \mathcal{F}$ holds iff $\mathcal{F} \subseteq \mathcal{L}^1$ and $\mathcal{V} \models A$ for all $A \in \mathcal{F}$. $\mathcal{T} \models \mathcal{F}$ holds iff $\mathcal{T}, \mathcal{F} \subseteq \mathcal{L}^1$ and $(\mathcal{V} \models \mathcal{T} \Rightarrow \mathcal{V} \models \mathcal{F})$ for all $\mathcal{L}^1$-valuations $\mathcal{V}$. We write $\mathcal{T} \models A$ for $\mathcal{T} \models \{A\}$, and $A \models \mathcal{F}$ for $\{A\} \models \mathcal{F}$, that is, $A \models B$ for $\{A\} \models \{B\}$. $\models \mathcal{F}$ stands for $\emptyset \models \mathcal{F}$.

(5) $\mathcal{M} \models A[\underline{a}/\underline{x}][\underline{b}/\underline{Y}]$ holds iff $\mathcal{M}$ is a $\mathcal{L}^1$-structure and all $\mathcal{L}^1$-valuations $\mathcal{V}$ with $\langle\mathcal{V}\rangle = \mathcal{M}$ are such that $\mathcal{V}[\underline{a}/\underline{x}][\underline{b}/\underline{Y}] \models A$. $\mathcal{M} \models A$ is $\mathcal{M} \models A[\,]$. $\mathcal{M} \models \mathcal{T}$ holds iff $\mathcal{T} \subseteq \mathcal{L}^1$ and $\mathcal{M} \models A$ for all $A \in \mathcal{T}$.

To simplify notation we use the same symbol, $\models$, for distinct relations in (4) and (5).

The following lemma shows that logical implication, $A \to B$, exactly corresponds to logical consequence, $A \models B$. The situation is different for provability, cf. Definition 12, for which we have that $\vdash A \to B$ implies $A \vdash B$, but the other direction only holds for closed formulas A.

Lemma 10.

For $A, B \in \mathcal{L}^1$ we have $\models A \to B$ iff $A \models B$.

Proof. By the definition of $\models$. □

The only part of the logic for $\mathcal{L}^1$ still missing, is the notion of formal proof. By formal proofs we derive finite sets of formulas Γ, and the existence of such a derivation means that the disjunction over all formulas in Γ holds (with respect to the specific axioms).

Definition 11.

Finite (possibly empty) sets of formulas are denoted by the greek letters Γ, Δ, Θ, Φ (with subscripts). If $\Gamma = \{A_0, ..., A_n\}$ and $\Delta = \{B_0, ..., B_m\}$ then Γ, C, Δ stands for $\{A_0, ..., A_n, C, B_0, ..., B_m\}$. We write $u \notin \Gamma$ for the variable u, if $u \notin A_i$ for all $i \leq n$, and otherwise $u \in \Gamma$ (analogously $U \notin \Gamma$ or $U \in \Gamma$), further

$$
\begin{array}{rcll}
\neg\Gamma & := & \{\neg A_0, ..., \neg A_n\}, & \\
\Gamma^{\vee} & := & ((\bot \vee A_0) ... \vee A_n), & (\sharp A_i < \sharp A_j \text{ for } i < j) \\
\mathsf{rk}(\Gamma) & := & \max\{\mathsf{rk}(A) \mid A \in \Gamma\}, & (\max\{\} = 0) \\
\Gamma[u/v] & := & \{A_0[u/v], ..., A_n[u/v]\}. & (\text{analogously } \Gamma[U/V])
\end{array}
$$

The inference rules for formal proofs consist of the common rules for classical logic, and the rules for the axioms $\mathcal{T} \subseteq \mathcal{L}^1$, and additional inference rules specified by some set $\mathcal{R} \subseteq \mathcal{L}^1$. The rules can be depicted as follows

$$\Gamma, A, \neg A \quad \text{with } A \text{ atomic},$$

$$\frac{\Gamma, A}{\Gamma, A \vee B}, \qquad \frac{\Gamma, B}{\Gamma, A \vee B},$$

$$\frac{\Gamma, A \qquad \Gamma, B}{\Gamma, A \wedge B},$$

$$\frac{\Gamma, A[v/u]}{\Gamma, \exists x A[x/u]}, \qquad \frac{\Gamma, A}{\Gamma, \forall x A[x/u]} \quad \text{with } u \notin \Gamma,$$

$$\frac{\Gamma, A[V/U]}{\Gamma, \exists X A[X/U]}, \qquad \frac{\Gamma, A}{\Gamma, \forall X A[X/U]} \quad \text{with } U \notin \Gamma,$$

$$\frac{\Gamma, A \qquad \Gamma, \neg A}{\Gamma},$$

$$\Gamma, A \quad \text{with } A \in \mathcal{T},$$

$$\frac{\Gamma, B}{\Gamma, A} \quad \text{with } (B \to A) \in \mathcal{R}.$$

In the following we define the provability relation $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n,j}_{\mathcal{C},l} \Gamma$, which captures provability by formal proofs. In addition to the axioms $\mathcal{T}$ and the additional inference rules $\mathcal{R}$, the relation also has control parameters n, j, $\mathcal{C}$, and l, such that $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n,j}_{\mathcal{C},l} \Gamma$ essentially means that

(1) there is a derivation of the finite set of formulas Γ, which possibly uses axioms in $\mathcal{T}$ and additional inference rules from $\mathcal{R}$,

(2) this derivation takes at most n steps (by definition we will have $n > 0$),

(3) any formula A occurring in this derivation has rank $\mathsf{rk}(A)$ at most j,

(4) the rank of any cut in this derivation is at most l with respect to $\mathcal{C}$, that is, any cut-formula A is such that $\mathsf{rk}_{\mathcal{C}}(A) \leq l$. If $l = 0$ then only cut-formulas in $\mathcal{C} \cup \neg\mathcal{C}$ are used.

Definition 12. (Formal Proof)

$\mathcal{T}\ [\mathcal{R}]\!\vdash^{n,j}_{\mathcal{C},l} \Phi$ holds iff $\mathcal{T}, \mathcal{R}, \mathcal{C}, \Phi \subseteq \mathcal{L}^1$ (Φ finite or empty), and $j, l, n \in \mathbb{N}$, and there are $A, B \in \mathcal{L}^1$, and $i, k, m \in \mathbb{N}$ with $m < n$, and $k \leq l$, and $\max\{i, \mathsf{rk}(\Phi)\} \leq j$, and there is $\Gamma \subseteq \mathcal{L}^1$, and free variables $u, U \notin \Phi$ and v, w, V, W, such that one of the following cases holds

$$\begin{array}{lll}
\Phi = \Gamma, A, \neg A & \text{and} & A \text{ is atomic,} \\
\Phi = \Gamma, A & \text{and} & A \in \mathcal{T}, \\
\Phi = \Gamma, A & \text{and} & \mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, B \text{ and } (B \to A) \in \mathcal{R}, \\
\Phi = \Gamma, A \vee B & \text{and} & (\mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, A \text{ or } \mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, B), \\
\Phi = \Gamma, A \wedge B & \text{and} & \mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, A \text{ and } \mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, B, \\
\Phi = \Gamma, \exists x A[x/v] & \text{and} & \mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, A[w/v], \\
\Phi = \Gamma, \exists X A[X/V] & \text{and} & \mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, A[W/V], \\
\Phi = \Gamma, \forall x A[x/u] & \text{and} & \mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, A, \\
\Phi = \Gamma, \forall X A[X/U] & \text{and} & \mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, A, \\
\multicolumn{3}{l}{\mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Phi, A \text{ and } \mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Phi, \neg A \text{ and } \mathsf{rk}_{\mathcal{C}}(A) \leq l.}
\end{array}$$

If $\mathcal{T}, [\mathcal{R}]$ or $\mathcal{C}$ is omitted in the notation $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Phi$ then $\mathcal{T} = \emptyset$, $\mathcal{R} = \emptyset$, $\mathcal{C} = \emptyset$, respectively. If m, i, k is omitted then this means that *there is some* unspecified m, i, k.

If $\Phi = \{A\}$ then we may write $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} A$ for $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Phi$. And if $\mathcal{F} \subseteq \mathcal{L}^1$ and $\mathcal{T}\ [\mathcal{R}]\vdash A$ holds for every $A \in \mathcal{F}$, then we may ambiguously write $\mathcal{T}\ [\mathcal{R}]\vdash \mathcal{F}$, but the meaning will always be clear from context, e.g., because $\mathcal{F}$ is an infinite set of formulas.

Clearly we want the provability relation to be adequate to the semantics of $\mathcal{L}^1$. The following theorem states soundness and completeness of provability.

Theorem 13. (Adequacy)

If $\mathcal{T} \subseteq \mathcal{L}^1$ is a set of *sentences* then $\mathcal{T} \vdash A$ iff $\mathcal{T} \models A$.

Proof. See Theorem 134 in the Appendix. $\square$

For the manipulation of proofs we have the following two lemmas about structural properties. There are a few unusual properties of our peculiar provability relation, i.e. parts 3, 4, and 8, of the next lemma, but all other properties are very common, like, weakening, substitution, the deduction theorem, i.e. compactness, and inversion.

Lemma 14. (Structural Properties)

(1) If $\mathcal{T}_0 \subseteq \mathcal{T}_1$, $\mathcal{R}_0 \subseteq \mathcal{R}_1$, $\mathcal{C}_0 \subseteq \mathcal{C}_1$, $m \leq n$, $\max\{i, \mathsf{rk}(\Delta)\} \leq j$, $k \leq l$, and $\mathcal{T}_0\ [\mathcal{R}_0]\vdash^{m,i}_{\mathcal{C}_0,k} \Gamma$, then $\mathcal{T}_1\ [\mathcal{R}_1]\vdash^{n,j}_{\mathcal{C}_1,l} \Delta, \Gamma$.

(2) If $\mathcal{T}, \mathcal{R}, \mathcal{C} \subseteq \mathcal{L}^1$ are *closed* under substitution and $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C},l} \Gamma$ then $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C},l} \Gamma[y/z]$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C},l} \Gamma[Y/Z]$ for all free variables y, z, Y, Z.

(3) $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C}\cup\neg\mathcal{C},l} \Gamma$ iff $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C},l} \Gamma$.

(4) If $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C},l} \Gamma$ then $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{j} \Gamma$.

(5) $\vdash_0 \neg A, A$.

(6) If $\mathcal{F}$ is a set of *sentences* and $\mathcal{F} \cup \mathcal{T}\ [\mathcal{R}]\vdash^{n}_{\mathcal{C},l} \Gamma$ then there is a finite set $\Delta \subseteq \mathcal{F}$ such that $\mathcal{T}\ [\mathcal{R}]\vdash_{\mathcal{C},l} \neg\Delta, \Gamma$.

(7) If $\mathcal{T} \cup \{\forall x A[x]\} \vdash^{n}_{\mathcal{C},l} \Gamma$ then $\mathcal{T} \cup \{A\} \vdash^{n+1}_{\mathcal{C},l} \Gamma$, and if $\mathcal{T} \cup \{A\} \vdash^{n} \Gamma$ and $\forall x A[x] \in \mathcal{L}^1$ then $\mathcal{T} \cup \{\forall x A[x]\} \vdash \Gamma$ (same for $\forall X A[X] \in \mathcal{L}^1$).

(8) If $\mathcal{T}\ [\mathcal{Q} \cup \mathcal{R}]\vdash^{n} \Gamma$ then $(\mathcal{T} \cup \mathcal{Q})\ [\mathcal{R}]\vdash \Gamma$, and if $(\mathcal{T} \cup \mathcal{Q})\ [\mathcal{R}]\vdash^{n}_{\mathcal{C},l} \Gamma$ and $\mathcal{Q} \subseteq \{(A \to B) \mid A, B \in \mathcal{L}^1\}$ then $\mathcal{T}\ [\mathcal{Q} \cup \mathcal{R}]\vdash_{\mathcal{C},l} \Gamma$.

Proof.

(1) By induction on m, considering all cases in Definition 12.

If $\mathcal{T}_0\ [\mathcal{R}_0]\vdash^{m_0,i_0}_{\mathcal{C}_0,k_0} \Gamma, A$, and $\mathcal{T}_0\ [\mathcal{R}_0]\vdash^{m_0,i_0}_{\mathcal{C}_0,k_0} \Gamma, \neg A$ where $m_0 < m$, $k_0 \leq k$, and $\mathsf{rk}_{\mathcal{C}_0}(A) \leq k$, and $\max\{i_0, \mathsf{rk}(\Gamma)\} \leq i$, then by i.h. for $j_0 = \max\{i_0, \mathsf{rk}(\Delta)\}$ we have that

$$\mathcal{T}_1\ [\mathcal{R}_1]\vdash^{m_0,j_0}_{\mathcal{C}_1,k_0} \Delta, \Gamma, A \text{ and } \mathcal{T}_1\ [\mathcal{R}_1]\vdash^{m_0,j_0}_{\mathcal{C}_1,k_0} \Delta, \Gamma, \neg A.$$

Because of $m_0 < m \leq n$, and $k_0 \leq k \leq l$, and

$$\max\{j_0, \mathsf{rk}(\Delta, \Gamma)\} = \max\{i_0, \mathsf{rk}(\Gamma), \mathsf{rk}(\Delta)\} \leq \max\{i, \mathsf{rk}(\Delta)\} \leq j,$$

and $\mathsf{rk}_{\mathcal{C}_1}(A) \leq \mathsf{rk}_{\mathcal{C}_0}(A) \leq k \leq l$ by Lemma 8, we get $\mathcal{T}_1\ [\mathcal{R}_1]\vdash^{n,j}_{\mathcal{C}_1,l} \Delta, \Gamma$.

Similarly for the other cases.

(2) By induction on n, considering all cases in Definition 12.

If $\Gamma = \Gamma_0, \exists x A[x/v]$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma_0, A[w/v]$ with $m < n$, and $k \leq l$, and $\max\{i, \mathsf{rk}(\Gamma)\} \leq j$, then $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma_0[y/z], A[w/v][y/z]$ by i.h., and we may assume $v \notin \{y, z, w\}$, because for any free variable $u \notin A$ and $B := A[u/v]$ we have that $\exists x A[x/v] = \exists x B[x/u]$ and $A[w/v] = B[w/u]$. We further have $\exists x A[x/v][y/z] = \exists x A[y/z][x/v]$ because of $y \neq v \neq z$, and

$$A[w/v][y/z] = \begin{cases} A[y/z][y/v] & w = z, \\ A[y/z][w/v] & w \neq z, \end{cases}$$

i.e. there is some $u \in \{y, w\}$ such that $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma_0[y/z], A[y/z][u/v]$, hence $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C},l} \Gamma_0[y/z], \exists x A[y/z][x/v]$, i.e. $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C},l} \Gamma[y/z]$.

If $\Gamma = \Gamma_0, \forall x A[x/u]$ with $u \notin \Gamma$, and $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma_0, A$ with $m < n$, $k \leq l$, $\max\{i, \mathsf{rk}(\Gamma)\} \leq j$, then we choose some free variable $v \notin \{y, z\}$ with $v \notin \Gamma$. For $B := A[v/u]$ we get $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma_0, B$ by i.h., and $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma_0[y/z], B[y/z]$ by i.h., and further

$$\forall x B[y/z][x/v] = \forall x B[x/v][y/z] = \forall x A[x/u][y/z].$$

We get $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C},l} \Gamma[y/z]$ because of $v \notin \Gamma_0[y/z], \forall x B[y/z][x/v]$.

If $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma, A$, and $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma, \neg A$, and $\mathsf{rk}_{\mathcal{C}}(A) \leq l$, $k \leq l$, $m < n$, and $\max\{i, \mathsf{rk}(\Gamma)\} \leq j$, then by i.h. we have that $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma[y/z], A[y/z]$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{m,i}_{\mathcal{C},k} \Gamma[y/z], \neg A[y/z]$. And because of $\mathsf{rk}_{\mathcal{C}}(D[y/z]) = \mathsf{rk}_{\mathcal{C}}(D)$ for every formula D, and $\mathsf{rk}(\Gamma[y/z]) = \mathsf{rk}(\Gamma)$, we get $\mathcal{T}\ [\mathcal{R}]\vdash^{n,j}_{\mathcal{C},l} \Gamma[y/z]$.

Similarly for the other cases.

(3) The claim holds trivially, because of $\mathsf{rk}_{\mathcal{C}} = \mathsf{rk}_{\mathcal{C}\cup\neg\mathcal{C}}$.

(4) By induction on n, considering all cases in Definition 12.

If $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, A$, and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},k} \Gamma, \neg A$ with $\mathsf{rk}_{\mathcal{C}}(A) \leq l$, $k \leq l$, $m < n$, and $\max\{i, \mathsf{rk}(\Gamma)\} \leq j$, then we get $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\emptyset,i} \Gamma, A$ and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\emptyset,i} \Gamma, \neg A$ by i.h., hence $\mathsf{rk}(\Gamma, A) \leq i$ and $\mathsf{rk}(A) \leq i \leq j$, hence $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n,j}_{\emptyset,j} \Gamma$.

Similarly for the other cases.

(5) By induction on the structure of the formula A.

If $A = B \vee C$ then $\vdash_0 \neg B, B$ and $\vdash_0 \neg C, C$ by i.h., further $\vdash_0 \neg B, B, C$ and $\vdash_0 \neg C, B, C$ by Part 1. Hence $\vdash_0 \neg B \wedge \neg C, B, C$ and $\vdash_0 \neg B \wedge \neg C, B \vee C$, i.e. $\vdash_0 \neg A, A$.

If $A = \exists x B[x/u]$ then $\vdash_0 \neg B, B$ by i.h., hence $\vdash_0 \exists x B[x/u], \neg B, B$, and $\vdash_0 \neg B, \exists x B[x/u]$, and $\vdash_0 \forall x \neg B[x/u], \exists x B[x/u]$ because of $u \notin \neg A, A$, that is $\vdash_0 \neg A, A$.

Similarly for the other cases.

(6) By induction on n, considering all cases in Definition 12.

If $\Gamma = \Gamma_0, A$ for some $A \in \mathcal{F}$, then by Part 5 we have $\vdash_0 \neg A, A$, hence $\mathcal{T}\ [\mathcal{R}]\!\vdash_{\mathcal{C},l} \neg A, A, \Gamma_0$ by Part 1., i.e. $\mathcal{T}\ [\mathcal{R}]\!\vdash_{\mathcal{C},l} \neg A, \Gamma$.

If $\Gamma = \Gamma_0, A \wedge B$, and $\mathcal{F} \cup \mathcal{T}\ [\mathcal{R}]\!\vdash^{m}_{\mathcal{C},k} \Gamma_0, A$, and $\mathcal{F} \cup \mathcal{T}\ [\mathcal{R}]\!\vdash^{m}_{\mathcal{C},k} \Gamma_0, B$, with $m < n$, $k \leq l$, then there are $\Delta_0 \subseteq \mathcal{F}$, $\Delta_1 \subseteq \mathcal{F}$ with $\mathcal{T}\ [\mathcal{R}]\!\vdash_{\mathcal{C},k} \neg\Delta_0, \Gamma_0, A$ and $\mathcal{T}\ [\mathcal{R}]\!\vdash_{\mathcal{C},k} \neg\Delta_1, \Gamma_0, B$ by i.h., hence if $\Delta := \Delta_0, \Delta_1$ then $\mathcal{T}\ [\mathcal{R}]\!\vdash_{\mathcal{C},k} \neg\Delta, \Gamma_0, A$ and $\mathcal{T}\ [\mathcal{R}]\!\vdash_{\mathcal{C},k} \neg\Delta, \Gamma_0, B$ by Part 1, therefore $\mathcal{T}\ [\mathcal{R}]\!\vdash_{\mathcal{C},l} \neg\Delta, \Gamma$.

If $\Gamma = \Gamma_0, \forall x A[x/u]$, and $\mathcal{F} \cup \mathcal{T}\ [\mathcal{R}]\!\vdash^{m}_{\mathcal{C},k} \Gamma_0, A$, with $m < n$, $k \leq l$, and $u \notin \Gamma$, then for some $\Delta \subseteq \mathcal{F}$ we get $\mathcal{T}\ [\mathcal{R}]\!\vdash_{\mathcal{C},k} \neg\Delta, \Gamma_0, A$ by i.h., hence $\mathcal{T}\ [\mathcal{R}]\!\vdash_{\mathcal{C},l} \neg\Delta, \Gamma$ because $\neg\Delta$ is a set of sentences and $u \notin \neg\Delta, \Gamma$.

Similarly for the other cases.

(7) By induction on n, considering all cases in Definition 12. The only interesting case is when the axiom $\forall x A[x/u]$ or A is used.

If $\mathcal{T} \cup \{\forall x A[x/u]\} \vdash^{n}_{\mathcal{C},l} \Gamma, \forall x A[x/u]$ then we also have $\mathcal{T} \cup \{A\} \vdash^{n}_{\mathcal{C},l} A$, and hence $\mathcal{T} \cup \{A\} \vdash^{n+1}_{\mathcal{C},l} \forall x A[x/u]$, that is $\mathcal{T} \cup \{A\} \vdash^{n+1}_{\mathcal{C},l} \Gamma, \forall x A[x/u]$ by Part 1.

If $\mathcal{T} \cup \{A\} \vdash^{n} \Gamma, A$ then we use Part 5 to get $\vdash_0 A, \neg A$, hence $\vdash_0 A, \neg\forall x A[x]$, and because of $\mathcal{T} \cup \{\forall x A[x]\} \vdash A, \forall x A[x]$ we have $\mathcal{T} \cup \{\forall x A[x]\} \vdash A$, that is $\mathcal{T} \cup \{\forall x A[x]\} \vdash \Gamma, A$ by Part 1.

(8) By induction on n, considering all cases in Definition 12. The only interesting case is when $\neg A \vee B \in \mathcal{Q}$ is used as an axiom or inference rule.

If $\mathcal{T}\ [\mathcal{Q} \cup \mathcal{R}]\!\vdash^{n} \Gamma, B$ and $\mathcal{T}\ [\mathcal{Q} \cup \mathcal{R}]\!\vdash^{m} \Gamma, A$ with $m < n$ then by i.h. and Part 1 we have $(\mathcal{T} \cup \mathcal{Q})\ [\mathcal{R}]\!\vdash \Gamma, A, B$. We get $(\mathcal{T} \cup \mathcal{Q})\ [\mathcal{R}]\!\vdash \Gamma, \neg B, B$ by Parts 1 and 5, hence $(\mathcal{T} \cup \mathcal{Q})\ [\mathcal{R}]\!\vdash \Gamma, A \wedge \neg B, B$, and clearly $(\mathcal{T} \cup \mathcal{Q})\ [\mathcal{R}]\!\vdash^{1} \Gamma, \neg A \vee B, B$, that is $(\mathcal{T} \cup \mathcal{Q})\ [\mathcal{R}]\!\vdash \Gamma, B$.

If $(\mathcal{T} \cup \mathcal{Q})\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, \neg A \vee B$ then we get $\mathcal{T}\ [\mathcal{Q} \cup \mathcal{R}]\!\vdash_{\mathcal{C},l} \Gamma, \neg A, A$ by Parts 1 and 5, hence $\mathcal{T}\ [\mathcal{Q} \cup \mathcal{R}]\!\vdash_{\mathcal{C},l} \Gamma, \neg A, B$, that is $\mathcal{T}\ [\mathcal{Q} \cup \mathcal{R}]\!\vdash_{\mathcal{C},l} \Gamma, \neg A \vee B$. □

Lemma 15. (Inversion)

If $\mathcal{T}, \mathcal{R}, \mathcal{C} \subseteq \mathcal{L}^1$ are *closed* under substitution and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, C$ with $C \notin \mathcal{R}^{\rightarrow} \cup \mathcal{T}$ then we have that

(1) If $C = A \vee B$ then $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, A, B$.
(2) If $C = A \wedge B$ then $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, A$ and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, B$.
(3) If $C = \forall x A[x/u]$ and $B = A[v/u] \in \mathcal{L}^1$ then $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, B$.
(4) If $C = \forall X A[X/U]$ and $B = A[V/U] \in \mathcal{L}^1$ then $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, B$.

Proof.

(1) By induction on n, considering all cases in Definition 12.

If $\Gamma, A \vee B = \Gamma_0, D, \neg D$ and D is atomic, then $\Gamma = \Gamma, D, \neg D$, hence we get $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, A, B$ by Lemma 14.(1).

If $\Gamma, A \vee B = \Gamma_0, D_0 \vee D_1$, and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m}_{\mathcal{C},k} \Gamma_0, D_i$ with $m < n$, $k \leq l$ then we have the following: If $A \vee B = D_0 \vee D_1$ then $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m}_{\mathcal{C},k} \Gamma_0, A, B$, and by i.h. we may assume $A \vee B \notin \Gamma_0$, hence $\Gamma_0 \subseteq \Gamma$ and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, A, B$. Otherwise there is Γ_0' such that $\Gamma_0 = \Gamma_0', A \vee B$ and $\Gamma = \Gamma_0', D \vee E$, and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m}_{\mathcal{C},k} \Gamma_0', A, B, D_i$ by i.h., hence $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, A, B$.

If $\Gamma, A \vee B = \Gamma_0, \exists x D[x/u]$ and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m}_{\mathcal{C},k} \Gamma_0, D[v/u]$ with $m < n$, $k \leq l$ then there is Γ_0' such that $\Gamma_0 = \Gamma_0', A \vee B$ and $\Gamma = \Gamma_0', \exists x D[x/u]$, and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m}_{\mathcal{C},k} \Gamma_0', A, B, D[v/u]$ by i.h., hence $\mathcal{T}\ [\mathcal{R}]\!\vdash^{n}_{\mathcal{C},l} \Gamma, A, B$.

Similarly for the other cases.

(2) Analogous to Part 1.

(3) By induction on n, considering all cases in Definition 12.

If $\Gamma, \forall x A[x/u] = \Gamma_0, \forall y D[y/w]$ and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m}_{\mathcal{C},k} \Gamma_0, D$ with $m < n$, $k \leq l$ and $w \notin \Gamma, \forall x A[x/u]$ then we have the following: If $\forall x A[x/u] = \forall y D[y/w]$ then $A = D[u/w]$, and $B = A[v/u] = D[u/w][v/u] = D[v/w]$ because

$u \notin D$, hence $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{m}{\mathcal{C},k}\ \Gamma_0, B$ by Lemma 14.(2). By i.h. we may assume $\forall x A[x/u] \notin \Gamma_0$, hence $\Gamma_0 \subseteq \Gamma$ and $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n}{\mathcal{C},l}\ \Gamma, B$. Otherwise there is Γ_0' sucht that $\Gamma_0 = \Gamma_0', \forall x A[x/u]$ and $\Gamma = \Gamma_0', \forall y D[y/w]$, and some free variable $z \notin D$ and $z \notin \Gamma, B$, such that $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{m}{\mathcal{C},k}\ \Gamma_0, D[z/w]$ by Lemma 14.(2), and $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{m}{\mathcal{C},k}\ \Gamma_0', B, D[z/w]$ by i.h., hence because of $\forall y D[y/w] = \forall y D[z/w][y/z]$ we get $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n}{\mathcal{C},l}\ \Gamma_0', B, \forall y D[y/w]$, that is $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n}{\mathcal{C},l}\ \Gamma, B$.

Similarly for the other cases (analogous to Part 1).

(4) Analogous to Part 3. $\square$

Another very important property of our provability relation is partial cut elimination, meaning that, if we have $\mathcal{T}\ [\mathcal{R}]\vdash \Gamma$ (with some slight restrictions on $\mathcal{T}$ and $\mathcal{R}$), then we also have $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{}{\mathcal{C},0}\ \Gamma$ for $\mathcal{C} = \mathcal{R}^{\rightarrow} \cup \mathcal{T}$, i.e., there is a derivation of Γ with all cut-formulas in $\mathcal{R}^{\rightarrow} \cup \mathcal{T}$. In the following we give a proof of syntactic cut reduction.

Theorem 16. (Cut-Reduction)

If $\mathcal{T}, \mathcal{R}, \mathcal{C}$ are *closed* under substitution, and $\mathcal{C} \supseteq \mathcal{R}^{\rightarrow} \cup \mathcal{T}$, then we have

(1) If $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n}{\mathcal{C},k}\ \Gamma, A$ and $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{m}{\mathcal{C},k}\ \Delta, \neg A$ and $0 < \mathsf{rk}_{\mathcal{C}}(A) \leq k+1$ then $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n+m}{\mathcal{C},k}\ \Gamma, \Delta$.

(2) If $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n}{\mathcal{C},k+1}\ \Gamma$ then $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{2^n}{\mathcal{C},k}\ \Gamma$.

Proof.

(1) By induction on $n+m$, considering all cases in Definition 12.

First we observe that the statement is symmetric in A and $\neg A$, because if $B = \neg A$ then $\neg B = A$ and $\mathsf{rk}_{\mathcal{C}}(B) = \mathsf{rk}_{\mathcal{C}}(A)$.

If $\Gamma, A = \Gamma_0, \neg D, D$ and D is atomic, then either $A \notin \{\neg D, D\} \subseteq \Gamma$, hence $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n+m}{\mathcal{C},k}\ \Gamma, \Delta$, or otherwise w.l.o.g. $A = D$, hence $\neg A \in \Gamma$, i.e. $\Delta, \neg A \subseteq \Gamma, \Delta$ and $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n+m}{\mathcal{C},k}\ \Gamma, \Delta$.

If $\Gamma, A = \Gamma_0, B$ and $B \in \mathcal{T}$, then we have $B \in \Gamma$ because of $0 < \mathsf{rk}_{\mathcal{C}}(A)$, i.e. $A \notin \mathcal{C} \supseteq \mathcal{T}$, hence $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n+m}{\mathcal{C},k}\ \Gamma, \Delta$.

If $\Gamma, A = \Gamma_0, B_0 \vee B_1$ and $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n_0}{\mathcal{C},k}\ \Gamma_0, B_i$ with $n_0 < n$ then we have the following: If $A \neq B_0 \vee B_1$ then there is Γ_0' such that $\Gamma_0 = \Gamma_0', A$ and $\Gamma = \Gamma_0', B_0 \vee B_1$, i.e. $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n_0}{\mathcal{C},k}\ \Gamma_0', A, B_i$, and $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n_0+m}{\mathcal{C},k}\ \Gamma_0', B_i, \Delta$ by i.h., hence $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n+m}{\mathcal{C},k}\ \Gamma_0', B_0 \vee B_1, \Delta$, that is $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n+m}{\mathcal{C},k}\ \Gamma, \Delta$. If $A = B_0 \vee B_1$, i.e. $\neg A = \neg B_0 \wedge \neg B_1$, then we have $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n_0}{\mathcal{C},k}\ \Gamma, A, B_i$ because of $\Gamma, A = \Gamma_0, A$, and hence $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{n_0+m}{\mathcal{C},k}\ \Gamma, \Delta, B_i$ by i.h.. Further we may assume $\mathcal{T}\ [\mathcal{R}]\vert\!\!\frac{m_0}{\mathcal{C},k}\ \Delta, \neg B_0$ and

$\mathcal{T}\ [\mathcal{R}]\vdash^{m_0}_{\mathcal{C},k} \Delta, \neg B_1$ for some $m_0 < m$, because any other case is treated elsewhere by symmetry. Finally $\mathcal{T}\ [\mathcal{R}]\vdash^{n+m}_{\mathcal{C},k} \Gamma, \Delta$ because of $\mathcal{T}\ [\mathcal{R}]\vdash^{m_0}_{\mathcal{C},k} \Gamma, \Delta, \neg B_i$ and $\mathsf{rk}_{\mathcal{C}}(B_i) \leq k$.

If $\Gamma, A = \Gamma_0, \forall x B[x/u]$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0}_{\mathcal{C},k} \Gamma_0, B$ with $n_0 < n$ and $u \notin \Gamma, A$ then we have the following: If $A \neq \forall x B[x/u]$ then there is Γ_0' and some free variable $v \notin \Gamma, A, \Delta$ such that $\Gamma_0 = \Gamma_0', A$, and $\Gamma = \Gamma_0', \forall x B[x/u]$, and $\forall x B[x/u] = \forall x B[v/u][x/v]$. By Lemma 14.(2) we have $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0}_{\mathcal{C},k} \Gamma_0', A, B[v/u]$, and we get $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0+m}_{\mathcal{C},k} \Gamma_0', B[v/u], \Delta$ by i.h., hence $\mathcal{T}\ [\mathcal{R}]\vdash^{n+m}_{\mathcal{C},k} \Gamma_0', \forall x B[v/u][x/v], \Delta$ because $v \notin \Gamma, \Delta$, that is $\mathcal{T}\ [\mathcal{R}]\vdash^{n+m}_{\mathcal{C},k} \Gamma, \Delta$. If $A = \forall x B[x/u]$, i.e. $\neg A = \exists x \neg B[x/u]$, then we have $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0}_{\mathcal{C},k} \Gamma, A, B$ because of $\Gamma, A = \Gamma_0, A$. Further we may assume $\mathcal{T}\ [\mathcal{R}]\vdash^{m_0}_{\mathcal{C},k} \Delta, \neg B[w/u]$ for some $m_0 < m$, because any other case is treated elsewhere by symmetry. We get $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0}_{\mathcal{C},k} \Gamma, A, B[w/u]$ by Lemma 14.(2), hence $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0+m}_{\mathcal{C},k} \Gamma, \Delta, B[w/u]$ by i.h.. Finally we get $\mathcal{T}\ [\mathcal{R}]\vdash^{n+m}_{\mathcal{C},k} \Gamma, \Delta$ because $\mathcal{T}\ [\mathcal{R}]\vdash^{m_0}_{\mathcal{C},k} \Gamma, \Delta, \neg B[w/u]$ and $\mathsf{rk}_{\mathcal{C}}(B[w/u]) \leq k$.

If $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0}_{\mathcal{C},k} \Gamma, A, B$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0}_{\mathcal{C},k} \Gamma, A, \neg B$ with $n_0 < n$ and $\mathsf{rk}_{\mathcal{C}}(B) \leq k$ then we can get $\mathcal{T}\ [\mathcal{R}]\vdash^{m}_{\mathcal{C},k} \Delta, \neg A, B$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{m}_{\mathcal{C},k} \Delta, \neg A, \neg B$ by Lemma 14.(1), and $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0+m}_{\mathcal{C},k} \Gamma, \Delta, B$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{n_0+m}_{\mathcal{C},k} \Gamma, \Delta, \neg B$ by i.h., hence $\mathcal{T}\ [\mathcal{R}]\vdash^{n+m}_{\mathcal{C},k} \Gamma, \Delta$ because of $\mathsf{rk}_{\mathcal{C}}(B) \leq k$.

Similarly for the other cases.

(2) By induction on n, considering all cases in Definition 12.

If $\mathcal{T}\ [\mathcal{R}]\vdash^{m}_{\mathcal{C},k+1} \Gamma, A$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{m}_{\mathcal{C},k+1} \Gamma, \neg A$ with $m < n$ and $\mathsf{rk}_{\mathcal{C}}(A) \leq k+1$ then $\mathcal{T}\ [\mathcal{R}]\vdash^{2^m}_{\mathcal{C},k} \Gamma, A$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{2^m}_{\mathcal{C},k} \Gamma, \neg A$ by i.h.; if $\mathsf{rk}_{\mathcal{C}}(A) = 0$ then $\mathcal{T}\ [\mathcal{R}]\vdash^{2^m+1}_{\mathcal{C},k} \Gamma$, otherwise $\mathcal{T}\ [\mathcal{R}]\vdash^{2^m+2^m}_{\mathcal{C},k} \Gamma, \Gamma$ by Part 1, hence $\mathcal{T}\ [\mathcal{R}]\vdash^{2^n}_{\mathcal{C},k} \Gamma$.

If $\Gamma = \Gamma_0, \forall x A[x/u]$ and $\mathcal{T}\ [\mathcal{R}]\vdash^{m}_{\mathcal{C},k+1} \Gamma_0, A$ with $m < n$ and $u \notin \Gamma$ then $\mathcal{T}\ [\mathcal{R}]\vdash^{2^m}_{\mathcal{C},k} \Gamma_0, A$ by i.h., hence $\mathcal{T}\ [\mathcal{R}]\vdash^{2^n}_{\mathcal{C},k} \Gamma$.

Similarly for the other cases. □

Corollary 17. (Partial Cut-Elimination)

If $\mathcal{T}, \mathcal{R}, \mathcal{C}$ are *closed* under substitution, and $\mathcal{C} \supseteq \mathcal{R}^{\rightarrow} \cup \mathcal{T}$, and $\mathcal{T}\ [\mathcal{R}]\vdash^{n}_{\mathcal{C},k} \Gamma$, then $\mathcal{T}\ [\mathcal{R}]\vdash^{2^n_k}_{\mathcal{C},0} \Gamma$, where $2^n_0 = n$ and $2^n_{m+1} = 2^{2^n_m}$.

2. Standard Theories of Sets and Classes

Having established the formal setting of the language and the logic for set theory in Section 1, we are now able to formulate the axioms of Von Neumann–Bernays–Gödel set theory (see e.g. Mendelson [14] for more detailed information about NBG set theory).

We first need to define a lot of syntactic abbreviations, because our language of set theory is very simple. As far as we can, we use common notation from standard set theory for these abbreviations, such that, e.g. Zermelo-Fraenkel set theory, which is also introduced below, can be defined as usual (see e.g. Takeuti [18] for more detailed information about ZFC set theory).

We present an infinite axiomatization of NBG, that is, we extend ZFC by the axiom schema of class comprehension, and by the axiom of replacement (formulated for class functions), and by the axiom of global choice, to get NBG. We show that the axiom of separation (also called axiom of subsets), and the extensionality axiom for classes are both derivable in NBG. We further show the existence of a global wellordering in NBG, and later, in Section 4, we also prove the foundation axiom for classes in NBG.

Definition 18. (Abbreviations)

We define expressions σ of the form $\cup x$, $\cap x$, $\langle x, y\rangle$, $x\backslash y$, x', $\{x_1, ..., x_n\}$, and abbreviations $\sigma \in z$, $\sigma \in \sigma$, and $z \in \sigma$, for formulas, such that $\sigma \in z := (\exists u \in z)\forall v(v \in u \leftrightarrow v \in \sigma)$, and $\sigma \in \sigma := (\exists u \in \sigma)\forall v(v \in u \leftrightarrow v \in \sigma)$, and

$$
\begin{aligned}
z \in \cup x &:= (\exists w \in x)z \in w, \\
z \in \cap x &:= (\forall w \in x)z \in w, \\
z \in \{x_1, ..., x_n\} &:= ((\perp \vee z = x_1) ... \vee z = x_n), \\
z \in \langle x, y\rangle &:= z = \{x\} \vee z = \{x, y\}, \\
z \in x\backslash y &:= z \in x \wedge z \notin y, \\
z \in x' &:= z \in x \vee z = x.
\end{aligned}
$$

We write $x \cup y$ for $\cup\{x, y\}$, and $x \cap y$ for $\cap\{x, y\}$, and we write $\overline{0}$ and $\emptyset$ for $\{\}$, and $\overline{n+1}$ for $\overline{n} \cup \{\overline{n}\}$ $(n \in \mathbb{N})$.

We also use all abbreviations with *set variables replaced by class variables*, whenever the resulting expressions translate into proper $\mathcal{L}^1$-formulas. If

$A \in \mathcal{L}^1$ and σ is some expression of the form $\cup x$, $\cap x$, $\langle x, y\rangle$, $x\backslash y$, x', $\{x_1, ..., x_n\}$, or of the same form but with class variables, then $A[\sigma/u]$ stands for A with all subformulas $z \in u$, $u \in z$, $u \in Z$ with $z \neq u$, replaced by the formulas $z \in \sigma$, $\sigma \in z$, $\sigma \in Z$, respectively, and with $u \in u$ replaced by $\sigma \in \sigma$, where all bound variables in $z \in \sigma$, $\sigma \in \sigma$, $\sigma \in z$, $\sigma \in Z$, do not occur in A. Analogously $A[\sigma/V]$ for class variables V. We write $A[x \cup y/u]$, i.e. $A[\cup\{x, y\}/u]$, for $A[\cup v/u][\{x, y\}/v]$ (where v is some fresh variable). Analogously we write $A[\sigma/u]$ and $A[\sigma/V]$ for other composed expressions σ, e.g. we write $A[\overline{n+1}/u]$ for the formula $A[\cup v_1/u][\{v_2, v_3\}/v_1][\{v_2\}/v_3][\overline{n}/v_2]$ (where v_1, v_2, v_3 are some fresh variables).

Definition 19. (Abbreviations)

We define the expressions $(x)_y$, $(x)_{u,v}$, $f(x)$, $f[x]$, $dom(f)$, $ran(f)$, $f{\restriction}x$, f^{-1}, $\langle\!\langle x_1, ..., x_n\rangle\!\rangle$, (analogous Definition 18) such that

$$\begin{array}{rcl}
z \in (x)_y & := & \langle z, y\rangle \in x, \\
z \in (x)_{u,v} & := & \langle\langle z, v\rangle, u\rangle \in x, \\
z \in f(x) & := & \exists w(\langle x, w\rangle \in f \wedge z \in w), \\
z \in f[x] & := & (\exists w \in x) f(w) = z, \\
z \in dom(f) & := & (\exists w)\langle z, w\rangle \in f, \\
z \in ran(f) & := & (\exists w)\langle w, z\rangle \in f, \\
z \in f{\restriction}x & := & z \in f \wedge (\exists w \in x)\exists y\langle w, y\rangle = z, \\
z \in f^{-1} & := & \exists x\exists y(z = \langle x, y\rangle \wedge \langle y, x\rangle \in f), \\
z \in \langle\!\langle x_0, ..., x_n\rangle\!\rangle & := & ((z = \langle\overline{0}, x_0\rangle \vee ...) \vee z = \langle\overline{n}, x_n\rangle).
\end{array}$$

$$\begin{array}{rcl}
Rel[x] & := & (\forall u \in x)\exists v\exists w(u = \langle v, w\rangle), \\
Fun[x] & := & Rel[x] \wedge (\forall u \in dom(x))\exists! v\langle u, v\rangle \in x.
\end{array}$$

Definition 20. Zermelo–Fraenkel Set Theory $\mathsf{ZFC} \subseteq \mathcal{L}^0$

(Extensionality)	$\forall x\forall y(x = y \rightarrow \forall z(x \in z \leftrightarrow y \in z))$,
(Pair)	$\forall x\forall y\exists z(z = \{x, y\})$,
(Union)	$\forall x\exists y(y = \cup x)$,
(Powerset)	$\forall x\exists y\forall z(z \in y \leftrightarrow z \subseteq x)$,
(Infinity)	$\exists x(\emptyset \in x \wedge (\forall y \in x)(y \cup \{y\} \in x)$,
(Foundation)	$\forall x(x = \emptyset \vee (\exists y \in x)(\forall z \in y)z \notin x)$,
(Replacement)	$(\forall)(\forall x\forall y\forall z(A[x, y] \wedge A[x, z] \rightarrow y = z) \rightarrow$ $\forall u\exists v\forall y(y \in v \leftrightarrow (\exists x \in u)A[x, y]))$ for any $A \in \mathcal{L}^0$ and distinct $x, y, z, u, v \notin A$, where $(\forall)B$ stands for the universal closure of B,
(Choice)	$\forall x\exists f(Fun[f] \wedge (\forall y \in x)(y = \emptyset \vee f(y) \in y))$.

Definition 21. Von Neumann–Bernays–Gödel Set Theory $\mathsf{NBG} \subseteq \mathcal{L}^1$

(ZFC)	all axioms of ZFC,
(Comprehension)	$\exists X \forall y (y \in X \leftrightarrow A[y])$ for any $A \in \Pi^1_0,\ X, y \notin A$,
(Replacement)	$\forall F (\mathit{Fun}[F] \to \forall x \exists y (y = F[x]))$,
(Global Choice)	$\exists F (\mathit{Fun}[F] \wedge \forall x (x = \emptyset \vee F(x) \in x))$.

According to a well-known result, see, e.g., Levy [13], NBG is a conservative extension of ZFC.

Theorem 22. (Conservative Extension)

For $A \in \mathcal{L}^0$ we have $\mathsf{NBG} \vdash A$ if and only if $\mathsf{ZFC} \vdash A$.

The following theorem, showing that separation holds in NBG, is proved mainly by using the replacement axiom. The subsequent corollary is important because it guarantees extensionality for NBG set theory (see also Appendix A).

Theorem 23. (Separation)

$\mathsf{NBG} \vdash \forall x \forall Y \exists z\, (z = Y \cap x)$.

Proof. If $A[v, U] = (\exists z \in U)(v = \langle z, z \rangle)$ then by logic we get

$$\forall x (x \in V \leftrightarrow A[x, U]) \to \mathit{Fun}[V] \wedge \mathit{dom}(V) = U \wedge (\forall z \in U)(V(z) = z).$$

By the axiom of replacement we have

$$\forall x (x \in V \leftrightarrow A[x, U]) \to \forall x \exists y (y = \{V(z) \mid z \in \mathit{dom}(V) \cap x\}),$$

and because of $\mathit{dom}(V) = U$ and $(\forall z \in U)(V(z) = z)$ we get

$$\forall Y (\forall x (x \in Y \leftrightarrow A[x, U]) \to \forall x \exists y (y = U \cap x)),\ \text{i.e.}$$

$$\exists Y \forall x (x \in Y \leftrightarrow A[x, U]) \to \forall x \exists y (y = U \cap x).$$

We have elementary comprehension, hence finally $\forall x \exists y (y = U \cap x)$. $\square$

Corollary 24. (Extensionality)

$\mathsf{NBG} \vdash \forall x \forall y (x = y \to \forall Z (x \in Z \leftrightarrow y \in Z))$.

Proof. Assume there are x, y, Z, such that $x = y$, and $x \in Z$, and $y \notin Z$. By separation we have that $Z \cap \{x\}$ is a set with $x \in Z \cap \{x\}$ and $y \notin Z \cap \{x\}$, that is in contradiction to the extensionality axiom in ZFC. $\square$

Besides the expressions, as, e.g., in Definition 18, which we use as syntactic abbreviations, we also want to use shorthand notation for some specific classes. The following definition is about the definability of classes, e.g., about elementarily definable classes. For a few definable classes we are going to introduce some sort of constants in the meta-language.

Definition 25. (Definable Classes)

We write $Def[X, A[y]]$ for $\forall y(y \in X \leftrightarrow A[y])$ if $y \notin A \in \mathcal{L}^1$ and if A has only *one free* variable. If the name of the variable y is not important then we just write $Def[X, A]$. The collection $\mathcal{C} \subseteq |\mathcal{M}|$ is called *definable* by A if $\mathcal{C} = \{a \in |\mathcal{M}| \mid \mathcal{M} \models A[a/x]\}$ and x is the only free variable in A. The class $b \in \|\mathcal{M}\|$ is called *definable* by A if $\mathcal{M} \models Def[X, A][b/X]$. $b \in \|\mathcal{M}\|$ is called *definable in* $\mathcal{T} \subseteq \mathcal{L}^1$ by A, if b is definable by A and $\mathcal{T} \vdash \exists X\, Def[X, A]$. Further, if $A \in \mathcal{L}^0$ then b and $\mathcal{C}$ are called *elementarily* definable (in $\mathcal{T}$).

By the comprehension axiom in NBG, we have the following elementarily definable classes.

Lemma 26.

The following classes are elementarily definable in NBG,

$$\begin{array}{rcl} \mathcal{V} & := & \{x \mid x = x\}, \\ \mathcal{O}n & := & \{x \mid \cup x \subseteq x \wedge (\forall y \in x)(\cup y \subseteq y)\}, \\ \in & := & \{x \mid \exists y \exists z(x = \langle y, z\rangle \wedge y \in z)\}, \\ \mathcal{P} & := & \{x \mid \exists y \exists z(x = \langle y, z\rangle \wedge \forall u(u \in z \leftrightarrow u \subseteq y))\}, \end{array}$$

$\mathcal{V}, \mathcal{O}n, \in, \mathcal{P}$ are the class of all sets, the class of all ordinals, the element relation on sets, and the powerset function, respectively. (We already used the symbol $\mathcal{V}$ to denote $\mathcal{L}^1$-valuations; it will always be clear from context whether $\mathcal{V}$ is the class of all sets or a valuation.)

In the following we use the expressions $\mathcal{V}$, $\mathcal{O}n$, $\in$, $\mathcal{P}$, denoting the definable classes in the previous lemma, analogous to the way we use the expressions σ in Definition 18.

By having the class of all ordinals $\mathcal{O}n$ defined, we are now going to use special letters to denote variables ranging over the ordinals.

Definition 27. (Ordinals)

We use the greek letters $\alpha, \beta, \gamma, \delta, \kappa$ (with subscripts) to denote ordinals. We write $\forall \alpha A[\alpha]$ and $\exists \alpha A[\alpha]$ for $(\forall x \in \mathcal{O}n)A[x]$ and $(\exists x \in \mathcal{O}n)A[x]$, respectively.

For putting the axiom of global choice to work, in most situations, it is very useful to have a global wellordering at hand. We define the formula Gl such that $Gl[F]$ means that F is a bijection $F : \mathcal{V} \to \mathcal{O}n$. The following lemma shows that NBG proves the existence of such global wellorderings F.

Definition 28. (Global Wellordering)

$$Gl[F] \quad := \quad Rel[F] \wedge \forall x \exists! \alpha \langle x, \alpha \rangle \in F \wedge \forall \alpha \exists! x \langle x, \alpha \rangle \in F.$$

Lemma 29. (Global Wellordering)

$\mathsf{NBG} \vdash \exists X\, Gl[X]$.

Proof. We show that there is a bijective function $W : \mathcal{O}n \to \mathcal{V}$ in NBG. By the recursion principle (definition by transfinite induction, see e.g. Mendelson [14]) we have the following two class functions

$$\begin{array}{rcl} V(\alpha) & := & \bigcup\{\mathcal{P}(V(\beta)) \mid \beta < \alpha\}, \\ R(x) & := & \bigcup\{R(y) + 1 \mid y \in x\}. \end{array}$$

We write $\rho(x)$ for $R(x) + 1$, and by transfinite induction we get

$$\rho(x) \leq \alpha \leftrightarrow x \in V(\alpha).$$

We extend some global choice function to get C, with $C(\emptyset) = \emptyset$, and by comprehension we have the following class functions

$$\begin{array}{rcl} F(x) & := & \bigcup\{\rho(y) \mid y \in ran(x)\}, \\ G(x) & := & \begin{cases} C(V(F(x)) \setminus ran(x)) & V(F(x)) \setminus ran(x) \neq \emptyset, \\ C(V(F(x)+1) \setminus ran(x)) & \text{otherwise.} \end{cases} \end{array}$$

We define the function W by the recursion principle, such that

$$W(\alpha) = G(W \restriction \alpha).$$

If $W(\alpha) = W(\beta)$ for $\alpha < \beta$ then $W(\beta) \in ran(W \restriction \beta)$ in contradiction to $W(\beta) = G(W \restriction \beta)$, hence W is one-one. For any set x and any cardinal $\kappa > |V(\rho(x))|$ we have that $|\,ran(W \restriction \kappa)| = \kappa$ because W is one-one, hence there is some $\alpha < \kappa$ with $\rho(W(\alpha)) > \rho(x)$, because otherwise $ran(W \restriction \kappa) \subseteq V(\rho(x))$ in contradiction to $|\,ran(W \restriction \kappa)| > |V(\rho(x))|$. If $\beta := \min\{\alpha \mid \rho(W(\alpha)) > \rho(x)\}$ and $\gamma := F(W \restriction \beta)$ then

$$\gamma \leq \rho(x) < \rho(W(\beta)),$$

and we have $V(\gamma) \setminus ran(W \restriction \beta) = \emptyset$ because otherwise we get a contradiction by the definition of W, i.e. $W(\beta) \in V(\gamma)$ and $\rho(W(\beta)) \leq \gamma$. By the definition of W we have $W(\beta) \in V(\gamma + 1)$, that is $\rho(W(\beta)) \leq \gamma + 1$, hence $\rho(x) = \gamma$ and $x \in V(\gamma) \subseteq ran(W \restriction \beta)$, i.e. any set x is in the range of W. $\square$

3. Choice Schemes

In this section we define the axiom schemata AC and DC (see e.g. Feferman and Sieg [3], Simpson [17]) for choice and dependent choice, respectively, and we show that a bunch of slightly different formulations of the schemata of dependent choice are all equivalent over NBG. By definition, we will have that dependent choice implies choice, and by a standard argument we further get that choice, AC, implies class comprehension CA for some formulas.

Many results of this section about AC and DC are well-established in second order arithmetic (see Feferman and Sieg [3], Simpson [17]), and all the arguments from arithmetic are easily translated into set theory, nevertheless we are going to give proves for all statements.

We further define the collection principle Col, which is shown to be equivalent to choice AC over NBG extended by class comprehension CA for appropriate formulas. The collection principle Col was introduced in Jäger and Krähenbühl [10] to deal with the asymmetric interpretation, i.e. to make the interpretation simpler. In this thesis we are not going to use the collection principle Col any further, because we directly deal with choice instead of collection in the asymmetric interpretations in Section 10.

Definition 30. (Abbreviations)

We define the expressions $(x)^y$ and $\langle x\rangle^y$ (analogous Definition 18) such that

$$\begin{array}{rcl} z \in (x)^y & := & z \in x \wedge (\exists w \in y)\exists v \langle v, w\rangle = z, \\ z \in \langle x\rangle^y & := & (\exists w \in y)\langle z, w\rangle \in x. \end{array}$$

Most of the choice schemes defined below are analogous to the choice schemes in second order arithmetic, as for example in Simpson [17].

Definition 31. (Choice Schemes)

For $x, y, z, Y, Z, \alpha \notin A \in \mathcal{L}^1$ with $x \neq z$, $Y \neq Z$ we define

$$\begin{array}{rcl} CA[A[x], Y] & := & \exists Y \forall x(x \in Y \leftrightarrow A[x]), \\ Col[A[x, Y], z] & := & \forall x \exists Y A[x, Y] \to \exists Y \forall x \exists z A[x, (Y)_z], \\ AC[A[x, Y]] & := & \forall x \exists Y A[x, Y] \to \exists Y \forall x A[x, (Y)_x], \\ AC_{\mathcal{O}n}[A[\alpha, Y]] & := & \forall \alpha \exists Y A[\alpha, Y] \to \exists Y \forall \alpha A[\alpha, (Y)_\alpha], \end{array}$$

$$\begin{array}{rcl}
DC[A[x,Y,Z]] & := & \forall x \forall Y \exists Z A[x,Y,Z] \to \exists Z \forall x A[x,(Z)^x,(Z)_x], \\
DC^{\Diamond}[A[x,Y,Z]] & := & \forall x \forall Y \exists Z A[x,Y,Z] \to \exists Z \forall x A[x,\langle Z\rangle^x,(Z)_x], \\
DC^{-}[A[Y,Z]] & := & \forall Y \exists Z A[Y,Z] \to \exists Z \forall x A[(Z)^x,(Z)_x], \\
DC^{\Diamond -}[A[Y,Z]] & := & \forall Y \exists Z A[Y,Z] \to \exists Z \forall x A[\langle Z\rangle^x,(Z)_x], \\
 & & \\
DC_{\mathcal{O}n}[A[\alpha,Y,Z]] & := & \forall \alpha \forall Y \exists Z A[\alpha,Y,Z] \to \exists Z \forall \alpha A[\alpha,(Z)^\alpha,(Z)_\alpha], \\
DC^{\Diamond}_{\mathcal{O}n}[A[\alpha,Y,Z]] & := & \forall \alpha \forall Y \exists Z A[\alpha,Y,Z] \to \exists Z \forall \alpha A[\alpha,\langle Z\rangle^\alpha,(Z)_\alpha], \\
DC^{-}_{\mathcal{O}n}[A[Y,Z]] & := & \forall Y \exists Z A[Y,Z] \to \exists Z \forall \alpha A[(Z)^\alpha,(Z)_\alpha], \\
DC^{\Diamond -}_{\mathcal{O}n}[A[Y,Z]] & := & \forall Y \exists Z A[Y,Z] \to \exists Z \forall \alpha A[\langle Z\rangle^\alpha,(Z)_\alpha].
\end{array}$$

If $\mathcal{F} \subseteq \mathcal{L}^1$ then $Col[\mathcal{F}] := \{Col[A[x,Y],z] \in \mathcal{L}^1 \mid x,Y,z \notin A \in \mathcal{F},\ x \neq z\}$, analogously $CA[\mathcal{F}]$, $AC[\mathcal{F}]$, $DC[\mathcal{F}]$, etc..

We need to say a few words about the notation we just used, e.g. about the arguments $A[x]$, Y in the construction of $CA[A[x],Y]$. We already know that $A[x]$ stands for $A[x/u]$ for some variable u. If u is a free variable with $u \in A \in \mathcal{L}^1$, and $x \notin A$ is a bound variable, then $A[x] \notin \mathcal{L}^1$. Even though $A[x]$ is not a formula, the expression $A[x]$ holds essential information for the construction of $CA[A[x],Y]$; that's why we write $CA[A[x],Y]$ instead of, e.g., $CA[A,x,Y]$.

The next two lemmas show that Π^1_n and Σ^1_n are essentially closed under the logical connectors $\wedge$, $\vee$, and the set quantifiers $\forall x$, $\exists x$, i.e. closed modulo equivalence of formulas. The choice principle AC is needed to have closure for the set quantifiers, e.g. if $A,B \in \Pi^1_n$ and C is one of the formulas $A \wedge B$, $A \vee B$, $\forall x A[x]$, or $\exists x A[x]$, then there is some $D \in \Pi^1_n$ such that $\mathsf{NBG} \cup AC[\Pi^1_{n-1}]$ proves $C \leftrightarrow D$. This closure property is repeatedly used throughout this section.

Lemma 32.

If $\mathcal{F} \in \{\Pi^1_n, \Sigma^1_n\}$ and $A,B \in \mathcal{F}$ then there are $C,D \in \mathcal{F}$ such that

(1) $\mathsf{NBG} \vdash C \leftrightarrow (A \wedge B)$,

(2) $\mathsf{NBG} \vdash D \leftrightarrow (A \vee B)$.

Proof. By induction on n. If $n = 0$ then $A \wedge B$ and $A \vee B$ are in $\mathcal{F}$. If $\mathcal{F} = \Sigma^1_{n+1}$ and $A = \exists X A_0[X]$ and $B = \exists Y B_0[Y]$ with $A_0, B_0 \in \Pi^1_n$ then by i.h. there is some $D_0[U] \in \Pi^1_n$ such that

$$\mathsf{NBG} \vdash D_0[U] \leftrightarrow (A_0[(U)_{\overline{0}}] \vee B_0[(U)_{\overline{1}}]).$$

For $D := \exists Z D_0[Z] \in \Sigma^1_{n+1}$ we get $\mathsf{NBG} \vdash D \leftrightarrow (A \vee B)$ because of

$$\mathsf{NBG} \vdash \exists Z (A_0[(Z)_{\overline{0}}] \vee B_0[(Z)_{\overline{1}}]) \leftrightarrow (A \vee B).$$

Analogously for $A \wedge B$ and all other cases of A, B and $\mathcal{F}$. $\square$

Lemma 33.

If $\mathcal{F} \in \{\Pi^1_n, \Sigma^1_n\}$ and $x \notin A \in \mathcal{F}$ then there are $C, D \in \mathcal{F}$ such that

(1) $\mathsf{NBG} \cup AC[\Pi^1_{n-1}] \vdash C \leftrightarrow \forall x A[x]$,

(2) $\mathsf{NBG} \cup AC[\Pi^1_{n-1}] \vdash D \leftrightarrow \exists x A[x]$,

where $\Pi^1_{n-1} = \emptyset$ for $n = 0$.

Proof. By induction on n. If $n = 0$ then $\forall x A[x]$ and $\exists x A[x]$ are in $\mathcal{F}$. If $\mathcal{F} = \Pi^1_{n+1}$ and $A = \forall Y A_0[Y]$ with $A_0 \in \Sigma^1_n$ then $\neg \exists x A[x] = \forall x \exists Y \neg A_0[x, Y]$ with $\neg A_0 \in \Pi^1_n$. We have $\forall x \exists Y \neg A_0[x, Y] \leftrightarrow \exists Y \forall x \neg A_0[x, (Y)_x]$ by $AC[\Pi^1_n]$, and by i.h. there is some $D_0[U] \in \Pi^1_n$ such that $D_0[U] \leftrightarrow \forall x \neg A_0[x, (U)_x]$, hence for $D := \neg \exists Y D_0[Y] \in \Pi^1_{n+1}$ we have $D \leftrightarrow \exists x A[x]$. By i.h. there is $C_0[U] \in \Sigma^1_n$ such that $C_0[U] \leftrightarrow \forall x A_0[x, U]$, hence for $C := \forall Y C_0[Y] \in \Pi^1_{n+1}$ we have $C \leftrightarrow \forall x A[x]$, because of $\forall Y \forall x A_0[x, Y] \leftrightarrow \forall x \forall Y A_0[x, Y]$.

Analogously for all other cases of A and $\mathcal{F}$. $\square$

The following lemma shows that if we have choice or dependent choice for Π^1_n formulas then we also get it for Σ^1_{n+1} formulas.

Lemma 34.

(1) $\mathsf{NBG} \cup AC[\Pi^1_n] \vdash AC[\Sigma^1_{n+1}]$,

(2) $\mathsf{NBG} \cup DC^{\diamondsuit}[\Pi^1_n] \vdash DC^{\diamondsuit}[\Sigma^1_{n+1}]$.

Proof.

(1) If $A \in \Sigma^1_{n+1}$ then $A \in \Pi^1_n$ or $A = \exists Z B[Z]$ for some $B \in \Pi^1_n$, such that $A[x, Y] = \exists Z B[x, Y, Z]$. If $C[x, Y] := B[x, (Y)_0, (Y)_1] \in \Pi^1_n$, then $\forall x \exists Y A[x, Y] \leftrightarrow \forall x \exists Y C[x, Y]$, and we have that the following are equivalent

(a) $\exists Y \forall x A[x, (Y)_x] = \exists Y \forall x \exists Z B[x, (Y)_x, Z]$,

(b) $\exists Y \exists Z \forall x B[x, (Y)_x, (Z)_x]$,

(c) $\exists Y \forall x B[x, ((Y)_{\overline{0}})_x, ((Y)_{\overline{1}})_x]$,

(d) $\exists Y \forall x B[x, ((Y)_x)_{\overline{0}}, ((Y)_x)_{\overline{1}}] = \exists Y \forall x C[x, (Y)_x]$,

that is (a) $\leftrightarrow$ (b) by $AC[\Pi^1_n]$, and (b) $\leftrightarrow$ (c), (c) $\leftrightarrow$ (d) by elementary comprehension. Hence we have $\mathsf{NBG} \cup AC[\Pi^1_n] \vdash AC[C[x, Y]] \to AC[A[x, Y]]$, that is $\mathsf{NBG} \cup AC[\Pi^1_n] \vdash AC[A[x, Y]]$.

(2) If $A[x, X, Y] = \exists Z B[x, X, Y, Z] \in \Sigma^1_{n+1}$ with $B \in \Pi^1_n$ and $C[x, X, Y] := B[x, (X)_0, (Y)_0, (Y)_1] \in \Pi^1_n$, then we have that

$$\forall x \forall X \exists Y A[x, X, Y] \leftrightarrow \forall x \forall X \exists Y C[x, X, Y],$$

and the following are equivalent

(a) $\exists Y \forall x A[x, \langle Y\rangle^x, (Y)_x] = \exists Y \forall x \exists Z B[x, \langle Y\rangle^x, (Y)_x, Z]$,

(b) $\exists Y \exists Z \forall x B[x, \langle Y\rangle^x, (Y)_x, (Z)_x]$,

(c) $\exists Y \forall x B[x, \langle (Y)_{\overline{0}}\rangle^x, ((Y)_{\overline{0}})_x, ((Y)_{\overline{1}})_x]$,

(d) $\exists Y \forall x B[x, (\langle Y\rangle^x)_{\overline{0}}, ((Y)_x)_{\overline{0}}, ((Y)_x)_{\overline{1}}] = \exists Y \forall x C[x, \langle Y\rangle^x, (Y)_x]$,

that is, (a) $\leftrightarrow$ (b) by $AC[\Pi^1_n]$ (we have $DC^\diamondsuit[\Pi^1_n] \vdash AC[\Pi^1_n]$), and (b) $\leftrightarrow$ (c), (c) $\leftrightarrow$ (d) by comprehension, that is if $Z := \{\langle\langle z, x\rangle, y\rangle \mid \langle\langle z, y\rangle, x\rangle \in Y\}$ then

$$((Z)_y)_x = ((Y)_x)_y \text{ and } (\langle Z\rangle^y)_x = \langle (Y)_x\rangle^y.$$

Hence we have that $\mathsf{NBG} \cup DC^\diamondsuit[\Pi^1_n] \vdash DC^\diamondsuit[C[x, Y]] \to DC^\diamondsuit[A[x, Y]]$, and finally $\mathsf{NBG} \cup DC^\diamondsuit[\Pi^1_n] \vdash DC^\diamondsuit[A[x, Y]]$. $\square$

The next theorem shows that we get class comprehension from choice. And together with the previous lemma we get that choice for Π^1_n formulas implies comprehension for all Π^1_n and Σ^1_n formulas (because $\Pi^1_n \cup \Sigma^1_n \subseteq \Pi^1_{n+1} \cap \Sigma^1_{n+1}$).

Theorem 35.

For $A \in \Sigma^1_n$ and $B \in \Pi^1_n$ we have

$$\mathsf{NBG} \cup AC[\Sigma^1_n] \vdash \forall x(A[x] \leftrightarrow B[x]) \to \exists Y \forall x(x \in Y \leftrightarrow A[x]).$$

Proof. By Lemma 32 there is some $C[u, U] \in \Sigma^1_n$ such that

$$C[u, U] \leftrightarrow ((A[u] \wedge \emptyset \in U) \vee (\neg B[u] \wedge \emptyset \notin U)).$$

We have $\forall x(A[x] \leftrightarrow B[x]) \to \forall x \exists Y C[x, Y]$, hence by $AC[\Sigma^1_n]$ we get

$$\forall x(A[x] \leftrightarrow B[x]) \to \exists Y \forall x C[x, (Y)_x].$$

If we define $D_0 := \forall x(\emptyset \in (Z)_x \leftrightarrow A[x])$, $D_1 := \exists Y \forall x(x \in Y \leftrightarrow \emptyset \in (Z)_x)$, and $D_2 := \exists Y \forall x(x \in Y \leftrightarrow A[x])$, then we have that

$$\forall x(A[x] \leftrightarrow B[x]) \wedge \forall x C[x, (Z)_x] \to D_0,$$

and further $D_0 \to (D_1 \to D_2)$, hence $D_0 \to D_2$ by elementary comprehension, i.e.

$$\forall x(A[x] \leftrightarrow B[x]) \wedge \exists Y \forall x C[x, (Y)_x] \to D_2,$$

and finally $\forall x(A[x] \leftrightarrow B[x]) \to D_2$. $\square$

Corollary 36.

$\mathsf{NBG} \cup AC[\Pi^1_n] \vdash CA[\Pi^1_n \cup \Sigma^1_n]$.

Proof. By Lemma 34 and Theorem 35. □

The next lemma is the main lemma to get the equivalence between the collection principle *Col* and choice *AC*. The lemma makes essential use of the existence of a global wellordering in NBG.

Lemma 37.

For $A \in \Pi^1_n$ we have that

$$\mathsf{NBG} \cup CA[\Pi^1_n] \vdash \exists Y \forall x \exists z A[x, (Y)_z] \to \exists Y \forall x A[x, (Y)_x].$$

Proof. We assume that $\forall x \exists y A[x, (Z)_y]$ for Z, and $W : \mathcal{O}n \to \mathcal{V}$ is such that $Gl[W^{-1}]$ (see Lemma 29). By $CA[\Pi^1_n]$ we get the classes

$$\begin{array}{rcl} X & := & \{\langle x, \alpha\rangle \mid A[x, (Z)_{W(\alpha)}]\}, \\ F & := & \{\langle x, \alpha\rangle \mid \langle x, \alpha\rangle \in X \wedge (\forall \beta < \alpha)\langle x, \alpha\rangle \notin X\}, \\ Y & := & \{\langle z, x\rangle \mid z \in (Z)_{W(F(x))}\}, \end{array}$$

such that $F : \mathcal{V} \to \mathcal{O}n$ and $\forall x A[x, (Z)_{W(F(x))}]$ by the definition of F, and hence $\forall x A[x, (Y)_x]$. □

Corollary 38.

$\mathsf{NBG} \cup CA[\Pi^1_n] \cup Col[\Pi^1_n] \vdash AC[\Pi^1_n]$.

Corollary 39.

For $\mathcal{F}_0, \mathcal{F}_1 \in \{\Pi^1_n, \Sigma^1_{n+1}\}$ we have that

$$\mathsf{NBG} \cup AC[\mathcal{F}_0] \vdash A \quad \text{iff} \quad \mathsf{NBG} \cup CA[\Pi^1_n] \cup Col[\mathcal{F}_1] \vdash A.$$

We have formulated the choice principles AC and DC for sets and also for ordinals, i.e. $AC_{\mathcal{O}n}$ and $DC_{\mathcal{O}n}$. The following two theorems show that these two formulations are essentially the same over NBG.

Theorem 40.

For $\mathcal{F} \in \{\Pi^1_n, \Sigma^1_n\}$ we have that

(1) $\mathsf{NBG} \cup AC[\mathcal{F}] \vdash AC_{\mathcal{O}n}[\mathcal{F}]$,

(2) $\mathsf{NBG} \cup AC_{\mathcal{O}n}[\mathcal{F}] \vdash AC[\mathcal{F}]$.

Proof.

(1) If $A[u, V] \in \mathcal{F}$ then by Lemma 32 there is $B[u, V] \in \mathcal{F}$ such that

$$B[u, V] \leftrightarrow (u \in \mathcal{O}n \rightarrow A[u, V]),$$

further if $\forall \alpha \exists Y A[\alpha, Y]$ then $\forall x \exists Y B[x, Y]$, hence $\exists Y \forall x B[x, (Y)_x]$ by $AC[\mathcal{F}]$, that is $\exists Y \forall \alpha A[\alpha, (Y)_\alpha]$.

(2) If $W : \mathcal{O}n \rightarrow \mathcal{V}$ is such that $Gl[W^{-1}]$ by Lemma 29, and $A[u, V] \in \mathcal{F}$ then $B[u, V] := A[W(u), V] \in \mathcal{F}$, and if $\forall x \exists Y A[x, Y]$ then clearly $\forall \alpha \exists Y B[\alpha, Y]$, and there is Y such that $\forall \alpha B[\alpha, (Y)_\alpha]$ by $AC_{\mathcal{O}n}[\mathcal{F}]$, i.e. $\forall x A[x, (Y)_{W^{-1}(x)}]$, and for $Z := \{\langle z, x\rangle \mid z \in (Y)_{W^{-1}(x)}\}$ we have $\forall x A[x, (Z)_x]$. □

Theorem 41.

For $\mathcal{F} \in \{\Pi^1_n, \Sigma^1_n\}$ we have that

(1) $\mathsf{NBG} \cup DC[\mathcal{F}] \vdash DC_{\mathcal{O}n}[\mathcal{F}]$,

(2) $\mathsf{NBG} \cup DC_{\mathcal{O}n}[\mathcal{F}] \vdash DC[\mathcal{F}]$.

Proof.

(1) Analogous to Theorem 40(1).

(2) By recursion we get $\rho : \mathcal{V} \rightarrow \mathcal{O}n$, such that $\rho(x) := \bigcup\{\rho(y) + 1 \mid y \in x\}$. If $B[u, U, V] := A[u, \{\langle z, y\rangle \mid y \in u \wedge \langle\langle z, y\rangle, \rho(y)\rangle \in U\}, V]$ with $A \in \mathcal{F}$ then $B[u, U, V] \in \mathcal{F}$, and if $\forall x \forall X \exists Y A[x, X, Y]$ then

$$\forall \alpha \forall x \forall X \exists Y (\alpha = \rho(x) \rightarrow B[x, X, Y]).$$

We trivially have $\mathsf{NBG} \cup DC_{\mathcal{O}n}[\mathcal{F}] \vdash AC_{\mathcal{O}n}[\mathcal{F}]$, hence by Theorem 40 and Lemma 32, and by using $AC[\mathcal{F}]$, we get

$$\forall \alpha \forall X \exists Y \forall x (\alpha = \rho(x) \rightarrow B[x, X, (Y)_x]),$$

and by $DC_{\mathcal{O}n}[\mathcal{F}]$ there is a class Y such that

$$\forall \alpha \forall x (\alpha = \rho(x) \rightarrow B[x, (Y)^\alpha, ((Y)_\alpha)_x]).$$

If $Z := \{\langle z, y\rangle \mid \langle\langle z, y\rangle, \rho(y)\rangle \in Y\}$ and $\alpha = \rho(x)$ then we have that

$$\begin{aligned}
(Z)_x &= ((Y)_\alpha)_x, \\
(Z)^x &= \{\langle z, y\rangle \mid y \in x \wedge \langle\langle z, y\rangle, \rho(y)\rangle \in (Y)^\alpha\}.
\end{aligned}$$

Hence we get $\forall \alpha \forall x (\alpha = \rho(x) \rightarrow A[x, (Z)^x, (Z)_x])$, and this is equivalent to $\forall x (\exists \alpha (\alpha = \rho(x)) \rightarrow A[x, (Z)^x, (Z)_x])$, i.e. $\forall x A[x, (Z)^x, (Z)_x]$. □

The next lemma is an intermediate step for the proofs of the two theorems that follow. The lemma shows that choice, AC, is a (not so obvious) implication of some "weak" forms of dependent choice.

Lemma 42.

For $\mathcal{F} \in \{\Pi^1_n, \Sigma^1_n\}$ we have that
(1) $\mathsf{NBG} \cup DC^{\Diamond\text{-}}[\mathcal{F}] \vdash AC[\mathcal{F}]$,
(2) $\mathsf{NBG} \cup DC^{\Diamond\text{-}}_{\mathcal{O}n}[\mathcal{F}] \vdash AC_{\mathcal{O}n}[\mathcal{F}]$.

Proof.

(1) If $B[U, V] := \forall x((U)_0 = x \to (V)_0 = \{x\} \wedge A[x, (V)_1])$ with $A \in \mathcal{F}$ then $B[U, V]$ is equivalent to some formula in $\mathcal{F}$ by Lemma 32 and because

$$B[U, V] \leftrightarrow (\exists x(x = (U)_0) \to (\forall x \in (V)_0)(x = (U)_0) \wedge A[(U)_0, (V)_1]).$$

If $\forall x \exists Y A[x, Y]$ then $\forall X \exists Y B[X, Y]$ and by $DC^{\Diamond\text{-}}[\mathcal{F}]$ there is Y such that

$$\forall y \forall x((\langle Y\rangle^y)_0 = x \to ((Y)_y)_0 = \{x\} \wedge A[x, ((Y)_y)_1]).$$

By induction on x we get $(\langle Y\rangle^x)_0 = x$, because $(\langle Y\rangle^\emptyset)_0 = \emptyset$ and for $x \neq \emptyset$ and $z \in x$ we have $((Y)_z)_0 = \{z\}$ by i.h., hence $(\langle Y\rangle^x)_0 = x$. We have $\forall x A[x, ((Y)_x)_1])$, and for $Z := \{\langle z, x\rangle \mid z \in ((Y)_x)_1\}$ we get $\forall x A[x, (Z)_x])$.

(2) Analogous to Part 1. □

The following two theorems together with Theorem 41 yields that all the variants of the dependent choice schemata are equivalent over NBG.

Theorem 43.

For $\mathcal{F} \in \{\Pi^1_n, \Sigma^1_n\}$ we have that
(1) $\mathsf{NBG} \cup DC[\mathcal{F}] \vdash DC^{\text{-}}[\mathcal{F}]$,
(2) $\mathsf{NBG} \cup DC^{\text{-}}[\mathcal{F}] \vdash DC^{\Diamond\text{-}}[\mathcal{F}]$,
(3) $\mathsf{NBG} \cup DC^{\Diamond\text{-}}[\mathcal{F}] \vdash DC^{\Diamond}[\mathcal{F}]$,
(4) $\mathsf{NBG} \cup DC^{\Diamond}[\mathcal{F}] \vdash DC[\mathcal{F}]$.

Proof.

(1) Trivial.

(2) If $B[U, V] := A[\{y \mid \exists z\langle y, z\rangle \in U\}, V]$ with $A \in \mathcal{F}$ then $B[U, V] \in \mathcal{F}$. and if $\forall X \exists Y A[X, Y]$ then $\forall X \exists Y B[X, Y]$, and by $DC^{\text{-}}[\mathcal{F}]$ there is some Y such that $\forall x B[(Y)^x, (Y)_x]$. We have that

$$\{y \mid \exists z\langle y, z\rangle \in (Y)^x\} = \{y \mid (\exists z \in x)\langle y, z\rangle \in Y\} = \langle Y\rangle^x,$$

hence $\forall x A[\langle Y\rangle^x, (Y)_x]$.

(3) If $B[U,V] := \forall x A[x, \langle U\rangle^x, (V)_x]$ with $A \in \mathcal{F}$ then B is equivalent to some formula in $\mathcal{F}$ by Lemma 42 and Lemma 33. If $\forall x\forall X\exists Y A[x,X,Y]$ then $\forall x\forall X\exists Y A[x, \langle X\rangle^x, Y]$, and we get $\forall X\exists Y\forall x A[x, \langle X\rangle^x, (Y)_x]$ by Lemma 42 and $AC[\mathcal{F}]$, hence $\forall X\exists Y B[X,Y]$. By $DC^{\diamondsuit\text{-}}[\mathcal{F}]$ we get some Y such that $\forall z B[\langle Y\rangle^z, (Y)_z]$, i.e. $\forall z\forall x A[x, \langle\langle Y\rangle^z\rangle^x, ((Y)_z)_x]$. If $Z := \{\langle y,x\rangle \mid \langle\langle y,x\rangle,x\rangle \in Y\}$ then we have that

$$(Z)_x = ((Y)_x)_x \text{ and } \langle Z\rangle^x = \langle\langle Y\rangle^x\rangle^x,$$

hence $\forall x A[x, \langle Z\rangle^x, (Z)_x]$.

(4) If $B[x,U,V] := A[x,U,(V)_x] \wedge ((U)^x = U \to (\forall z \in V)\exists y\langle y,x\rangle = z)$ with $A \in \mathcal{F}$ then B is equivalent to some formula in $\mathcal{F}$ by Lemma 32. If $\forall x\forall X\exists Y A[x,X,Y]$ then $\forall x\forall X\exists Y B[x,X,Y]$, and by $DC^{\diamondsuit}[\mathcal{F}]$ we get some Y such that $\forall x B[x, \langle Y\rangle^x, (Y)_x]$, hence we have that $\forall x A[x, \langle Y\rangle^x, ((Y)_x)_x]$ and $\forall x((\langle Y\rangle^x)^x = \langle Y\rangle^x \to (\forall z \in (Y)_x)\exists y\langle y,x\rangle = z)$. By induction on x we get $(\forall z \in (Y)_x)\exists y\langle y,x\rangle = z$, and if $Z := \{\langle y,x\rangle \mid \langle\langle y,x\rangle,x\rangle \in Y\}$ then

$$(Z)_x = ((Y)_x)_x \text{ and } (Z)^x = \langle Y\rangle^x,$$

hence $\forall x A[x, (Z)^x, (Z)_x]$. □

Theorem 44.

For $\mathcal{F} \in \{\Pi^1_n, \Sigma^1_n\}$ we have that

1. $\mathsf{NBG} \cup DC_{\mathcal{O}n}[\mathcal{F}] \vdash DC^{\text{-}}_{\mathcal{O}n}[\mathcal{F}]$,
2. $\mathsf{NBG} \cup DC^{\text{-}}_{\mathcal{O}n}[\mathcal{F}] \vdash DC^{\diamondsuit\text{-}}_{\mathcal{O}n}[\mathcal{F}]$,
3. $\mathsf{NBG} \cup DC^{\diamondsuit\text{-}}_{\mathcal{O}n}[\mathcal{F}] \vdash DC^{\diamondsuit}_{\mathcal{O}n}[\mathcal{F}]$,
4. $\mathsf{NBG} \cup DC^{\diamondsuit}_{\mathcal{O}n}[\mathcal{F}] \vdash DC_{\mathcal{O}n}[\mathcal{F}]$.

Proof. Analogous to Theorem 43. □

We can state the following corollary by collecting together all the results about choice principles we got so far.

Corollary 45.

For $\mathcal{C}_0, \mathcal{C}_1 \in \{\Pi^1_n, \Sigma^1_{n+1}\}$ and

$$\begin{array}{rcl} \mathcal{T}_i & \in & \left\{ \begin{array}{c} DC[\mathcal{C}_i], DC^{\text{-}}[\mathcal{C}_i], DC^{\diamondsuit\text{-}}[\mathcal{C}_i], DC^{\diamondsuit}[\mathcal{C}_i], \\ DC_{\mathcal{O}n}[\mathcal{C}_i], DC^{\text{-}}_{\mathcal{O}n}[\mathcal{C}_i], DC^{\diamondsuit\text{-}}_{\mathcal{O}n}[\mathcal{C}_i], DC^{\diamondsuit}_{\mathcal{O}n}[\mathcal{C}_i] \end{array} \right\}, \\ \mathcal{F}_i & \in & \{AC[\mathcal{C}_i], AC_{\mathcal{O}n}[\mathcal{C}_i]\}, \end{array}$$

we have that

1. $\mathsf{NBG} \cup \mathcal{T}_0 \vdash \mathcal{T}_1 \cup \mathcal{F}_1$,
2. $\mathsf{NBG} \cup \mathcal{F}_0 \vdash \mathcal{F}_1 \cup CA[\Pi^1_n \cup \Sigma^1_n]$.

4. Well-founded Class Relations

Wellfoundedness and transfinite induction are two well-known and closely related concepts, see e.g. Forster [5]. In this section we give a general definition of the two concepts within theories of sets and classes, and we prove some familiar statements about wellfoundedness of the element relation $\in$ in NBG. Further, the definitions are such that we can easily state some theorems about the exact relationship of the two notions.

Definition 46. (Wellfoundedness)

Y is *well-founded* on Z for $C[x]$ in $\mathcal{T}$, if Y, Z are definable in $\mathcal{T}$ by A, B, respectively, and

$$\mathcal{T} \vdash \mathit{Def}[Y, A] \wedge \mathit{Def}[Z, B] \to \mathit{Wf}_Y^Z[C[x]]$$

where $\mathit{Wf}_Y^Z[C[x]]$ is the formula

$$(\exists x \in Z)C[x] \to (\exists x \in Z)(C[x] \wedge (\forall y \in Z)(C[y] \to \langle y, x\rangle \notin Y)),$$

i.e. $\mathcal{T} \vdash \mathit{Wf}_A^B[C[x]]$ where $\mathit{Wf}_A^B[C[x]]$ is

$$\exists x(B[x] \wedge C[x]) \to \exists x(B[x] \wedge C[x] \wedge \forall y(B[y] \wedge C[y] \to \neg A[\langle y, x\rangle])).$$

If $\mathcal{D} \subseteq \mathcal{L}^1$ then we write $\mathit{Wf}_A^B[\mathcal{D}]$ for $\{\mathit{Wf}_A^B[C[x]] \in \mathcal{L}^1 \mid C \in \mathcal{D}\}$.

In ZFC set theory, i.e. by the axiom of foundation, we have that $\in$ is well-founded on sets. In NBG we get that $\in$ is well-founded on all classes, by using the axiom of global choice. Further, because of class comprehension in NBG, we have that $\in$ is wellfounded for all formulas in Π^1_0.

Theorem 47. (Class Foundation)

If A and B are the formulas $\forall Z(Z \neq \emptyset \to (\exists x \in Z)(\forall y \in Z)\ \langle y, x\rangle \notin Y)$ and $\forall z(z \neq \emptyset \to (\exists x \in z)(\forall y \in z)\ \langle y, x\rangle \notin Y)$, respectively, then

$$\mathsf{NBG} \vdash A \leftrightarrow B.$$

Proof. We have $A \to B$ because of $\mathsf{NBG} \vdash \forall z \exists X(z = X)$. And we show the contrapositive of $B \to A$. Let $Z \neq \emptyset$ such that $(\forall x \in Z)(\exists y \in Z)\langle y, x\rangle \in Y$. We define the elementary class G such that

$$G := \{f \mid \mathit{Fun}[f] \wedge \mathit{dom}(f) = \omega \wedge f[\omega] \subseteq Z \wedge \forall p \langle f(p'), f(p)\rangle \in Y\}.$$

For any $f \in G$ and $z = f[\omega] \neq \emptyset$ we have that $(\forall x \in z)(\exists y \in z)\langle y, x\rangle \in Y$. We need to show $G \neq \emptyset$. Let W be a global wellordering (see Definition 28, Lemma 29) and let $C[x, y]$ be an elementary formula expressing that "$x, y \in Z$, and $\langle y, x\rangle \in Y$, and y is the least such set with respect to W". For any $x_0 \in Z$ and for the elementary class F with

$$F := \{\langle p, f\rangle \mid \mathit{Fun}[f] \wedge \mathit{dom}(f) = p' \wedge f(\emptyset) = x_0 \wedge (\forall q \in p)C[f(q'), f(q)]\},$$

we have that $\mathit{dom}(F) = \omega \wedge \mathit{Fun}[F] \wedge \forall p(F(p) \subseteq F(p'))$ and $\cup(F[\omega]) \in G$. □

Corollary 48. (Π^1_0 Foundation)

(1) $\mathsf{NBG} \vdash \forall z(\mathit{Wf}^{\mathcal{V}}_{\in}[x \in z]) \leftrightarrow \forall Z(\mathit{Wf}^{\mathcal{V}}_{\in}[x \in Z])$,

(2) $\mathsf{NBG} \vdash \mathit{Wf}^{\mathcal{V}}_{\in}[\Pi^1_0]$,

(3) $\mathsf{NBG} \vdash \mathit{Wf}^{X}_{\in}[\Pi^1_0]$.

Proof.

(1) For A, B as in Theorem 47 (and $\in$ as in Lemma 26) we have that

$$\mathsf{NBG} \vdash A[\in/Y] \leftrightarrow \forall Z(\mathit{Wf}^{\mathcal{V}}_{\in}[x \in Z]),$$
$$\mathsf{NBG} \vdash B[\in/Y] \leftrightarrow \forall z(\mathit{Wf}^{\mathcal{V}}_{\in}[x \in z]).$$

(2) We have $\mathsf{NBG} \vdash \forall z(\mathit{Wf}^{\mathcal{V}}_{\in}[x \in z])$ because of the axiom of foundation, and we have $\mathsf{NBG} \cup \{\forall Z(\mathit{Wf}^{\mathcal{V}}_{\in}[x \in Z])\} \vdash \mathit{Wf}^{\mathcal{V}}_{\in}[\Pi^1_0]$ because of the comprehension axiom.

(3) For $C \in \Pi^1_0$ we have $\mathsf{NBG} \vdash \mathit{Wf}^{\mathcal{V}}_{\in}[y \in X \wedge C[y]]$ by (2). Hence the claim holds because of $\mathit{Wf}^{\mathcal{V}}_{\in}[y \in X \wedge C[y]] \to \mathit{Wf}^{X}_{\in}[C[y]]$. □

We define the schemata for transfinite induction analogously to the schemata for wellfoundedness. The subsequent lemma shows that transfinite induction is the dual of wellfoundedness, and that under appropriate assumptions the two concepts become essentially the same.

Definition 49. (Transfinite Induction)

Transfinite induction along Y on Z for $C[x]$ in $\mathcal{T}$ holds, if Y, Z are definable in $\mathcal{T}$ by A, B, respectively, and

$$\mathcal{T} \vdash \mathit{Def}[Y, A] \wedge \mathit{Def}[Z, B] \to \mathit{TI}_Y^Z[C[x]]$$

where $\mathit{TI}_Y^Z[C[x]]$ is the formula

$$(\forall x \in Z)((\forall y \in Z)(\langle y, x\rangle \in Y \to C[y]) \to C[x]) \to (\forall x \in Z)C[x],$$

i.e. $\mathcal{T} \vdash \mathit{TI}_A^B[C[x]]$ where

$$\mathit{TI}_A^B[C[x]] := \mathit{Prog}_A^B[C[x]] \to \forall x(B[x] \to C[x]),$$
$$\mathit{Prog}_A^B[C[x]] := \forall x(B[x] \wedge \forall y(B[y] \wedge A[\langle y, x\rangle] \to C[y]) \to C[x]).$$

If $\mathcal{D} \subseteq \mathcal{L}^1$ then we write $\mathit{TI}_A^B[\mathcal{D}]$ for $\{\mathit{TI}_A^B[C[x]] \in \mathcal{L}^1 \mid C \in \mathcal{D}\}$, and for technical reasons we further define $\mathit{TI}_\in[\mathcal{D}] \subseteq \mathcal{L}^1$ to be the set containing all formulas $\forall x((\forall y \in x)A[y] \to A[x]) \to \forall x A[x]$ for every $A \in \mathcal{D}$.

Clearly we have $\mathit{TI}_\in[\mathcal{D}] \neq \mathit{TI}_\in^{\mathcal{V}}[\mathcal{D}]$, but the two sets are semantically equivalent, i.e., $\mathit{TI}_\in[\mathcal{D}] \vdash \mathit{TI}_\in^{\mathcal{V}}[\mathcal{D}]$ and $\mathit{TI}_\in^{\mathcal{V}}[\mathcal{D}] \vdash \mathit{TI}_\in[\mathcal{D}]$.

Lemma 50. (Duality)

(1) $\vdash \mathit{Wf}_A^B[\neg C[x]] \leftrightarrow \mathit{TI}_A^B[C[x]]$,

(2) If $\neg\mathcal{D} \subseteq \mathcal{D} \subseteq \mathcal{L}^1$ then $\mathcal{T} \vdash \mathit{Wf}_A^B[\mathcal{D}]$ iff $\mathcal{T} \vdash \mathit{TI}_A^B[\mathcal{D}]$,

(3) If $C := \forall Y \exists Z \forall x(x \in Z \leftrightarrow x \notin Y)$ then

$$\vdash C \to (\forall Y(\mathit{Wf}_A^B[x \in Y]) \leftrightarrow \forall Y(\mathit{TI}_A^B[x \in Y])).$$

Proof.

(1) The contrapositive of $\mathit{Wf}_Y^Z[\neg C[x]]$ is

$$\neg(\exists x \in Z)(\neg C[x] \wedge (\forall y \in Z)(\neg C[y] \to \langle y, x\rangle \notin Y)) \to \neg(\exists x \in Z)\neg C[x]$$

and this formula is logically equivalent to

$$(\forall x \in Z)((\forall y \in Z)(\langle y, x\rangle \in Y \to C[y]) \to C[x]) \to (\forall x \in Z)C[x],$$

i.e. equivalent to $\mathit{TI}_Y^Z[C[x]]$.

(2) Follows directly from (1).

(3) Assume $\forall Y(\mathit{Wf}_A^B[x \in Y])$ and let U be any class. According to $\mathcal{T}$ there is a class Z such that $x \in U \leftrightarrow x \notin Z$. We have $\mathit{Wf}_A^B[x \in Z]$, hence $\mathit{TI}_A^B[x \notin Z]$, i.e. $\mathit{TI}_A^B[x \in U]$, by Part 1. □

5. Notation System for Ordinals

In this section we define a notation system for ordinals, which is later used in a generic way in Section 6 to define the well-founded linear ordering $(E_0, \lhd)$, going far beyond the ordinals. In Jäger and Krähenbühl [10], a very similar notation system for E_0 has already been defined without detour over notations for ordinals; the aim of this section is to emphasize the tight connection between the notation system for ordinals and E_0, e.g., the arithmetic operations defined on the ordinal notations can later be easily lifted to the notations in E_0 (again in a generic way). That is, we can get a good understanding of addition, multiplication, and exponentiation, in E_0, by just coping with common ordinal arithmetic (i.e., by coping with the operations on ordinals in Cantor normal form).

By a notation system we mean that each ordinal can be represented in a uniform way by an object which essentially consists only of finitely many basic symbols, i.e. in our system the notations consist of hereditarily finite sequences of basic symbols. In addition to the finite presentabiltity of the notations, the notation system must be such that the basic operations on the ordinals, e.g. ordinal arithmetic, are also achieved by uniform operations on the notations.

The definition of the notation system is inspired by the Cantor normal form of ordinals, and it is a slight generalisation of the standard notation system for the ordinal ε_0, as e.g. in Pohlers [15] or Schütte [16]. (In Appendix C we also use a construction very similar to the one in this section, to define a notation system based on the binary Veblen function, leading to a generalisation of the standard notation system for the ordinal Γ_0, as e.g. in Pohlers [15].)

First, we define the notation system by a top-down approach, that is, we use the ordinals themselves, and the Cantor normal form, to recursively define the notation $\widehat{\alpha}$ for each ordinal α, such that

$$\widehat{\alpha} \quad := \quad \begin{cases} \alpha & \alpha = \omega^{\alpha} \text{ or } \alpha = \emptyset, \\ \langle\!\langle \widehat{\gamma_0}, ..., \widehat{\gamma_r} \rangle\!\rangle & \alpha =_{\mathsf{CNF}} \omega^{\gamma_0} + ... + \omega^{\gamma_r} \neq \omega^{\alpha}. \end{cases}$$

Clearly, we need a proper class of basic symbols to represent all the ordinals, and as we can see, all ordinals α with $\alpha = \omega^{\alpha}$ and $\emptyset$ are the basic symbols.

Secondly, and more important, we define the same notations by a bottom-up approach, that is we give an inductive definition of the notations by just using hereditarily finite sequences and the basic symbols, without recourse to the ordinals. This definition will be given in a generic way (i.e., with two class parameters involved), such that it is easily adapted to other classes of basic symbols, e.g. leading to notation systems that go beyond the ordinals, like the notations in E_0, in Section 6. In addition to the notations we also define operations on these notations, i.e. arithmetical operations and an ordering relation, which correspond to the original arithmetic operations on ordinals and the element relation $\in$ on ordinals.

In the following definitions we *explicitly* show, i.e., by just using *elementary* formulas, that the ordinal notation system, and the operations and the ordering relation on notations, are elementarily definable (in both cases, the top-down and the bottom-up approach). The definability of the notation system by elementary formulas is important later on, because we are going to give proofs by induction for statments where these notations are involved; e.g. in NBG, where we have induction for elementary formulas only.

E.g. addition $\alpha+\beta$ can be defined by $\alpha+\beta := F_\alpha(\beta)$ where the class function $F_\alpha : \mathit{On} \to \mathit{On}$ is such that $F_\alpha(\emptyset) := \alpha$ and $F_\alpha(\beta) := \bigcup_{\gamma<\beta} F_\alpha(\gamma)'$ for $\beta \neq \emptyset$. Analogously, multiplication $\alpha \cdot \beta$ is defined by $\alpha \cdot \beta := G_\alpha(\beta)$ where $G_\alpha(\beta) := \bigcup_{\gamma<\beta}(G_\alpha(\gamma)+\alpha)$, and exponentiation α^β is such that $\alpha^\beta := H_\alpha(\beta)$ where $H_\alpha(\emptyset) := \overline{1}$ and $H_\alpha(\beta) := \bigcup_{\gamma<\beta}(H_\alpha(\gamma) \cdot \alpha)$ for $\beta \neq \emptyset$. By using the class functions F_α, G_α, and H_α, we can explain the meaning of the elementary formulas *Add*, *Mult*, and *Exp*, in the following definition, that is, $\mathit{Add}[f,\alpha]$ is equivalent to $\exists\beta(\emptyset \in \beta \wedge f = F_\alpha \restriction \beta)$, and analogously for *Mult* and *Exp*. Further, the general sum $\Sigma_\alpha f$ is defined such that $\Sigma_\alpha f := \bigcup_{\gamma<\alpha}(\Sigma_\gamma f + f(\gamma))$; the formula $\mathit{Sum}[f,g,\alpha]$, which is used for the definition of $\Sigma_\alpha f$, means that the function g is such that $g(\emptyset) = f(\emptyset)$, and for any γ with $\emptyset < \gamma < \alpha$ we have $g(\gamma) = \bigcup_{\beta<\gamma} g(\beta) + f(\beta')$, that is, $\Sigma_\alpha f = \bigcup_{\gamma<\alpha} g(\gamma)$.

Definition 51. (Ordinal Arithmetic)

We define the expressions $\alpha+\beta$, $\alpha \cdot \beta$, and α^β, (analogous Definition 18) such that

$$\begin{aligned}
z \in \alpha+\beta &:= \exists f(\mathit{Add}[f,\alpha] \wedge \beta \in \mathit{dom}(f) \wedge z \in f(\beta)),\\
z \in \alpha\cdot\beta &:= \exists f(\mathit{Mult}[f,\alpha] \wedge \beta \in \mathit{dom}(f) \wedge z \in f(\beta)),\\
z \in \alpha^\beta &:= \exists f(\mathit{Exp}[f,\alpha] \wedge \beta \in \mathit{dom}(f) \wedge z \in f(\beta)),\\
z \in \cup_\alpha f &:= (\exists\beta \in \alpha) z \in f(\beta),\\
z \in \Sigma_\alpha f &:= \exists g(\mathit{Sum}[f,g,\alpha] \wedge z \in \cup_\alpha g),
\end{aligned}$$

$$
\begin{array}{rcl}
Add[f,\alpha] & := & Fun[f] \wedge \emptyset \in dom(f) \wedge f(\emptyset) = \alpha \wedge \\
& & \exists\beta\exists\gamma(dom(f) = \beta \wedge ran(f) \subseteq \gamma \wedge (\forall\beta_1 \in \beta\backslash\{\emptyset\})(\\
& & \forall x(x \in f(\beta_1) \leftrightarrow (\exists\beta_0 \in \beta_1)x \in f(\beta_0)'))),
\end{array}
$$

$$
\begin{array}{rcl}
Mult[f,\alpha] & := & Fun[f] \wedge \emptyset \in dom(f) \wedge f(\emptyset) = \emptyset \wedge \\
& & \exists\beta\exists\gamma(dom(f) = \beta \wedge ran(f) \subseteq \gamma \wedge (\forall\beta_1 \in \beta\backslash\{\emptyset\})(\\
& & \forall x(x \in f(\beta_1) \leftrightarrow (\exists\beta_0 \in \beta_1)x \in f(\beta_0) + \alpha))),
\end{array}
$$

$$
\begin{array}{rcl}
Exp[f,\alpha] & := & Fun[f] \wedge \emptyset \in dom(f) \wedge f(\emptyset) = \overline{1} \wedge \\
& & \exists\beta\exists\gamma(dom(f) = \beta \wedge ran(f) \subseteq \gamma \wedge (\forall\beta_1 \in \beta\backslash\{\emptyset\})(\\
& & \forall x(x \in f(\beta_1) \leftrightarrow (\exists\beta_0 \in \beta_1)x \in f(\beta_0) \cdot \alpha))),
\end{array}
$$

$$
\begin{array}{rcl}
Sum[f,g,\alpha] & := & Fun[f] \wedge Fun[g] \wedge \emptyset \in \alpha \wedge \alpha \subseteq dom(f) \cap dom(g) \wedge \\
& & \exists\beta(ran(f) \subseteq \beta) \wedge g(\emptyset) = f(\emptyset) \wedge (\forall\alpha_1 \in \alpha\backslash\{\emptyset\})(\\
& & \forall x(x \in g(\alpha_1) \leftrightarrow (\exists\alpha_0 \in \alpha_1)x \in g(\alpha_0) + f({\alpha_0}'))).
\end{array}
$$

(See e.g. Takeuti [18] for the properties of $\alpha + \beta$, $\alpha \cdot \beta$, and α^β.)

Lemma 52. (Ordinal Arithmetic)

(1) $\mathsf{NBG} \vdash \forall\alpha\forall\beta(\alpha + \emptyset = \alpha \wedge \alpha + \beta' = (\alpha + \beta)' \wedge$
$\alpha \cdot \emptyset = \emptyset \wedge \alpha \cdot \beta' = (\alpha \cdot \beta) + \alpha \wedge$
$\alpha^\emptyset = \overline{1} \wedge \alpha^{\beta'} = \alpha^\beta \cdot \alpha)$,

(2) $\mathsf{NBG} \vdash \forall\alpha\forall\beta\forall f(Fun[f] \wedge \beta \subseteq dom(f) \wedge \beta \neq \emptyset \rightarrow$
$((\forall\gamma \in \beta)(f(\gamma) = \alpha + \gamma') \rightarrow \cup_\beta f = \alpha + \beta) \wedge$
$((\forall\gamma \in \beta)(f(\gamma) = \alpha \cdot \gamma') \rightarrow \cup_\beta f = \alpha \cdot \beta) \wedge$
$((\forall\gamma \in \beta)(f(\gamma) = \alpha^{\gamma'}) \rightarrow \cup_\beta f = \alpha^\beta)$.

Proof. By Definition (i.e. by induction on the ordinals). □

The following definition (see Takeuti [18]) is about the least infinite ordinal ω, i.e., the least nonempty ordinal not being a successor of any ordinal.

Definition 53.

We define the expression ω (analogous Definition 18) such that

$$z \in \omega \quad := \quad \forall x(x \in z' \rightarrow x = \emptyset \vee \exists\alpha(x = \alpha')).$$

(See e.g. Takeuti [18] for the properties of ω.)

We use the letters p, q, r, s, t (with subscripts) to denote ordinals in ω. We write $\forall pA[p]$ and $\exists pA[p]$ for $(\forall x \in \omega)A[x]$ and $(\exists x \in \omega)A[x]$, respectively.

To be able to deal with the Cantor normal form in the language $\mathcal{L}^0$, we define the formula *CNF*, such that $CNF[f,p,\alpha]$ holds if the sequence f of ordinals, $\gamma_1, ..., \gamma_p$, is such that $\alpha = \omega^{\gamma_1} + ... + \omega^{\gamma_p}$ with $\gamma_1 \geq ... \geq \gamma_p$. The subsequent theorem is the Cantor normal form theorem, which states that this normal form exists for any ordinal $\alpha \neq \emptyset$ and really is unique.

Definition 54. (Cantor Normal Form)

$$\begin{array}{lll} CNF[f,p,\alpha] & := & Fun[f] \wedge dom(f) = p \wedge \exists h(\forall p_1 \in p)(\\ & & \exists\beta(f(p_1) = \beta \wedge h(p_1) = \omega^\beta) \wedge \\ & & (\forall p_0 \in p_1)h(p_1) \in h(p_0)' \wedge \alpha = \Sigma_p h). \end{array}$$

Theorem 55. (Cantor Normal Form)

(1) $\mathsf{NBG} \vdash \forall\alpha(\alpha = \emptyset \vee \exists! f \exists p CNF[f,p,\alpha])$,

(2) $\mathsf{NBG} \vdash \forall\alpha\forall f\forall p(CNF[f,p,\alpha] \rightarrow \alpha = \emptyset \vee \alpha = \omega^\alpha \vee f(\emptyset) \in \alpha)$.

Proof. See e.g. Pohlers [15] or Jech [12]. □

Having at hand all the ingredients for the formal definition of the notation system, i.e., ordinal arithmetic and the formula *CNF*, we are now ready to state the *elementary* formula which characterizes the notations $\widehat{\alpha}$ as we described it at the beginning of this section. And we finish the first part of this section, i.e. the top-down approach to the notations, by stating the subsequent lemma about the defining recursion of the notation system.

The formula OT_ε in the following definition is such that $OT_\varepsilon[f]$ holds if the domain of f is some ordinal α and $f(\gamma) = \widehat{\gamma}$ for all $\gamma \in \alpha$.

Definition 56. (Ordinal Notation System)

We define the expression $\widehat{\alpha}$ (analogous Definition 18) such that

$$z \in \widehat{\alpha} \quad := \quad \exists f(OT_\varepsilon[f] \wedge \alpha \in dom(f) \wedge z \in f(\alpha)),$$

$$\begin{array}{lll} OT_\varepsilon[f] & := & Fun[f] \wedge \exists\alpha(dom(f) = \alpha \wedge (\forall\alpha_0 \in \alpha)(\\ & & (\alpha_0 = \emptyset \wedge f(\alpha_0) = \alpha_0) \vee (\alpha_0 = \omega^{\alpha_0} \wedge f(\alpha_0) = \alpha_0) \vee \\ & & (\alpha_0 \neq \emptyset \wedge \alpha_0 \neq \omega^{\alpha_0} \wedge Fun[f(\alpha_0)] \wedge \exists g\exists p(CNF[g,p,\alpha_0] \wedge \\ & & dom(f(\alpha_0)) = p \wedge (\forall p_0 \in p) f(\alpha_0)(p_0) = f(g(p_0)))). \end{array}$$

We simply write $\widehat{n}$ for $\widehat{\widehat{n}}$.

Lemma 57. (Ordinal Notation System)

$$\begin{array}{r} \mathsf{NBG} \vdash \forall\alpha\forall f\forall p(\alpha \neq \emptyset \wedge \alpha \neq \omega^\alpha \wedge CNF[f,p,\alpha] \rightarrow \qquad\qquad \\ Fun[\widehat{\alpha}] \wedge dom(\widehat{\alpha}) = p \wedge (\forall p_1 \in p)\widehat{\alpha}(p_1) = \widehat{f(p_1)}). \end{array}$$

Proof. By Definition (i.e. by induction on the ordinals). □

The formula *Lin* in the following definition is such that $Lin[U, V]$ means that the relation V is a strict linear ordering on U.

Definition 58. (Strict Total Order)

$$\begin{aligned} Lin[U,V] \quad :=& \quad (\forall x \in U)(\forall y \in U)(\forall z \in U)\langle x,x\rangle \notin V \wedge \\ & \quad (\langle x,y\rangle \in V \wedge \langle y,z\rangle \in V \to \langle x,z\rangle \in V) \wedge \\ & \quad (\langle x,y\rangle \in V \vee x = y \vee \langle y,x\rangle \in V). \end{aligned}$$

The following definitions are heading towards the inductive definition of our generic notation system, based on hereditarily finite sequences only, i.e. the bottom-up approach to the notation.

We first define a somewhat cryptic ordering relation $\widehat{Y}_X$ on the hereditarily finite sequences $\mathcal{H}_X$ of basic symbols in X, such that any strict linear ordering relation Y on X is extended to the strict linear ordering $\widehat{Y}_X$ on $\mathcal{H}_X \cup X$. The inspiration for $\widehat{Y}_X$ comes from the inherited ordering relation on the notations $\widehat{\alpha}$, corresponding to the ordering of the ordinals, i.e. the definition of $\widehat{Y}_X$ is similar to the definition of the ordering relation on the standard notation system for ε_0, see e.g. Pohlers [15].

The formula *Hed* in the following definition is such that $Hed[X, y]$ holds if the set y consists of hereditarily finite sequences, such that for any $f \in y$ we have $f(p) \in y \cup X$ for all $p \in dom(f)$. By induction on the set-theoretic rank we can easily see that y consists of nothing but hereditarily finite sequences, because the rank of $f(p)$ is smaller than the rank of f for all $f, f(p) \in y$.

Definition 59. (Hereditarily finite Sequences)

We define the expression $\mathcal{H}_X$ (analogous Definition 18) such that

$$z \in \mathcal{H}_X \quad := \quad \exists y(Hed[X,y] \wedge z \in y),$$

$$\begin{aligned} Hed[X,y] \quad :=& \quad (\forall f \in y)(Fun[f] \wedge \exists p(dom(f) = p \wedge \\ & \quad (\forall p_0 \in p)f(p_0) \in y \cup X)). \end{aligned}$$

The previous definition can be easily translated into an inductive definition of the class $\mathcal{H}_X$ (induction on the set-theoretic rank), i.e., $\emptyset \in \mathcal{H}_X$, and every function f with finite domain p, such that $f(p_0) \in \mathcal{H}_X \cup X$ for all $p_0 \in p$, is in $\mathcal{H}_X$.

The expression $\mathcal{H}_X$ is defined for arbitrary X, but in the following we always assume $\mathcal{H}_X \cap X = \emptyset$.

The formula *Ex* in the following definition is such that $Ex[X, Y, y]$ holds if $y \subseteq \widehat{Y}_X$ and if for any pair $\langle f, g\rangle \in y$ the set y contains all relevant pairs $\langle h_1, h_2\rangle \in \widehat{Y}_X$, which are used to decide whether $\langle f, g\rangle$ belongs to $\widehat{Y}_X$.

Definition 60. (Ordering Relation)

We define the expression $\widehat{Y}_X$ (analogous Definition 18) such that

$$z \in \widehat{Y}_X \quad := \quad \exists y(Ex[X, Y, y] \wedge z \in y),$$

$$\begin{aligned} Ex[X, Y, y] \quad := \quad & (\forall x \in y)\exists f \exists g(x = \langle f, g\rangle \wedge \{f, g\} \subseteq \mathcal{H}_X \cup X \wedge \\ & (f \in X \wedge g \in X \wedge \langle f, g\rangle \in Y) \vee \\ & (f \in X \wedge g \notin X \wedge g \neq \emptyset \wedge (\langle f, g(\emptyset)\rangle \in y \vee f = g(\emptyset)))\vee \\ & (f \notin X \wedge g \in X \wedge (f = \emptyset \vee \langle f(\emptyset), g\rangle \in y) \vee \\ & (f \notin X \wedge g \notin X \wedge Lex[y, f, g])), \end{aligned}$$

$$\begin{aligned} Lex[y, f, g] \quad := \quad & \exists p \exists q(dom(f) = p \wedge dom(g) = q \wedge \\ & ((p \in q \wedge (\forall p_0 \in p)f(p_0) = g(p_0)) \vee (\exists p_1 \in p \cap q)(\\ & \langle f(p_1), g(p_1)\rangle \in y \wedge (\forall p_0 \in p_1)f(p_0) = g(p_0)). \end{aligned}$$

The previous definition can be easily translated into an inductive definition of the binary relation $\widehat{Y}_X$, i.e., $\widehat{Y}_X$ consists of pairs $\langle f, g\rangle$ with $f, g \in \mathcal{H}_X \cup X$, such that one of the following is the case

$$\begin{aligned} & f \in X \wedge g \in X \wedge \langle f, g\rangle \in Y, \\ & f \in X \wedge g \notin X \wedge g \neq \emptyset \wedge (\langle f, g(\emptyset)\rangle \in \widehat{Y}_X \vee f = g(\emptyset)), \\ & f \notin X \wedge g \in X \wedge (f = \emptyset \vee \langle f(\emptyset), g\rangle \in \widehat{Y}_X, \\ & f \notin X \wedge g \notin X \wedge Lex[\widehat{Y}_X, f, g]. \end{aligned}$$

Lemma 61. (Strict Total Order)

$\mathsf{NBG} \vdash Lin[X, Y] \wedge \mathcal{H}_X \cap X = \emptyset \to Lin[\mathcal{H}_X \cup X, \widehat{Y}_X].$

Proof. The proof is similar to the purely combinatorial proof (i.e. without using the set theoretic background) for the standard notation system of ε_0 (cf. Remark 3.3.19 in Pohlers [15]). The most difficult part of the proof is transitivity, $\langle x, y\rangle \in \widehat{Y}_X \wedge \langle y, z\rangle \in \widehat{Y}_X \to \langle x, z\rangle \in \widehat{Y}_X$, and this is proved by induction on the sum of depths of x, y and z. The tedious technical details are left to the reader. $\square$

Now we use the relation $\widehat{Z}_X$, i.e. $Y := \widehat{Z}_X$, to define the class of all notations $\widehat{\mathcal{O}t}_{X,Y} \subseteq \mathcal{H}_X \cup X$ (actually, we define $\widehat{\mathcal{O}t}^0_{X,Y} \subseteq \mathcal{H}_X$ such that $\widehat{\mathcal{O}t}_{X,Y} = \widehat{\mathcal{O}t}^0_{X,Y} \cup X$). Further, to get the arithmetic operations on notations, we first define $\widehat{\mathcal{O}t}^1_{X,Y} := \widehat{\mathcal{O}t}^0_{X,Y} \cup \{\langle\!\langle w\rangle\!\rangle \mid w \in X\} \subseteq \mathcal{H}_X$ and arithmetic operations on $\widehat{\mathcal{O}t}^1_{X,Y}$, that is, $\hat{+}^1_Y$, $\hat{\cdot}^1_{X,Y}$, and $\hat{\wedge}^1_{X,Y}$ (see definitions below). After that, we use mappings $\pi_X : \widehat{\mathcal{O}t}_{X,Y} \to \widehat{\mathcal{O}t}^1_{X,Y}$ and $\pi^{inv}_X : \widehat{\mathcal{O}t}^1_{X,Y} \to \widehat{\mathcal{O}t}_{X,Y}$, to get the arithmetic operations on $\widehat{\mathcal{O}t}_{X,Y}$. At the end of this section, we will see that by putting $X := \{\alpha \mid \alpha = \omega^\alpha\}$, and $Z = \{\langle x, y\rangle \mid x \in y\}$, and $Y := \widehat{Z}_X$, we get the notations $\widehat{\alpha}$, that is, $\{\widehat{\alpha} \mid \alpha \in \mathcal{O}n\} = \widehat{\mathcal{O}t}_{X,Y}$. In the previous lemma, we have already seen that $\widehat{Z}_X$ is a strict linear ordering on $\widehat{\mathcal{O}t}_{X,Y}$ for any X and Z with $Lin[X, Z]$, hence the ordering relation on the notations will be $\widehat{Z}_X$.

The notation system $\widehat{\mathcal{O}t}_{X,Y}$ is defined similarly to the standard notation system for ε_0, see e.g. Pohlers [15].

Definition 62. (Generic Notation System)

We define the expressions $\widehat{\mathcal{O}t}^0_{X,Y}$, $\widehat{\mathcal{O}t}^1_{X,Y}$ $\widehat{\mathcal{O}t}_{X,Y}$, π_X, π^{inv}_X (analogous Definition 18) such that

$$\begin{array}{rcl}
z \in \widehat{\mathcal{O}t}^0_{X,Y} & := & \exists y(\hat{OT}[X, Y, y] \wedge z \in y), \\
z \in \widehat{\mathcal{O}t}^1_{X,Y} & := & z \in \widehat{\mathcal{O}t}^0_{X,Y} \vee (\exists w \in X) z = \langle\!\langle w \rangle\!\rangle, \\
z \in \widehat{\mathcal{O}t}_{X,Y} & := & z \in \widehat{\mathcal{O}t}^0_{X,Y} \vee z \in X, \\
z \in \pi_X & := & \exists y((y \in X \wedge z = \langle y, \langle\!\langle y \rangle\!\rangle\rangle) \vee (y \notin X \wedge z = \langle y, y\rangle)), \\
z \in \pi^{inv}_X & := & z \in \pi^{-1}_X \wedge (\forall y \in X) z \neq \langle \langle\!\langle y \rangle\!\rangle, \langle\!\langle y \rangle\!\rangle\rangle,
\end{array}$$

$$\begin{array}{rcl}
\hat{OT}[X, Y, y] & := & (\forall f \in y) f \in \mathcal{H}_X \wedge (\forall p \in dom(f))(f(p) \in y \cup X \wedge \\
& & (\forall q \in p)(f(q') = f(q) \vee \langle f(q'), f(q)\rangle \in Y)) \wedge \\
& & ((dom(f) \neq \overline{1} \vee f(\emptyset) \notin X)).
\end{array}$$

Again, the previous definition can be easily translated into a proper inductive definition of $\widehat{\mathcal{O}t}^0_{X,Y}$ (analogous to the preceding definitions).

There are purely combinatorial operations on the notations of ordinals in Cantor normal form, which represent the arithmetic operations on ordinals, see e.g. Takeuti [18] for a detailed account. E.g. if $\alpha =_{\mathsf{CNF}} \omega^\gamma$ and $\beta =_{\mathsf{CNF}} \omega^{\delta_0} + ... + \omega^{\delta_n}$ then

$$\alpha + \beta =_{\mathsf{CNF}} \begin{cases} \omega^{\delta_0} + ... + \omega^{\delta_n} & \gamma < \delta_0, \\ \omega^\gamma + \omega^{\delta_0} + ... + \omega^{\delta_n} & \gamma \geq \delta_0, \end{cases}$$

$$\beta \cdot \alpha =_{\mathsf{CNF}} \begin{cases} \omega^{\delta_0} + ... + \omega^{\delta_n} & \gamma = \emptyset, \\ \omega^{\delta_0 + \gamma} & \gamma > \emptyset, \end{cases}$$

$$\beta^\alpha =_{\mathsf{CNF}} \begin{cases} \omega^{\delta_0} + ... + \omega^{\delta_n} & \gamma = \emptyset, \\ \omega^{\delta_0 \cdot \alpha} & \gamma > \emptyset, \delta_0 > \emptyset, \\ \omega^\emptyset & \gamma > \emptyset, \delta_0 = \emptyset, n = 0, \\ \omega^{(\omega^{\gamma - 1})} & \omega > \gamma > \emptyset, \delta_0 = \emptyset, n > 0, \\ \omega^\alpha & \gamma \geq \omega, \delta_0 = \emptyset, n > 0, \end{cases}$$

and if α has a more complex form, i.e. $\alpha =_{\mathsf{CNF}} \omega^{\gamma_0} + ... + \omega^{\gamma_m}$, then we can use this operations, to find the Cantor normal form of $\omega^{\gamma_0} + (... + (\omega^{\gamma_m} + \beta))$, and $\beta \cdot \omega^{\gamma_0} + ... + \beta \cdot \omega^{\gamma_m}$, and $\beta^{(\omega^{\gamma_0})} \cdot ... \cdot \beta^{(\omega^{\gamma_m})}$, that is, the normal forms of $\alpha + \beta$, and $\beta \cdot \alpha$, and β^α, respectively.

In the following three definitions we capture theses operations for addition, multiplication and exponentiation in context of the generic notation system $\widehat{\mathcal{O}t}_{X,Z}$, which is also based on the Cantor normal form.

Definition 63. (Addition)

We define the expressions $f^\frown g$ and $f \mathbin{\hat{+}}_{X,Y} g$ (analogous Definition 18) such that

$$
\begin{array}{rcl}
z \in f^\frown g & := & z \in f \vee (\exists p \in dom(g)) z = \langle dom(f) + p, g(p) \rangle, \\
z \in x \mathbin{\hat{+}}^0_Y g & := & (\langle x, g(\emptyset)\rangle \in Y \wedge g \neq \emptyset \wedge z \in g) \vee \\
 & & (\langle x, g(\emptyset)\rangle \notin Y \vee g = \emptyset) \wedge z \in \langle\!\langle x \rangle\!\rangle^\frown g, \\
z \in f \mathbin{\hat{+}}^1_Y g & := & \exists p \exists h (p = dom(f) \wedge dom(h) = p' \wedge h(p) = g \wedge \\
 & & z \in h(\emptyset) \wedge (\forall q \in p) h(q) = f(q) \mathbin{\hat{+}}^0_Y h(q')), \\
z \in f \mathbin{\hat{+}}_{X,Y} g & := & z \in \pi^{inv}_X(\pi_X(f) \mathbin{\hat{+}}^1_Y \pi_X(g)).
\end{array}
$$

Definition 64. (Multiplication)

We define the expression $f \mathbin{\hat{\cdot}}_{X,Y} g$ (analogous Definition 18) such that

$$
\begin{array}{rcl}
z \in f \mathbin{\hat{\cdot}}^0_{X,Y} x & := & (f \neq \emptyset \wedge x = \emptyset \wedge z \in f) \vee \\
 & & (f \neq \emptyset \wedge x \neq \emptyset \wedge z \in \langle\!\langle f(\emptyset) \mathbin{\hat{+}}_{X,Y} x \rangle\!\rangle), \\
z \in f \mathbin{\hat{\cdot}}^1_{X,Y} g & := & \exists p \exists h (p = dom(g) \wedge dom(h) = p' \wedge h(\emptyset) = \emptyset \wedge \\
 & & z \in h(p) \wedge (\forall q \in p) h(q') = h(q) \mathbin{\hat{+}}^1_Y (f \mathbin{\hat{\cdot}}^0_{X,Y} g(q))), \\
z \in f \mathbin{\hat{\cdot}}_{X,Y} g & := & z \in \pi^{inv}_X(\pi_X(f) \mathbin{\hat{\cdot}}^1_{X,Y} \pi_X(g)).
\end{array}
$$

Definition 65. (Exponentiation)

We define the expression $f \mathbin{\hat{\wedge}}_{X,Y} x$ (analogous Definition 18) such that

$$
\begin{array}{rcl}
z \in f^- & := & (\langle \emptyset, \emptyset \rangle \in f \wedge \exists p (p' \in dom(f) \wedge z = \langle p, f(p') \rangle)) \vee \\
 & & (\langle \emptyset, \emptyset \rangle \notin f \wedge z \in f), \\
z \in f \mathbin{\hat{\wedge}}^0_{X,Y} x & := & (x = \emptyset \wedge z \in f) \vee (f = \emptyset \wedge z \in f) \vee \\
 & & (x \neq \emptyset \wedge f(\emptyset) = \emptyset \wedge ((dom(f) = \overline{1} \wedge z \in f) \vee \\
 & & \qquad (dom(f) \notin \overline{2} \wedge z \in \langle\!\langle \pi_X(\langle\!\langle x^- \rangle\!\rangle) \rangle\!\rangle)) \vee \\
 & & (x \neq \emptyset \wedge f(\emptyset) \neq \emptyset \wedge z \in \langle\!\langle f(\emptyset) \mathbin{\hat{\cdot}}_{X,Y} \langle\!\langle x \rangle\!\rangle \rangle\!\rangle), \\
z \in f \mathbin{\hat{\wedge}}^1_{X,Y} g & := & \exists p \exists h (p = dom(g) \wedge dom(h) = p' \wedge h(\emptyset) = \langle\!\langle \emptyset \rangle\!\rangle \wedge \\
 & & z \in h(p) \wedge (\forall q \in p) h(q') = h(q) \mathbin{\hat{\cdot}}^1_{X,Y} (f \mathbin{\hat{\wedge}}^0_{X,Y} g(q))), \\
z \in f \mathbin{\hat{\wedge}}_{X,Y} g & := & z \in \pi^{inv}_X(\pi_X(f) \mathbin{\hat{\wedge}}^1_{X,Y} \pi_X(g)),
\end{array}
$$

We end this section with the next theorem, which brings together the two approaches to our notation system for ordinals. We connect $\widehat{\mathcal{O}t}_{X,Z}$ to the notations $\widehat{\alpha}$, by taking the appropriate instance of the generic system $\widehat{\mathcal{O}t}_{X,Z}$, i.e. by putting $X := \{\alpha \mid \alpha = \omega^\alpha\}$ and $Z := \widehat{Y}_X$ where $Y := \{\langle x, y\rangle \mid x \in y\}$. The theorem also shows that the arithmetic operations and the ordering relation on $\widehat{\mathcal{O}t}_{X,Z}$ correspond to ordinal arithmetic and to the element relation on ordinals, respectively.

Theorem 66. (Ordinal Notation System)

Let $\widehat{\mathcal{O}n}$, $\mathcal{O}n_\varepsilon$ and $<_\varepsilon$ be elementarily definable classes in NBG, such that

$$\begin{array}{rcl} \widehat{\mathcal{O}n} & := & \{x \mid \exists\alpha(x = \widehat{\alpha})\}, \\ \mathcal{O}n_\varepsilon & := & \{x \mid \exists\alpha(x = \alpha \wedge \alpha = \omega^\alpha)\}, \\ <_\varepsilon & := & \widehat{\in}_{\mathcal{O}n_\varepsilon}, \end{array}$$

where $\widehat{\in}_{\mathcal{O}n_\varepsilon}$ is as in Definition 60 with $\in$ as in Lemma 26. If we write $+_\varepsilon$, $\cdot_\varepsilon$, $\wedge_\varepsilon$ for $\hat{+}_{\mathcal{O}n_\varepsilon,<_\varepsilon}$, $\hat{\cdot}_{\mathcal{O}n_\varepsilon,<_\varepsilon}$, $\hat{\wedge}_{\mathcal{O}n_\varepsilon,<_\varepsilon}$, respectively, then we have that

(1) $\mathsf{NBG} \vdash \forall\alpha\forall\beta(\alpha \in \beta \leftrightarrow \widehat{\alpha} <_\varepsilon \widehat{\beta})$,

(2) $\mathsf{NBG} \vdash \forall\alpha(\widehat{\alpha} \in \widehat{\mathcal{O}t}_{\mathcal{O}n_\varepsilon \cap \alpha', <_\varepsilon})$,

(3) $\mathsf{NBG} \vdash \widehat{\mathcal{O}n} = \widehat{\mathcal{O}t}_{\mathcal{O}n_\varepsilon, <_\varepsilon}$,

(4) $\mathsf{NBG} \vdash \forall\alpha\forall\beta(\widehat{\alpha + \beta} = \widehat{\alpha} +_\varepsilon \widehat{\beta} \wedge \widehat{\alpha \cdot \beta} = \widehat{\alpha} \cdot_\varepsilon \widehat{\beta} \wedge \widehat{\alpha^\beta} = \widehat{\alpha} \wedge_\varepsilon \widehat{\beta})$.

Proof. (1) is proved by induction on the natural sum (Hessenberg sum) of α and β. (2) is proved by induction on α. (3) follows by induction on the depth of x for $x \in \widehat{\mathcal{O}t}_{\mathcal{O}n_\varepsilon,<_\varepsilon}$, and by (2). The proof of (4) goes along the line of combinatorial properties of ordinal arithmetic that is captured in the Definitions 63–65 (see Takeuti [18] for an account of ordinal arithmetic for ordinals in Cantor normal form). We leave the tedious technical details of this proof to the reader. $\square$

6. Wellorderings beyond the Ordinals

We are now going to define the linear ordering $(E_0, \lhd)$, by using the generic notation system from Section 5. In some way, $(E_0, \lhd)$ can be seen as the analogue of $(\varepsilon_0, \in)$, with the set of the natural numbers, i.e. the ordinal ω, replaced by the class of all ordinals, i.e. the notation Ω. The ordering is shown to have well-founded initial segments up to the specific bounds Ω_n for any n, and this sequence of initial segments, i.e., iterated class comprehension along this segments, is later used in Section 10 for the proof-theoretic characterization of the choice principles AC and DC over NBG with full induction.

In the following definition, we extend the class of all epsilon numbers $\mathcal{O}n_\varepsilon$ with a new "virtual" epsilon number Ω on top of all others, to get $\mathcal{O}n_\Omega$, i.e., $\mathcal{O}n_\Omega$ has a top element Ω, in contrast to $\mathcal{O}n_\varepsilon$. Based on the symbols in $\mathcal{O}n_\Omega$ and the ordering on $\mathcal{O}n_\Omega$ we build the notation system E_0 and the ordering relation $\lhd$ according to the construction of the generic notation system in the previous section. From this construction, we also get the operations $\hat{+}$, $\hat{\cdot}$, and $\hat{\wedge}$, on the notations in E_0, corresponding to addition, multiplication and exponentiation, respectively.

Clearly, E_0 extends the class of all notations $\widehat{\alpha}$ of the ordinals, because of $\mathcal{O}n_\varepsilon \subseteq \mathcal{O}n_\Omega$. Actually, the class E_0 consists exactly of the notations $\widehat{\alpha}$ and $\widehat{\alpha}[\Omega]$ for all ordinals α, where $\widehat{\alpha}[\Omega]$ is just $\widehat{\alpha}$ with *all* the occurrences of the biggest epsilon number ε_β in $\widehat{\alpha}$ being replaced by Ω.

For any $\mathfrak{a} \in E_0$ we either have $\mathfrak{a} = \widehat{\alpha}$ for some α, or we have plenty of ordinals α such that $\mathfrak{a} = \widehat{\alpha}[\Omega]$. We can use this fact to simulate the operations $\hat{+}$, $\hat{\cdot}$, $\hat{\wedge}$, by the operations $+_\varepsilon$, $\cdot_\varepsilon$, $\wedge_\varepsilon$ (from Theorem 66), respectively. If we write $\varepsilon_{\mathsf{max}}(\widehat{\alpha})$ for the biggest epsilon number in $\widehat{\alpha}$, and $\varepsilon_{\mathsf{max}}(\widehat{\alpha}) = \emptyset$ for $\alpha < \varepsilon_0$, then, e.g., for any $\mathfrak{a}, \mathfrak{b} \in E_0$ we have that

$$\mathfrak{a} \mathbin{\hat{\cdot}} \mathfrak{b} = \begin{cases} \widehat{\alpha} \cdot_\varepsilon \widehat{\beta} & \mathfrak{a} = \widehat{\alpha} \wedge \mathfrak{b} = \widehat{\beta}, \\ & \mathfrak{a} = \widehat{\alpha}[\Omega] \wedge \mathfrak{b} = \widehat{\beta}[\Omega] \wedge \varepsilon_{\mathsf{max}}(\widehat{\alpha}) = \varepsilon_{\mathsf{max}}(\widehat{\beta}) > \emptyset \vee \\ (\widehat{\alpha} \cdot_\varepsilon \widehat{\beta})[\Omega] & \mathfrak{a} = \widehat{\alpha} \wedge \mathfrak{b} = \widehat{\beta}[\Omega] \wedge \varepsilon_{\mathsf{max}}(\widehat{\alpha}) < \varepsilon_{\mathsf{max}}(\widehat{\beta}) \vee \\ & \mathfrak{a} = \widehat{\alpha}[\Omega] \wedge \mathfrak{b} = \widehat{\beta} \wedge \varepsilon_{\mathsf{max}}(\widehat{\alpha}) > \varepsilon_{\mathsf{max}}(\widehat{\beta}), \end{cases}$$

and analogously we can compute addition and exponentiation on E_0.

For the ordering relation $\lhd$ we can use exactly the same trick, that is, we have $\mathfrak{a} \lhd \mathfrak{b}$ iff there are α, β, which represent $\mathfrak{a}$ and $\mathfrak{b}$, like in one of the four cases above, and $\widehat{\alpha} <_\varepsilon \widehat{\beta}$. Because of this very correspondence between $\hat{+}$, $\hat{\cdot}$, $\hat{\wedge}$, $\lhd$, and $+_\varepsilon$, $\cdot_\varepsilon$, $\wedge_\varepsilon$, $<_\varepsilon$, and because of Theorem 66, we also may transfer all the usual inequalities and equalities about ordinal arithmetic, from the ordinals to the notations in E_0.

Definition 67.

We define the expressions Ω, E_0, $\lhd$ (analogous Definition 18) such that

$$\begin{array}{rcl} z \in \Omega & := & z \in \langle \emptyset, \emptyset \rangle, \\ z \in \mathcal{O}n_\Omega & := & z \in \mathcal{O}n_\varepsilon \vee z = \Omega, \\ z \in \prec & := & \exists\alpha\exists\beta((z = \langle \alpha, \beta \rangle \wedge \alpha \in \beta) \vee z = \langle \alpha, \Omega \rangle), \\ z \in \prec_\Omega & := & z \in \widehat{\prec}_{\mathcal{O}n_\Omega}, \\ z \in E_0 & := & z \in \widehat{\mathcal{O}t}_{\mathcal{O}n_\Omega, \prec_\Omega}, \\ z \in \lhd & := & z \in \prec_\Omega \wedge (\exists x \in E_0)(\exists y \in E_0) z = \langle x, y \rangle, \\ z \in \unlhd & := & z \in \lhd \vee (\exists x \in E_0) z = \langle x, x \rangle. \end{array}$$

We use the letters $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}, \mathfrak{e}, \mathfrak{u}, \mathfrak{v}, \mathfrak{w}$ to denote elements in E_0. We write $\forall\mathfrak{a}A[\mathfrak{a}]$ and $\exists\mathfrak{a}A[\mathfrak{a}]$ for $(\forall x \in E_0)A[x]$ and $(\exists x \in E_0)A[x]$, respectively. Sometimes we drop the universal quantifier $\forall\mathfrak{a}$ in front of $A[\mathfrak{a}]$. In this case $A[\mathfrak{a}]$ stands for $x \in E_0 \to A[x]$, e.g. we may write $\mathcal{T} \vdash A[\mathfrak{a}]$ instead of $\mathcal{T} \vdash \forall\mathfrak{a}A[\mathfrak{a}]$.

We write $\hat{+}$, $\hat{\cdot}$, $\hat{\wedge}$ for $\hat{+}_{\mathcal{O}n_\Omega, \prec_\Omega}$, $\hat{\cdot}_{\mathcal{O}n_\Omega, \prec_\Omega}$, $\hat{\wedge}_{\mathcal{O}n_\Omega, \prec_\Omega}$, respectively. We write $\mathfrak{a}^\mathfrak{b}$ for $\mathfrak{a} \hat{\wedge} \mathfrak{b}$, and $x \lhd y$ and $x \unlhd y$ for $\langle x, y \rangle \in \lhd$ and $\langle x, y \rangle \in \unlhd$, respectively. Furter we write $(\exists x \lhd y)A[x]$ and $(\forall x \lhd y)A[x]$ for $\exists x(x \lhd y \wedge A[x])$ and $\forall x(x \lhd y \to A[x])$, respectively.

And finally, we write Ω_0 for $\Omega \hat{+} \widehat{1}$, and Ω_{n+1} for $\widehat{\omega}^{\Omega_n}$.

The following lemma states some simple properties about Ω, i.e. part 1 is used for part 2 of the subsequent theorem, about the linear ordering $(E_0, \lhd)$, which is a direct consequence of Lemma 61.

Lemma 68.

(1) $\mathsf{NBG} \vdash \Omega \notin \mathcal{O}n \wedge \Omega \notin \mathcal{H}_X \wedge \mathcal{H}_X \cap \mathcal{O}n = \{\emptyset\}$,

(2) $\mathsf{NBG} \vdash \Omega_0 = \langle\!\langle \Omega, \emptyset \rangle\!\rangle \wedge \Omega_{n+1} = \langle\!\langle \Omega_n \rangle\!\rangle$

Proof. By definition, by computation, and by induction on n. $\square$

Theorem 69. (Linear Ordering)

(1) E_0 and $\lhd$ are elementarily definable classes in NBG,

(2) $\mathsf{NBG} \vdash Lin[E_0, \lhd]$.

Proof. By definition and by Lemma 61. □

In the following we write $TI^{\mathfrak{a}}_{\lhd}[C[x]]$ for $TI^{x \lhd \mathfrak{a}}_{\lhd}[C[x]]$, and analogously we write $Wf^{\mathfrak{a}}_{\lhd}[C[x]]$ and $Prog^{\mathfrak{a}}_{\lhd}[C[x]]$.

The formula $Prog^{\mathfrak{a}}_{\lhd}[C[x]]$ is a subformula, i.e. the premise, of the formula $TI^{\mathfrak{a}}_{\lhd}[C[x]]$. We show that $Prog^{\mathfrak{a}}_{\lhd}[C[x]]$ can be replaced by the stronger statement $Prog^{E_0}_{\lhd}[C[x]]$, without changing the meaning of $TI^{\mathfrak{a}}_{\lhd}[C[x]]$. Clearly, this is the case, because $(E_0, \lhd)$ is a linear ordering.

Lemma 70.

$$\mathsf{NBG} \vdash TI^{\mathfrak{a}}_{\lhd}[C[x]] \leftrightarrow (Prog^{E_0}_{\lhd}[C[x]] \rightarrow (\forall x \lhd \mathfrak{a})C[x]).$$

Proof. We assume $Prog^{E_0}_{\lhd}[C[x]] \rightarrow (\forall x \lhd \mathfrak{a})C[x]$, i.e. its contrapositive

$$(\exists x \lhd \mathfrak{a})\neg C[x] \rightarrow (\exists x \in E_0)(\neg C[x] \wedge (\forall y \in E_0)(\neg C[y] \rightarrow y \not\lhd x),$$

and we show $TI^{\mathfrak{a}}_{\lhd}[C[x]]$. If there is some $\mathfrak{b}$ such that $\mathfrak{b} \lhd \mathfrak{a} \wedge \neg C[\mathfrak{b}]$ then by assumption we find $\mathfrak{d}$ such that $\neg C[\mathfrak{d}]$ and $\forall y \in E_0(\neg C[y] \rightarrow y \not\lhd \mathfrak{d})$, i.e. $\mathfrak{b} \not\lhd \mathfrak{d}$. We get $\mathfrak{d} \unlhd \mathfrak{b}$ because of totality, and $\mathfrak{d} \lhd \mathfrak{a}$ because of transitivity. Hence $(\exists x \lhd \mathfrak{a})(\neg C[x] \wedge (\forall y \lhd \mathfrak{a})(\neg C[y] \rightarrow y \not\lhd x)$, i.e. we have shown the contrapositive of $TI^{\mathfrak{a}}_{\lhd}[C[x]]$. The other direction follows because of $Prog^{E_0}_{\lhd}[C[x]] \rightarrow Prog^{\mathfrak{a}}_{\lhd}[C[x]]$; for any $\mathfrak{b}$ with $\mathfrak{b} \lhd \mathfrak{a}$ and $(\forall \mathfrak{c} \lhd \mathfrak{a})\mathfrak{c} \lhd \mathfrak{b} \rightarrow C[\mathfrak{c}]$ we get $(\forall \mathfrak{c} \in E_0)\mathfrak{c} \lhd \mathfrak{b} \rightarrow C[\mathfrak{c}]$ by totality, i.e. $(\forall \mathfrak{c} \in E_0)\mathfrak{c} \lhd \mathfrak{a} \vee \mathfrak{a} \unlhd \mathfrak{c}$, and because of $\mathfrak{a} \unlhd \mathfrak{c} \rightarrow \mathfrak{c} \not\lhd \mathfrak{b}$, by transitivity. Hence $C[\mathfrak{b}]$ by $Prog^{E_0}_{\lhd}[C[x]]$, that is $Prog^{\mathfrak{a}}_{\lhd}[C[x]]$. □

The following two lemmas and the subsequent theorem are in complete analogy to the standard wellordering proof for the common notation system for ε_0, see e.g. Pohlers [15].

Lemma 71.

(1) $\mathsf{NBG} \vdash \mathfrak{a} \hat{+} (\mathfrak{b} \hat{+} \mathfrak{c}) = (\mathfrak{a} \hat{+} \mathfrak{b}) \hat{+} \mathfrak{c}$,

(2) $\mathsf{NBG} \vdash \mathfrak{c} \neq \emptyset \wedge \mathfrak{a} \lhd \mathfrak{b} \hat{+} \widehat{\omega}^{\mathfrak{c}} \rightarrow (\exists \mathfrak{d} \lhd \mathfrak{c})(\exists p \in \omega)\, \mathfrak{a} \lhd \mathfrak{b} \hat{+} \widehat{\omega}^{\mathfrak{d}} \hat{\cdot} \widehat{p}$.

Proof. The proof is in analogy to the standard notation system for $(\varepsilon_0, <)$, where the notations are also based on the Cantor normal form of ordinals. The tedious combinatorial arguments are left to the reader. $\square$

Definition 72. (Abbreviations)

$$\begin{array}{rcl} \mathfrak{a} \subset C[y] & := & (\forall y \lhd \mathfrak{a}) C[y], \\ C_x^{\mathfrak{a}}[y] & := & (\forall \mathfrak{b} \lhd \mathfrak{a})(\mathfrak{b} \subset C[z/x] \to \mathfrak{b} \mathbin{\hat{+}} \widehat{\omega}^y \subset C[z/x]). \end{array}$$

Lemma 73.

For $\mathcal{F} \in \{\Pi_0^1, \mathcal{L}^1\}$ and $C \in \mathcal{F}$ we have that

(1) $\mathsf{NBG} \cup \mathit{Wf}_{\in}^{\omega}[\mathcal{F}] \vdash \mathit{Prog}_{\lhd}^{\widehat{\omega}^{\mathfrak{a}}}[C[y/x]] \to \mathit{Prog}_{\lhd}^{\mathfrak{a} \hat{+} \widehat{1}}[C_x^{\widehat{\omega}^{\mathfrak{a}}}[y]]$,

(2) $\mathsf{NBG} \cup \mathit{Wf}_{\in}^{\omega}[\mathcal{F}] \vdash \mathit{TI}_{\lhd}^{\mathfrak{a}}[C_x^{\widehat{\omega}^{\mathfrak{a}}}[y]] \to \mathit{TI}_{\lhd}^{\widehat{\omega}^{\mathfrak{a}}}[C[y/x]]$.

(We have $\mathsf{NBG} \vdash \mathit{Wf}_{\in}^{\omega}[\Pi_0^1]$ by Corollary 48.)

Proof. In the following we use the shorthand $C[y]$ for $C[y/x]$.

(1) For any $\mathfrak{a}$ with

(a) $\mathit{Prog}_{\lhd}^{\widehat{\omega}^{\mathfrak{a}}}[C[y]]$, i.e. $(\forall \mathfrak{b} \lhd \widehat{\omega}^{\mathfrak{a}})(\mathfrak{b} \subset C[z] \to C[\mathfrak{b}])$,

and for any $\mathfrak{b}$, $\mathfrak{c}$, with

(b) $\mathfrak{b} \lhd \mathfrak{a} \mathbin{\hat{+}} \widehat{1} \wedge \mathfrak{b} \subset C_x^{\widehat{\omega}^{\mathfrak{a}}}[y]$,

(c) $\mathfrak{c} \lhd \widehat{\omega}^{\mathfrak{a}} \wedge \mathfrak{c} \subset C[z]$,

we show that $\mathfrak{c} \mathbin{\hat{+}} \widehat{\omega}^{\mathfrak{b}} \subset C[z]$, i.e. $C_x^{\widehat{\omega}^{\mathfrak{a}}}[\mathfrak{b}]$ and hence $\mathit{Prog}_{\lhd}^{\mathfrak{a} \hat{+} \widehat{1}}[C_x^{\widehat{\omega}^{\mathfrak{a}}}[y]]$. If $\mathfrak{a} = \emptyset$ then $\mathfrak{c} \mathbin{\hat{+}} \widehat{\omega}^{\mathfrak{b}} = \widehat{1}$, and $C[\emptyset]$ by (a) because of $\emptyset \subset C[z]$. If $\emptyset \lhd \mathfrak{a}$ and $\mathfrak{e} \lhd \mathfrak{c} \mathbin{\hat{+}} \widehat{\omega}^{\mathfrak{b}}$, and if $\mathfrak{b} = \emptyset$ then $\mathfrak{e} \unlhd \mathfrak{c}$, hence $C[\mathfrak{e}]$ by (c)+(a), otherwise $\emptyset \lhd \mathfrak{b}$ and by Lemma 71 there are $\mathfrak{d} \lhd \mathfrak{b}$, $\mathfrak{d}_1 \lhd \mathfrak{a}$ and $p_0, p_1 \in \omega$ such that $\mathfrak{e} \lhd \mathfrak{c} \mathbin{\hat{+}} \widehat{\omega}^{\mathfrak{d}} \mathbin{\hat{\cdot}} \widehat{p_0}$ and $\mathfrak{c} \lhd \widehat{\omega}^{\mathfrak{d}_1} \mathbin{\hat{\cdot}} \widehat{p_1}$. We have $\mathfrak{c} \mathbin{\hat{+}} \widehat{\omega}^{\mathfrak{d}} \mathbin{\hat{\cdot}} \widehat{p} \lhd \widehat{\omega}^{\mathfrak{d}_1} \mathbin{\hat{\cdot}} \widehat{p_1} \mathbin{\hat{+}} \widehat{\omega}^{\mathfrak{d}} \mathbin{\hat{\cdot}} \widehat{p} \lhd \widehat{\omega}^{\mathfrak{a}}$ for any $p \in \omega$ because of $\mathfrak{d} \lhd \mathfrak{a}$, and further $C_x^{\widehat{\omega}^{\mathfrak{a}}}[\mathfrak{d}]$ by (b), hence $\forall p(\mathfrak{c} \mathbin{\hat{+}} \widehat{\omega}^{\mathfrak{d}} \mathbin{\hat{\cdot}} \widehat{p} \subset C[z])$ by induction on p, which is available because of $\mathit{Wf}_{\in}^{\omega}[\mathcal{F}]$, and finally $C[\mathfrak{e}]$.

(2) We assume

(d) $\mathit{Prog}_{\lhd}^{\mathfrak{a}}[C_x^{\widehat{\omega}^{\mathfrak{a}}}[y]] \to \mathfrak{a} \subset C_x^{\widehat{\omega}^{\mathfrak{a}}}[y]$,

(e) $\mathit{Prog}_{\lhd}^{\widehat{\omega}^{\mathfrak{a}}}[C[y]]$,

and we show $\widehat{\omega}^{\mathfrak{a}} \subset C[y]$. We have $\mathit{Prog}_{\lhd}^{\mathfrak{a} \hat{+} \widehat{1}}[C_x^{\widehat{\omega}^{\mathfrak{a}}}[y]]$ by (e)+(1), hence $\mathfrak{a} \subset C_x^{\widehat{\omega}^{\mathfrak{a}}}[y]$ by (d), and $C_x^{\widehat{\omega}^{\mathfrak{a}}}[\mathfrak{a}]$ by $\mathit{Prog}_{\lhd}^{\mathfrak{a} \hat{+} \widehat{1}}[C_x^{\widehat{\omega}^{\mathfrak{a}}}[y]]$. Finally $\widehat{\omega}^{\mathfrak{a}} \subset C[y]$ by $C_x^{\widehat{\omega}^{\mathfrak{a}}}[\mathfrak{a}]$ and because of $\emptyset \subset C[y]$.

$\square$

Finally, by using the previous main lemma, we can show transfinite induction along $\lhd$ up to any Ω_n, in NBG, for any elementary formula. And if we add full induction to NBG then we even get transfinite induction along $\lhd$ up to any Ω_n for all formulas.

Theorem 74.

(1) $\mathsf{NBG} \vdash \mathit{TI}_{\lhd}^{\Omega_n}[\Pi_0^1]$,

(2) $\mathsf{NBG} \cup \mathit{Wf}_{\in}^{\mathcal{V}}[\mathcal{L}^1] \vdash \mathit{TI}_{\lhd}^{\Omega_n}[\mathcal{L}^1]$.

Proof.

(1) We have $\mathsf{NBG} \vdash \mathit{Wf}_{\in}^{\mathcal{O}n}[\Pi_0^1]$ by Corollary 48, hence $\mathsf{NBG} \vdash \mathit{Wf}_{\lhd}^{\Omega}[\Pi_0^1]$ by Theorem 66 and because of $\widehat{\mathcal{O}n} = \{\mathfrak{a} \mid \mathfrak{a} \lhd \Omega\}$. We get $\mathsf{NBG} \vdash \mathit{TI}_{\lhd}^{\Omega}[\Pi_0^1]$ by Lemma 50. To show $\mathsf{NBG} \vdash \mathit{TI}_{\lhd}^{\Omega_0}[\Pi_0^1]$, i.e. $\mathsf{NBG} \vdash \mathit{TI}_{\lhd}^{\Omega \hat{+} \widehat{1}}[\Pi_0^1]$, we assume $\mathit{Prog}_{\lhd}^{\Omega \hat{+} \widehat{1}}[C[y]]$, and we get $\Omega \subset C[y]$ by $\mathit{TI}_{\lhd}^{\Omega}[\Pi_0^1]$. To show $\Omega \mathbin{\hat{+}} \widehat{1} \subset C[y]$, we need $C[\Omega]$. But $C[\Omega]$ holds because of $\mathit{Prog}_{\lhd}^{\Omega \hat{+} \widehat{1}}[C[y]]$ and $\Omega \subset C[y]$. Finally, $\mathsf{NBG} \vdash \mathit{TI}_{\lhd}^{\Omega_n}[\Pi_0^1]$ is shown by induction on n, by applying Lemma 73, i.e. $\mathit{TI}_{\lhd}^{\Omega_{n-1}}[C_x^{\Omega_n}[y]] \to \mathit{TI}_{\lhd}^{\Omega_n}[C[y/x]]$, where we use that $C_x^{\Omega_n}[z] \in \Pi_0^1$ for $C[z] \in \Pi_0^1$.

(2) Analogous to (1). □

7. Iterated Class Comprehension

In this section we define the schemata of iterated class comprehension, which is used to inductively define hierarchies of classes. We show that the choice schemes AC and DC can be used to prove the existence of such class hierarchies defined by iterated class comprehension along initial segments of the linear ordering $(E_0, \lhd)$ from Section 6.

Iterated class comprehension is later used in Section 8 to define truth and proof predicates in NBG and extensions thereof, and it is used in Section 9 to define cumulative hierarchies of classes for the asymmetric interpretations.

The following definition is a generalization of the expressions $(X)^y$ and $\langle X\rangle^y$ from Definition 30, such that, e.g., $(X)^y$ now also becomes $(X)^{\in,y}$, and such that $(X)^{Z,y}$ is defined for arbitrary binary relations Z.

Definition 75. (Abbreviations)

We define the expressions $(X)^{Z,y}$ and $\langle X\rangle^{Z,y}$ (analogous Definition 18) such that

$$\begin{aligned} z \in (X)^{Z,y} &\quad := \quad z \in X \wedge \exists v \exists w (z = \langle v, w\rangle \wedge \langle w, y\rangle \in Z), \\ z \in \langle X\rangle^{Z,y} &\quad := \quad \exists w (\langle z, w\rangle \in X \wedge \langle w, y\rangle \in Z). \end{aligned}$$

Iterated comprehension essentially allows us to build hierarchies of classes U, such that $(U)_y$ depends on the levels $(U)_z$ where z is any predecessor of y, i.e. a predecessor with respect to some fixed binary relation Z. On the level y in the hierarchy we have that $(U)_y = \{x \mid C[(U)^{Z,y}, x, y]\}$ for some fixed formula C, such that C and the relation Z actually determine the whole hierarchy, and if Z is well-founded then the hierarchy is even uniquely determined.

Definition 76. (Iterated Comprehension)

For $A, B \in \mathcal{L}^1$ with one free set variable and no other free variables, and $\mathcal{D} \subseteq \mathcal{L}^1$, we define

$$\begin{aligned} \mathit{Hier}_Z^Y[C[U,x,y]] &\quad := \quad (\forall y \in Y)\forall x (x \in (U)_y \leftrightarrow C[(U)^{Z,y}, x, y]), \\ \mathit{Hier}_A^B[C[U,x,y]] &\quad := \quad \forall y (B[y] \to \forall x (x \in (U)_y \leftrightarrow C[(U)^{A,y}, x, y])), \\ \exists\, \mathit{Hier}_A^B[\mathcal{D}] &\quad := \quad \{\exists U (\mathit{Hier}_A^B[C[U,x,y]]) \mid C \in \mathcal{D}\}. \end{aligned}$$

In the following we write $\mathit{Hier}_\lhd^{\mathfrak{a}}[C[U,x,y]]$ for $\mathit{Hier}_\lhd^{y \lhd \mathfrak{a}}[C[U,x,y]]$.

The following lemma shows the essential uniqueness of class hierarchies built by iterated comprehension along well-founded class relations.

Lemma 77.

If $A, B, C \in \mathcal{L}^1$ and $D := \forall x \forall y(A[\langle y, x\rangle] \wedge B[x] \rightarrow B[y])$, and $\mathsf{NBG} \subseteq \mathcal{T}$, and $\mathcal{T} \vdash \mathit{TI}^B_A[\Pi^1_0]$, then

$$\mathcal{T} \vdash \mathit{Hier}^B_A[C[U, x, y]] \wedge \mathit{Hier}^B_A[C[V, x, y]] \wedge D \rightarrow \forall y(B[y] \rightarrow (U)_y = (V)_y).$$

Proof. If we assume $\mathit{Hier}^B_A[C[U, x, y]]$ and $\mathit{Hier}^B_A[C[V, x, y]]$, and if $E[y] := (U)_y = (V)_y$, then for any x with $B[x]$ we have $\forall y(A[\langle y, x\rangle] \rightarrow B[y])$ and

$$\forall y(B[y] \wedge A[\langle y, x\rangle] \rightarrow E[y]) \rightarrow (U)^{A,x} = (V)^{A,x}.$$

Hence $\mathit{Prog}^B_A[E[y]]$, and the claim follows by $\mathit{TI}^B_A[E[y]]$. $\square$

The previous lemma can be easily applied to well-founded relations in NBG.

Corollary 78.

For all $C \in \mathcal{L}^1$ we have that

(1) $\mathsf{NBG} \vdash \mathit{Hier}^{\mathcal{O}n}_{\in}[C[U, x, y]] \wedge \mathit{Hier}^{\mathcal{O}n}_{\in}[C[V, x, y]] \rightarrow \forall \alpha((U)_\alpha = (V)_\alpha)$,

(2) $\mathsf{NBG} \vdash \mathit{Hier}^{\Omega^{\widehat{n}}}_{\lhd}[C[U, x, y]] \wedge \mathit{Hier}^{\Omega^{\widehat{n}}}_{\lhd}[C[V, x, y]] \rightarrow (\forall \mathfrak{a} \lhd \Omega^{\widehat{n}})\,(U)_{\mathfrak{a}} = (V)_{\mathfrak{a}}$,

(3) $\mathsf{NBG} \vdash \mathit{Hier}^{\Omega_n}_{\lhd}[C[U, x, y]] \wedge \mathit{Hier}^{\Omega_n}_{\lhd}[C[V, x, y]] \rightarrow (\forall \mathfrak{a} \lhd \Omega_n)\,(U)_{\mathfrak{a}} = (V)_{\mathfrak{a}}$.

The following theorem shows that if we add appropriate choice schemes and induction to NBG, then we can prove the existence of class hierarchies defined by iterated comprehension along initial segments of the linear ordering $(E_0, \lhd)$. The argument for the proof is similar to that for analogous statements in second order arithmetic, see e.g. Cantini [1], and Feferman and Sieg [4].

Theorem 79.

(1) $\mathsf{NBG} \cup \mathit{CA}[\Pi^1_n] \vdash \exists\, \mathit{Hier}^{\overline{m}}_{\in}[\Pi^1_n]$,

(2) $\mathsf{NBG} \cup \mathit{DC}[\Pi^1_n] \vdash \exists\, \mathit{Hier}^{\Omega^{\widehat{m}}}_{\lhd}[\Pi^1_n]$,

(3) $\mathsf{NBG} \cup \mathit{AC}[\Pi^1_n] \cup \mathit{Wf}^{\mathcal{V}}_{\in}[\mathcal{L}^1] \vdash \exists\, \mathit{Hier}^{\Omega_m}_{\lhd}[\Pi^1_n]$.

Proof.

(1) We trivially have $\exists\, Hier_{\in}^{\overline{0}}[\Pi^1_n]$, and for any $C \in \Pi^1_n$ we show that

$$Hier_{\in}^{\overline{m}}[C[V,x,y]] \to \exists U(Hier_{\in}^{\overline{m+1}}[C[U,x,y]]),$$

hence we get $\exists U(Hier_{\in}^{\overline{m}}[C[U,x,y]]$ by induction on m. If $Hier_{\in}^{\overline{m}}[C[V,x,y]]$, then by $CA[\Pi^1_n]$ we get some W such that

$$\forall z(z \in W \leftrightarrow C[(V)^{\overline{m}}, z, \overline{m}]),$$

and by comprehension we get U such that $Hier_{\in}^{\overline{m+1}}[C[U,x,y]]$, i.e.

$$U \quad := \quad \{\langle z,p\rangle \mid (p \in \overline{m} \wedge \langle z,p\rangle \in V) \vee (p = \overline{m} \wedge z \in W)\}.$$

(2) Proof by induction on m. $\Omega^{\widehat{0}} = \widehat{1}$, hence $\exists\, Hier_{\lhd}^{\Omega^{\widehat{0}}}[\Pi^1_n]$ by $CA[\Pi^1_n]$ which we have by Corollary 36. Let $E[\mathfrak{a}, U] := Hier_{\lhd}^{\mathfrak{a}}[A[U,x,y]]$ for some $A \in \Pi^1_n$, that is, $E[\mathfrak{a}, U]$ is equivalent to formulas in Π^1_{n+1} and Σ^1_{n+1}. We assume $\exists\, Hier_{\lhd}^{\Omega^{\widehat{m}}}[\Pi^1_n]$ and we show $(\exists Z)E[\Omega^{\widehat{m+1}}, Z]$; by case distinction on α, i.e. $\alpha = \emptyset$, or α is a successor, or α is a limit, we get

$$\forall\alpha\forall Y\exists Z((\forall\beta \in \alpha)E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, Y] \to E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\alpha}, Z]). \qquad (*)$$

I.e. if $\alpha = \emptyset$ then $\Omega^{\widehat{m}} \hat{\cdot} \widehat{\alpha} = \emptyset$ and $E[\emptyset, Z]$ holds trivially. If α is a successor, $\alpha = \beta'$, then we may assume $E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, Y]$, and by using Y and because of $\exists\, Hier_{\lhd}^{\Omega^{\widehat{m}}}[\Pi^1_n]$ we can build Z such that $E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta} \hat{+} \Omega^{\widehat{m}}, Z]$. If $\alpha \neq \emptyset$ is a limit ordinal then for any $\mathfrak{a} \lhd \Omega^{\widehat{m}} \hat{\cdot} \widehat{\alpha}$ there is some $\beta \in \alpha$ such that $\mathfrak{a} \lhd \Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}$, hence we have $(\forall\beta \in \alpha)E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, Y] \to E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\alpha}, Y]$. We have $DC^{\Diamond}_{\mathcal{O}n}[\Sigma^1_{n+1}]$ by Corollary 45, and we get $\exists Z\forall\alpha B[Z,\alpha]$ with

$$B[Z,\alpha] \quad := \quad (\forall\beta \in \alpha)E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, \langle Z\rangle^{\alpha}] \to E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\alpha}, (Z)_{\alpha}]$$

by $DC^{\Diamond}_{\mathcal{O}n}[\Sigma^1_{n+1}]$ and (*). If we define

$$\begin{aligned} C[Z,\gamma] \quad &:= \quad (\forall\alpha \in \gamma)(\forall\beta \in \alpha)E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, \langle Z\rangle^{\alpha}] \to \\ &\qquad\qquad (\forall\beta \in \gamma)E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, (Z)_{\beta}], \\ D[Z,\gamma] \quad &:= \quad (\forall\alpha \in \gamma)(\forall\beta \in \alpha)E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, \langle Z\rangle^{\alpha}] \to \\ &\qquad\qquad (\forall\beta \in \gamma)E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, \langle Z\rangle^{\gamma}]. \end{aligned}$$

then we have $(\forall\alpha \in \gamma)B[Z,\alpha] \to C[Z,\gamma]$, i.e. $\forall\alpha B[Z,\alpha] \to \forall\gamma C[Z,\gamma]$, and further we have $C[Z,\gamma] \to D[Z,\gamma]$. By Theorem 35 we have comprehension for the formula E, hence because of $TI_{\in}^{\mathcal{O}n}[\Pi^1_0]$ we get

$$\forall\gamma D[Z,\gamma] \to \forall\gamma(\forall\beta \in \gamma)E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, (Z)^{\gamma}].$$

We get $\exists Z \forall \alpha B[Z,\alpha] \to \exists Z \forall \gamma (\forall \beta \in \gamma) E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, (Z)^{\gamma}]$ by the three preceding implications. We have that $\exists Z \forall \alpha B[Z,\alpha]$, hence there is some Z such that $\forall \beta E[\Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}, (Z)^{\beta'}]$, i.e. we get $U := \{z \mid \exists \alpha \langle z, \alpha\rangle \in Z\}$ by comprehension, such that $E[\Omega^{\widehat{m+1}}, U]$, because there is some β with $\mathfrak{a} \lhd \Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta}$ for any $\mathfrak{a} \lhd \Omega^{\widehat{m+1}}$.

(3) Let $E[\mathfrak{a}, U] := \mathit{Hier}^{\mathfrak{a}}_{\lhd}[A[U,x,y]]$ for some $A \in \Pi^1_n$. If we show

$$(\forall \mathfrak{b} \lhd \mathfrak{a}) \exists Y\, E[\mathfrak{b}, Y] \to \exists Y E[\mathfrak{a}, Y] \text{ for any } \mathfrak{a} \lhd \Omega_{m+1},$$

then because of $TI^{\Omega_{m+1}}_{\lhd}[\mathcal{L}^1]$ we get $(\forall \mathfrak{b} \lhd \Omega_{m+1}) \exists Y\, E[\mathfrak{b}, Y]$, i.e. $\exists Y E[\Omega_m, Y]$. If $\mathfrak{a} = \emptyset$ then $E[\mathfrak{a}, Y]$ holds trivially. If $\mathfrak{a} = \mathfrak{c} \hat{+} \widehat{1}$ then we may assume V such that $E[\mathfrak{c}, V]$. By Corollary 36 we have $CA[\Pi^1_n]$, hence we get Y such that

$$Y \quad := \quad \{\langle x, \mathfrak{b}\rangle \mid (\mathfrak{b} = \mathfrak{c} \wedge A[(V)^{\mathfrak{c}}, x, \mathfrak{c}]) \vee (\mathfrak{b} \lhd \mathfrak{c} \wedge x \in (V)_{\mathfrak{b}})\},$$

and $(Y)^{\mathfrak{c}} = (V)^{\mathfrak{c}}$, and $(Y)_{\mathfrak{c}} = \{x \mid A[(Y)^{\mathfrak{c}}, x, \mathfrak{c}]\}$, i.e., $E[\mathfrak{a}, Y]$. If $\mathfrak{a} \lhd \Omega_{m+1}$ is a limit, that is $(\forall \mathfrak{b} \lhd \mathfrak{a}) \mathfrak{b} \hat{+} \widehat{1} \lhd \mathfrak{a}$, and if we assume that $(\forall \mathfrak{b} \lhd \mathfrak{a}) \exists Y E[\mathfrak{b}, Y]$, then there is V such that $(\forall \mathfrak{b} \lhd \mathfrak{a}) E[\mathfrak{b}, (V)_{\mathfrak{b}}]$ by $AC[\Pi^1_n]$. By elementary comprehension we get

$$Y \quad := \quad \{\langle x, \mathfrak{b}\rangle \mid \mathfrak{b} \lhd \mathfrak{a} \wedge \langle x, \mathfrak{b}\rangle \in (V)_{\mathfrak{b} \hat{+} \widehat{1}}\}$$

and we show that $E[\mathfrak{a}, Y]$. For $\mathfrak{b} \lhd \mathfrak{a}$ we have $E[\mathfrak{b} \hat{+} \widehat{1}, (V)_{\mathfrak{b} \hat{+} \widehat{1}}]$, and for $\mathfrak{c} \lhd \mathfrak{b}$ we further get $E[\mathfrak{c} \hat{+} \widehat{1}, (V)_{\mathfrak{b} \hat{+} \widehat{1}}]$ and $E[\mathfrak{c} \hat{+} \widehat{1}, (V)_{\mathfrak{c} \hat{+} \widehat{1}}]$, hence $((V)_{\mathfrak{c} \hat{+} \widehat{1}})^{\mathfrak{c} \hat{+} \widehat{1}} = ((V)_{\mathfrak{b} \hat{+} \widehat{1}})^{\mathfrak{c} \hat{+} \widehat{1}}$ by Lemma 77, and $(\forall \mathfrak{b} \lhd \mathfrak{a})\, (Y)^{\mathfrak{b}} = ((V)_{\mathfrak{b} \hat{+} \widehat{1}})^{\mathfrak{b}}$, because for $\mathfrak{c} \lhd \mathfrak{b} \lhd \mathfrak{a}$ we have that

$$((Y)^{\mathfrak{b}})_{\mathfrak{c}} = ((V)_{\mathfrak{c} \hat{+} \widehat{1}})_{\mathfrak{c}} = ((V)_{\mathfrak{b} \hat{+} \widehat{1}})_{\mathfrak{c}} = (((V)_{\mathfrak{b} \hat{+} \widehat{1}})^{\mathfrak{b}})_{\mathfrak{c}}.$$

For $\mathfrak{b} \lhd \mathfrak{a}$ we have $E[\mathfrak{b} \hat{+} \widehat{1}, (V)_{\mathfrak{b} \hat{+} \widehat{1}}]$, hence for any x we finally get

$$A[(Y)^{\mathfrak{b}}, x, \mathfrak{b}] \text{ iff } A[((V)_{\mathfrak{b} \hat{+} \widehat{1}})^{\mathfrak{b}}, x, \mathfrak{b}] \text{ iff } \langle x, \mathfrak{b}\rangle \in (V)_{\mathfrak{b} \hat{+} \widehat{1}} \text{ iff } x \in (Y)_{\mathfrak{b}},$$

that is $E[\mathfrak{a}, Y]$. $\square$

8. Gödelization, Formalized Truth and Proof

In this section we introduce *formalized* versions of the notions of formula, truth, and proof, in complete analogy to the notions defined in Section 1. The formalized version $\mathcal{G}_{\mathcal{L}^1}$ of the language $\mathcal{L}^1$, and the predicates for truth and proof are defined within NBG (the predicates usually only exist as classes in extensions of NBG). In contrast to $\mathcal{L}^1$, we will now also have constants for classes and constants for all sets in the formalized language $\mathcal{G}_{\mathcal{L}^1}$. And the formalized proof system will have an additional infinitary inference rule for the universal set quantifier.

The aim of introducing predicates for truth and proofs is twofold; for the asymmetric interpretation in Section 10 we use a cumulative hierarchy of classes, which resembles the constructible hierarchy in set theory, that is, any level of the class hierarchy contains all classes that can be built by elementary comprehension with class parameters from lower levels of the hierarchy (see Section 9). Clearly, for the definition of this hierarchy we need an appropriate truth predicate within NBG, which at least reflects the truth of elementary formulas. And further, we need a formalized proof predicate, because an infinitary proof system is used to deal with full induction $TI_{\in}[\mathcal{L}^1]$ in the asymmetric interpretation (as described in the Introduction).

But, first of all we want to represent formulas as sets in NBG. This is achieved by the following Gödelization. Analogous to the language $\mathcal{L}^1$ we have codes for free and bound set variables $\dot{\mathtt{v}}_x, \dot{\mathtt{u}}_x$, and free and bound class variables $\dot{\mathtt{V}}_x, \dot{\mathtt{U}}_x$, and further we have set constants $\dot{\mathtt{c}}_x$, and class constants $\dot{\mathtt{C}}_x$. And in analogy to the logical connectives in $\mathcal{L}^1$, we have operations on codes, i.e., $\dot{\sim}$, $\dot{\in}$, $\dot{\vee}$, $\dot{\wedge}$, $\dot{\exists}$, $\dot{\forall}$.

Definition 80. (Gödelization)

We define the following expressions (for variables x, y)

$$
\begin{array}{rcl@{\qquad}rcl}
\dot{\mathtt{v}}_x & := & \langle\!\langle \overline{0}, x \rangle\!\rangle, & \dot{\sim} x & := & \langle\!\langle \overline{6}, x \rangle\!\rangle, \\
\dot{\mathtt{c}}_x & := & \langle\!\langle \overline{1}, x \rangle\!\rangle, & x \mathbin{\dot{\in}} y & := & \langle\!\langle \overline{7}, x, y \rangle\!\rangle, \\
\dot{\mathtt{V}}_x & := & \langle\!\langle \overline{2}, x \rangle\!\rangle, & x \mathbin{\dot{\vee}} y & := & \langle\!\langle \overline{8}, x, y \rangle\!\rangle, \\
\dot{\mathtt{C}}_x & := & \langle\!\langle \overline{3}, x \rangle\!\rangle, & x \mathbin{\dot{\wedge}} y & := & \langle\!\langle \overline{9}, x, y \rangle\!\rangle, \\
\dot{\mathtt{u}}_x & := & \langle\!\langle \overline{4}, x \rangle\!\rangle, & \dot{\exists}\, x\, y & := & \langle\!\langle \overline{10}, x, y \rangle\!\rangle, \\
\dot{\mathtt{U}}_x & := & \langle\!\langle \overline{5}, x \rangle\!\rangle, & \dot{\forall}\, x\, y & := & \langle\!\langle \overline{11}, x, y \rangle\!\rangle.
\end{array}
$$

We write $\dot{x}$, that is $\dot{\mathtt{v}}_n$ and $\dot{\mathtt{u}}_n$, for $\dot{\mathtt{v}}_{\overline{n}}$ and $\dot{\mathtt{u}}_{\overline{n}}$, respectively, and analogously $\dot{Y}$ for class variables Y. For formulas $C \in \mathcal{L}^1$ we inductively define the expression $\ulcorner C \urcorner$ such that

$$\begin{array}{rcl} \ulcorner x \in y \urcorner & := & \dot{x} \mathrel{\dot{\in}} \dot{y}, \\ \ulcorner x \in Y \urcorner & := & \dot{x} \mathrel{\dot{\in}} \dot{Y}, \\ \ulcorner \sim A \urcorner & := & \dot{\sim} \ulcorner A \urcorner, \\ \ulcorner (A \vee B) \urcorner & := & \ulcorner A \urcorner \mathbin{\dot{\vee}} \ulcorner B \urcorner, \\ \ulcorner (A \wedge B) \urcorner & := & \ulcorner A \urcorner \mathbin{\dot{\wedge}} \ulcorner B \urcorner, \end{array} \qquad \begin{array}{rcl} \ulcorner \exists x A[x/u] \urcorner & := & \dot{\exists}\, \dot{x}(\ulcorner A \urcorner[\dot{x}/\dot{u}]), \\ \ulcorner \forall x A[x/u] \urcorner & := & \dot{\forall}\, \dot{x}(\ulcorner A \urcorner[\dot{x}/\dot{u}]), \\ \ulcorner \exists X A[X/U] \urcorner & := & \dot{\exists}\, \dot{X}(\ulcorner A \urcorner[\dot{X}/\dot{U}]), \\ \ulcorner \forall X A[X/U] \urcorner & := & \dot{\forall}\, \dot{X}(\ulcorner A \urcorner[\dot{X}/\dot{U}]), \end{array}$$

where $\ulcorner A \urcorner[\dot{x}/\dot{u}]$ and $\ulcorner A \urcorner[\dot{X}/\dot{U}]$ stand for the expressions that are obtained from $\ulcorner A \urcorner$ by replacing all occurrences of $\dot{u}$ and $\dot{U}$ by $\dot{x}$ and $\dot{X}$, respectively. We use the shorthand notation $\ulcorner \Gamma \urcorner$ for $\{\ulcorner A \urcorner \mid A \in \Gamma\}$.

Having defined the particular sets representing the formulas in $\mathcal{L}^1$, we further define the whole class of Gödel-codes $\mathcal{G}_{\mathcal{L}^1}$, e.g., which additionally contains the codes of formulas where variables are replaced by constants. We further define some useful operations on these codes $\phi \in \mathcal{G}_{\mathcal{L}^1}$, like substitution of "variables", $\phi[\dot{\mathtt{v}}_y/\dot{\mathtt{v}}_x]$, $\phi[\dot{\mathtt{u}}_y/\dot{\mathtt{v}}_x]$, $\phi[\dot{\mathtt{c}}_y/\dot{\mathtt{v}}_x]$, $\phi[\dot{\mathtt{V}}_y/\dot{\mathtt{V}}_x]$, $\phi[\dot{\mathtt{U}}_y/\dot{\mathtt{V}}_x]$, $\phi[\dot{\mathtt{C}}_y/\dot{\mathtt{V}}_x]$, complementation $\neg\phi$, a rank function $rk_X(\phi)$, and a function $term(\phi)$ to unveil all variables and constants in ϕ, i.e., $term(\phi)$ is the set of all codes, $\dot{\mathtt{v}}_y, \dot{\mathtt{u}}_y, \dot{\mathtt{c}}_y, \dot{\mathtt{V}}_y, \dot{\mathtt{U}}_y, \dot{\mathtt{C}}_y$, occurring in ϕ. All operations on $\mathcal{G}_{\mathcal{L}^1}$ are in complete analogy to the operations on $\mathcal{L}^1$.

The formula $Sub[f, x, y]$ in the following definition is just a compact form to write that f is a function such that $f(\phi) = (\phi[x/y])$ for any $\phi \in dom(f)$, and if $\phi \in \mathcal{G}_{\mathcal{L}^1} \cap dom(f)$ then any "subformula" (involved in the construction) of ϕ is in $dom(f)$ and $term(\phi) \subseteq dom(f)$. Similarly $Ter[f]$ means that f is a function such that $f(\phi) = term(\phi)$ for any $\phi \in dom(f)$, and if $\phi \in \mathcal{G}_{\mathcal{L}^1} \cap dom(f)$ then any "subformula" (involved in the construction) of ϕ is in $dom(f)$ and $term(\phi) \subseteq dom(f)$. And the formula $Goe[x]$ means that $x \subseteq \mathcal{G}_{\mathcal{L}^1}$, and if $\phi \in x$ then any "subformula" (involved in the construction) of ϕ is in x; see Lemma 82 for the contents of $\mathcal{G}_{\mathcal{L}^1}$.

Definition 81. (Abbreviations)

We define the expressions $u[x/y]$, $term(u)$, $\mathcal{G}_{\mathcal{L}^1}$ (analogous Definition 18) such that

$$\begin{array}{rcl} z \in u[x/y] & := & \exists f(Sub[f, x, y] \wedge u \in dom(f) \wedge z \in f(u)), \\ z \in term(u) & := & \exists f(Ter[f] \wedge u \in dom(f) \wedge z \in f(u)), \\ z \in \mathcal{G}_{\mathcal{L}^1} & := & \exists x(Goe[x] \wedge z \in x), \end{array}$$

$$
\begin{aligned}
Sub[f,x,y] \quad := \quad & Fun[f] \wedge (\forall g \in dom(f))\exists p(\\
& \exists v(g = \langle\!\langle p, v\rangle\!\rangle \wedge p \in \overline{6} \wedge \\
& ((g = y \wedge f(g) = x) \vee (g \neq y \wedge f(g) = g))) \vee \\
& (\exists g_1 \in dom(f))(\exists g_2 \in dom(f))(p \in \overline{12}\backslash\overline{7} \wedge \\
& ((g = \langle\!\langle \overline{6}, g_1\rangle\!\rangle \wedge f(g) = \langle\!\langle \overline{6}, f(g_1)\rangle\!\rangle) \vee \\
& (g = \langle\!\langle p, g_1, g_2\rangle\!\rangle \wedge f(g) = \langle\!\langle p, f(g_1), f(g_2)\rangle\!\rangle)))),
\end{aligned}
$$

$$
\begin{aligned}
Ter[f] \quad := \quad & Fun[f] \wedge (\forall g \in dom(f))\exists p(\\
& \exists v(g = \langle\!\langle p, v\rangle\!\rangle \wedge p \in \overline{6} \wedge f(g) = \{g\}) \vee \\
& (\exists g_1 \in dom(f))(\exists g_2 \in dom(f))(p \in \overline{12}\backslash\overline{7} \wedge \\
& ((g = \langle\!\langle \overline{6}, g_1\rangle\!\rangle \wedge f(g) = f(g_1)) \vee \\
& (g = \langle\!\langle p, g_1, g_2\rangle\!\rangle \wedge f(g) = f(g_1) \cup f(g_2)))),
\end{aligned}
$$

$$
\begin{aligned}
Goe[x] \quad := \quad & (\forall f \in x)(\\
& (\exists p \in \overline{2})(\exists q \in \overline{4})\exists u \exists v((f = \langle\!\langle \overline{7}, \langle\!\langle p, u\rangle\!\rangle, \langle\!\langle q, v\rangle\!\rangle\rangle\!\rangle \vee \\
& f = \langle\!\langle \overline{6}, \langle\!\langle \overline{7}, \langle\!\langle p, u\rangle\!\rangle, \langle\!\langle q, v\rangle\!\rangle\rangle\!\rangle\rangle\!\rangle) \wedge (p \neq \overline{0} \vee u \in \omega) \wedge \\
& (q \notin \{\overline{0}, \overline{2}\} \vee v \in \omega)) \vee \\
& (\exists f_1 \in x)(\exists f_2 \in x)(\exists p \in \overline{10}\backslash\overline{8}) f = \langle\!\langle p, f_1, f_2\rangle\!\rangle \vee \\
& (\exists f_1 \in x)(\exists p \in \overline{12}\backslash\overline{10})\exists q \exists r \exists u \exists v(\\
& ((u = \dot{\mathtt{u}}_q \wedge v = \dot{\mathtt{v}}_r) \vee (u = \dot{\mathtt{U}}_q \wedge v = \dot{\mathtt{V}}_r)) \wedge \\
& u \notin term(f_1) \wedge f = \langle\!\langle p, v, f_1[u/v]\rangle\!\rangle)).
\end{aligned}
$$

We use the letters ϕ, ψ, θ, ξ to denote elements in $\mathcal{G}_{\mathcal{L}^1}$. We write $\forall\phi A[\phi]$ and $\exists\phi A[\phi]$ for $(\forall x \in \mathcal{G}_{\mathcal{L}^1})A[x]$ and $(\exists x \in \mathcal{G}_{\mathcal{L}^1})A[x]$, respectively.

For the following expressions we just give an informal description; the formal definitions would be similar to the previous ones:

We define the expressions $\mathcal{G}_{\Pi^1_n}$, $\mathcal{G}_{\Sigma^1(\Pi^1_n)}$, etc., in analogy to $\mathcal{G}_{\mathcal{L}^1}$. E.g. $\mathcal{G}_{\Pi^1_n}$ is the class of Gödel codes representing the set of formulas Π^1_n, in the same way as $\mathcal{G}_{\mathcal{L}^1}$ represents $\mathcal{L}^1$, that is, we have $A \in \Pi^1_n$ iff $\ulcorner A \urcorner \in \mathcal{G}_{\Pi^1_n}$. Further, if $\phi \in \mathcal{G}_{\Pi^1_n}$ then $\phi[\dot{\mathtt{c}}_x/\dot{\mathtt{v}}_p], \phi[\dot{\mathtt{C}}_x/\dot{\mathtt{V}}_p] \in \mathcal{G}_{\Pi^1_n}$ for any x and $p \in \omega$ (i.e. $\mathcal{G}_{\Pi^1_n}$ is closed under substitution with constants).

For the formalized proof predicate the Gödel-codes without class constants will play an important role, hence we define the expression $\mathcal{G}^-_{\mathcal{L}^1}$ such that $\mathcal{G}^-_{\mathcal{L}^1} = \{\phi \in \mathcal{G}_{\mathcal{L}^1} \mid \forall x(\dot{\mathtt{C}}_x \notin term(\phi))\}$.

For functions f, g, we define the expression $\phi[f,g]_\infty$ such that $\phi[f,g]_\infty \in \mathcal{G}_{\mathcal{L}^1}$ is just ϕ with all "free variables" $\dot{\mathtt{v}}_p, \dot{\mathtt{V}}_q \in term(\phi)$ replaced by the constants $\dot{\mathtt{c}}_{f(p)}$, $\dot{\mathtt{C}}_{g(q)}$, respectively, i.e., if $\dot{\mathtt{v}}_s, \dot{\mathtt{V}}_s \notin term(\phi)$ for every $s > r$ then

$$\phi[f,g]_\infty = \phi[\dot{\mathtt{c}}_{f(0)}/\dot{\mathtt{v}}_0]...[\dot{\mathtt{c}}_{f(r)}/\dot{\mathtt{v}}_r][\dot{\mathtt{C}}_{g(0)}/\dot{\mathtt{V}}_0]...[\dot{\mathtt{C}}_{f(r)}/\dot{\mathtt{V}}_r].$$

Analogously we define $\phi[f]_\infty$, where only "free set variables" $\dot{\mathtt{v}}_p \in term(\phi)$ are replaced by the specified constants.

The rank $rk_X(\phi)$ is defined analogous to the formula rank, $\mathsf{rk}_{\mathcal{C}}$, e.g., such that $rk_{\mathcal{G}_{\mathcal{C}}}(\ulcorner A \urcorner) = \mathsf{rk}_{\mathcal{C}}(A)$. We write rk for $rk_{\emptyset}$.

The complement $\neg\phi$ is defined in analogy to the complementation of formulas, e.g., such that $\neg\ulcorner A \urcorner = \ulcorner \neg A \urcorner$.

We use the shorthand $\phi \dot\rightarrow \psi$ for $(\neg\phi) \dot\vee \psi$, and $\phi \dot\leftrightarrow \psi$ for $(\phi \dot\rightarrow \psi) \dot\wedge (\psi \dot\rightarrow \phi)$, and for sets $u \subseteq \mathcal{G}_{\mathcal{L}^1}$ we write $u[\![x/y]\!]$ for $\{\phi[x/y] \mid \phi \in u\}$, and analogously we write $u[\![f, g]\!]_\infty$ and $u[\![f]\!]_\infty$.

The previous cryptic definition of $\mathcal{G}_{\mathcal{L}^1}$, which explicitly shows that $\mathcal{G}_{\mathcal{L}^1}$ is elementarily definable, does not really reveal the content of $\mathcal{G}_{\mathcal{L}^1}$. Therefore we state the following lemma, which unfolds the recursive definition of the class of Gödel-codes.

Lemma 82.

$\mathsf{NBG} \vdash \forall z(z \in \mathcal{G}_{\mathcal{L}^1} \leftrightarrow A[z])$, where

$$\begin{aligned} A[z] \quad := \quad & \exists p \exists q \exists x \exists y \exists u \exists \phi \exists \psi (z \in \{(\phi \dot\vee \psi), (\phi \dot\wedge \psi)\} \vee \\ & (\dot{\mathsf{u}}_p \notin \mathit{term}(\phi) \wedge z \in \{\dot\exists\, \dot{\mathsf{u}}_p(\phi[\dot{\mathsf{u}}_p/\dot{\mathsf{v}}_q]), \dot\forall\, \dot{\mathsf{u}}_p(\phi[\dot{\mathsf{u}}_p/\dot{\mathsf{v}}_q])\}) \vee \\ & (\dot{\mathsf{U}}_p \notin \mathit{term}(\phi) \wedge z \in \{\dot\exists\, \dot{\mathsf{U}}_p(\phi[\dot{\mathsf{U}}_p/\dot{\mathsf{V}}_q]), \dot\forall\, \dot{\mathsf{U}}_p(\phi[\dot{\mathsf{U}}_p/\dot{\mathsf{V}}_q])\}) \vee \\ & (z \in \{u, \dot\sim u\} \wedge (u \in \{\dot{\mathsf{v}}_p \mathrel{\dot\in} \dot{\mathsf{v}}_q, \dot{\mathsf{v}}_p \mathrel{\dot\in} \dot{\mathsf{c}}_y, \dot{\mathsf{c}}_x \mathrel{\dot\in} \dot{\mathsf{v}}_q, \dot{\mathsf{c}}_x \mathrel{\dot\in} \dot{\mathsf{c}}_y\} \vee \\ & \qquad\qquad\qquad u \in \{\dot{\mathsf{v}}_p \mathrel{\dot\in} \dot{\mathsf{V}}_q, \dot{\mathsf{v}}_p \mathrel{\dot\in} \dot{\mathsf{C}}_y, \dot{\mathsf{c}}_x \mathrel{\dot\in} \dot{\mathsf{V}}_q, \dot{\mathsf{c}}_x \mathrel{\dot\in} \dot{\mathsf{C}}_y\})). \end{aligned}$$

Proof. By the definition of $\mathcal{G}_{\mathcal{L}^1}$. □

We define truth predicates that cover formulas up to some specified formula rank, that is, we can gradually increase the rank for which the formulas are properly reflected by the truth predicate. Further, truth is defined such that class quantifiers range over all classes $(U)_x$ of some class universe U, and such that the class constant $\dot{\mathsf{C}}_x$ stands for the corresponding class $(U)_x$. In contrast to the class quantifiers, the set quantifiers range over the universe of all sets, as usual, and any set constant $\dot{\mathsf{c}}_x$ stands for the set x itself.

The formula $Tr[\alpha, U, V]$ in the following definition means that V is a truth predicate containing all "true" *closed* Gödel-codes $\phi \in \mathcal{G}_{\mathcal{L}^1}$ with rank $rk_(\phi) \leq \alpha'$, and with all class quantifiers and class constants in ϕ interpreted with respect to the class universe U.

Definition 83. (Formalized Truth)

We define $Tr[\alpha, U, V]$ to be the formula,

$$\forall\phi\forall\psi\forall x\forall y\forall p\forall q(rk(\phi) \in \alpha' \wedge rk(\psi) \in \alpha' \rightarrow A),$$

where A is the conjunction of the following formulas

$$\begin{array}{rcl}
(\dot{\mathtt{c}}_x \mathrel{\dot{\in}} \dot{\mathtt{c}}_y) \in V & \leftrightarrow & x \in y, \\
\dot{\sim}(\dot{\mathtt{c}}_x \mathrel{\dot{\in}} \dot{\mathtt{c}}_y) \in V & \leftrightarrow & x \notin y, \\
(\dot{\mathtt{c}}_x \mathrel{\dot{\in}} \dot{\mathtt{C}}_y) \in V & \leftrightarrow & x \in (U)_y, \\
\dot{\sim}(\dot{\mathtt{c}}_x \mathrel{\dot{\in}} \dot{\mathtt{C}}_y) \in V & \leftrightarrow & x \notin (U)_y, \\
(\phi \mathrel{\dot{\vee}} \psi) \in V & \leftrightarrow & \phi \in V \vee \psi \in V, \\
(\phi \mathrel{\dot{\wedge}} \psi) \in V & \leftrightarrow & \phi \in V \wedge \psi \in V, \\
\dot{\exists}\, \dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q]) \in V & \leftrightarrow & \exists z(\phi[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q] \in V) \wedge \dot{\mathtt{u}}_p \notin \mathit{term}(\phi), \\
\dot{\forall}\, \dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q]) \in V & \leftrightarrow & \forall z(\phi[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q] \in V) \wedge \dot{\mathtt{u}}_p \notin \mathit{term}(\phi), \\
\dot{\exists}\, \dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q]) \in V & \leftrightarrow & \exists z(\phi[\dot{\mathtt{C}}_z/\dot{\mathtt{V}}_q] \in V) \wedge \dot{\mathtt{U}}_p \notin \mathit{term}(\phi), \\
\dot{\forall}\, \dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q]) \in V & \leftrightarrow & \forall z(\phi[\dot{\mathtt{C}}_z/\dot{\mathtt{V}}_q] \in V) \wedge \dot{\mathtt{U}}_p \notin \mathit{term}(\phi).
\end{array}$$

The following technical lemma shows that the formalized notion of truth is properly defined, i.e., if $Tr[\alpha, U, V]$ holds then V is essentially unique with respect to α and U. Clearly, the truth predicates also perfectly reflect the complementation of formulas.

Lemma 84.

(1) $\mathsf{NBG} \vdash \forall\alpha(\forall\beta \in \alpha)(Tr[\alpha, U, X] \to Tr[\beta, U, X])$,

(2) $\mathsf{NBG} \vdash \forall\phi(Tr[rk(\phi), U, X] \wedge Tr[rk(\phi), U, Y] \to (\phi[f, g]_\infty \in X \leftrightarrow \phi[f, g]_\infty \in Y))$,

(3) $\mathsf{NBG} \vdash (\forall\phi \in \mathcal{G}_{\Pi^1_0})(Tr[rk(\phi), U, X] \wedge Tr[rk(\phi), V, Y] \to (\sharp[\phi, f_0, g_0, f_1, g_1, U, V] \to (\phi[f_0, g_0]_\infty \in X \leftrightarrow \phi[f_1, g_1]_\infty \in Y))$,

where $\sharp[\phi, f_0, g_0, f_1, g_1, U, V]$ is the formula

$$\begin{aligned}(\forall x)((&\dot{\mathtt{C}}_x \in \mathit{term}(\phi) \to (U)_x = (V)_x) \wedge \\ &(\dot{\mathtt{V}}_x \in \mathit{term}(\phi) \to (U)_{g_0(x)} = (V)_{g_1(x)}) \wedge \\ &(\dot{\mathtt{v}}_x \in \mathit{term}(\phi) \to f_0(x) = f_1(x))),\end{aligned}$$

(4) $\mathsf{NBG} \vdash Tr[rk(\phi), U, Y] \to (\phi[f, g]_\infty \in Y \leftrightarrow (\neg\phi)[f, g]_\infty \notin Y)$.

Proof.

(1) By the definition of $Tr[\alpha, U, X]$, i.e. because of the transitivity of $\in$ on ordinals.

(2) By induction on p we get

$$\begin{aligned}&(\forall p \in \omega)\, Tr[p, U, X] \wedge Tr[p, U, Y] \to \\ &\qquad \forall\phi(rk(\phi) \in p' \to (\phi[f, g]_\infty \in X \leftrightarrow \phi[f, g]_\infty \in Y)).\end{aligned}$$

(3) Analogous to (2).

(4) Analogous to (2). □

Truth predicates for formulas, up to some fixed maximum formula rank, are shown to exist in NBG. If we want to have a truth predicate for all formulas, i.e. if we want the truth predicate to exist as a class, then NBG needs to be extended by iterated elementary comprehension up to ω.

Lemma 85.

(1) $\mathsf{NBG} \cup \exists\, \mathit{Hier}_{\in}^{\omega}[\Pi_0^1] \vdash \forall X \exists Y\, \mathit{Tr}[\omega, X, Y]$,

(2) $\mathsf{NBG} \vdash \forall X \exists Y\, \mathit{Tr}[\overline{n}, X, Y]$.

Proof.

(1) If we define $A[V, z, r, U] := \exists\phi\exists\psi\exists x\exists y\exists p\exists q(z \in \mathcal{G}_{\mathcal{L}^1} \wedge rk(z) = r \wedge B)$, where B is the disjunction of the following formulas

$$\begin{aligned}
&z = (\dot{\mathsf{c}}_x \mathrel{\dot{\in}} \dot{\mathsf{c}}_y) \wedge x \in y,\\
&z = \dot{\sim}(\dot{\mathsf{c}}_x \mathrel{\dot{\in}} \dot{\mathsf{c}}_y) \wedge x \notin y,\\
&z = (\dot{\mathsf{c}}_x \mathrel{\dot{\in}} \dot{\mathsf{C}}_y) \wedge x \in (U)_y,\\
&z = \dot{\sim}(\dot{\mathsf{c}}_x \mathrel{\dot{\in}} \dot{\mathsf{C}}_y) \wedge x \notin (U)_y,\\
&z = (\phi \mathrel{\dot{\vee}} \psi) \wedge (\phi \in \langle V\rangle^r \vee \psi \in \langle V\rangle^r),\\
&z = (\phi \mathrel{\dot{\wedge}} \psi) \wedge \phi \in \langle V\rangle^r \wedge \psi \in \langle V\rangle^r,\\
&z = \dot{\exists}\, \dot{\mathsf{u}}_p(\phi[\dot{\mathsf{u}}_p/\dot{\mathsf{v}}_q]) \wedge \exists u(\phi[\dot{\mathsf{c}}_u/\dot{\mathsf{v}}_q] \in \langle V\rangle^r) \wedge \dot{\mathsf{u}}_p \notin term(\phi),\\
&z = \dot{\forall}\, \dot{\mathsf{u}}_p(\phi[\dot{\mathsf{u}}_p/\dot{\mathsf{v}}_q]) \wedge \forall u(\phi[\dot{\mathsf{c}}_u/\dot{\mathsf{v}}_q] \in \langle V\rangle^r) \wedge \dot{\mathsf{u}}_p \notin term(\phi),\\
&z = \dot{\exists}\, \dot{\mathsf{U}}_p(\phi[\dot{\mathsf{U}}_p/\dot{\mathsf{V}}_q]) \wedge \exists u(\phi[\dot{\mathsf{C}}_u/\dot{\mathsf{V}}_q] \in \langle V\rangle^r) \wedge \dot{\mathsf{U}}_p \notin term(\phi),\\
&z = \dot{\forall}\, \dot{\mathsf{U}}_p(\phi[\dot{\mathsf{U}}_p/\dot{\mathsf{V}}_q]) \wedge \forall u(\phi[\dot{\mathsf{C}}_u/\dot{\mathsf{V}}_q] \in \langle V\rangle^r) \wedge \dot{\mathsf{U}}_p \notin term(\phi),
\end{aligned}$$

then by $\exists\, \mathit{Hier}_{\in}^{\omega}[\Pi_0^1]$ there is some V such that

$$(\forall r \in \omega)\forall z(z \in (V)_r \leftrightarrow A[(V)^r, z, r, U]).$$

By induction on r we get $(\forall r \in \omega)\, \mathit{Tr}[r, U, \langle V\rangle^\omega]$, hence $\mathit{Tr}[\omega, U, \langle V\rangle^\omega]$. See e.g. Jäger and Krähenbühl [10] for a similar construction.

(2) Analogous to (1). We use that $\mathsf{NBG} \vdash \exists\, \mathit{Hier}_{\in}^{\overline{n}}[\Pi_0^1]$ by Theorem 79. □

In Section 10 we extensively use truth predicates, hence the following handy notation will be of great use.

Definition 86.

We define the following abbreviations for formulas,

$$\begin{aligned}
\mathit{True}[\phi, U] &:= \exists Y(\mathit{Tr}[rk(\phi), U, Y] \wedge \phi \in Y),\\
\phi[f, g]_\infty^U &:= \mathit{True}[\phi[f, g]_\infty, U].
\end{aligned}$$

The following two lemmas are just restating some facts in the context of the new notation, i.e. by using $\phi[f, g]_\infty^U$.

Lemma 87.

(1) $\mathsf{NBG} \vdash \forall\phi(\exists Y\, Tr[rk(\phi), U, Y] \rightarrow (\phi[f,g]^U_\infty \leftrightarrow \neg((\neg\phi)[f,g]^U_\infty)))$,

(2) $\mathsf{NBG} \cup \exists\, Hier^\omega_\in[\Pi^1_0] \vdash \forall\phi(\phi[f,g]^U_\infty \leftrightarrow \neg((\neg\phi)[f,g]^U_\infty))$,

(3) $\mathsf{NBG} \vdash \ulcorner A \urcorner[f,g]^U_\infty \leftrightarrow \neg(\ulcorner\neg A\urcorner[f,g]^U_\infty)$.

Proof.

(1) If $\phi[f,g]^U_\infty$ and $\psi = \phi[f,g]_\infty$ then $\exists Y(\, Tr[rk(\psi), U, Y] \wedge \psi \in Y)$, and $\forall Y(\, Tr[rk(\psi), U, Y] \rightarrow \psi \in Y)$ by Lemma 84(2), hence because of $rk(\psi) = rk(\neg\psi)$ and by Lemma 84(4) we get $\forall Y(\, Tr[rk(\neg\psi), U, Y] \rightarrow (\neg\psi) \notin Y)$, that is $\neg((\neg\phi)[f,g]^U_\infty)$. On the other hand, if $\neg((\neg\phi)[f,g]^U_\infty)$ then we have $\forall Y(\, Tr[rk(\neg\psi), U, Y] \rightarrow (\neg\psi) \notin Y)$. And because of $\exists Y\, Tr[rk(\phi), U, Y]$ we get $\exists Y(\, Tr[rk(\psi), U, Y] \wedge (\neg\psi) \notin Y)$, hence by Lemma 84(4) we get $\exists Y(\, Tr[rk(\psi), U, Y] \wedge \psi \in Y)$, that is $\phi[f,g]^U_\infty$.

(2) By Part 1 and Lemma 85(1).

(3) By Part 1 and Lemma 85(2). □

Lemma 88.

$\mathsf{NBG} \vdash \forall\phi\forall\psi\forall p\forall q(A)$,

where A is the conjunction of the following formulas

$$
\begin{array}{rcl}
(\dot{\mathtt{c}}_x \mathrel{\dot\in} \dot{\mathtt{c}}_y)[f,g]^U_\infty & \leftrightarrow & x \in y, \\
\dot\sim(\dot{\mathtt{c}}_x \mathrel{\dot\in} \dot{\mathtt{c}}_y)[f,g]^U_\infty & \leftrightarrow & x \notin y, \\
(\dot{\mathtt{c}}_x \mathrel{\dot\in} \dot{\mathtt{C}}_y)[f,g]^U_\infty & \leftrightarrow & x \in (U)_y, \\
\dot\sim(\dot{\mathtt{c}}_x \mathrel{\dot\in} \dot{\mathtt{C}}_y)[f,g]^U_\infty & \leftrightarrow & x \notin (U)_y, \\
(\phi \mathrel{\dot\vee} \psi)[f,g]^U_\infty & \leftrightarrow & \phi[f,g]^U_\infty \vee \psi[f,g]^U_\infty, \\
(\phi \mathrel{\dot\wedge} \psi)[f,g]^U_\infty & \leftrightarrow & \phi[f,g]^U_\infty \wedge \psi[f,g]^U_\infty, \\
\dot\exists\, \dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])[f,g]^U_\infty & \leftrightarrow & \exists z(\phi[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q][f,g]^U_\infty) \wedge \dot{\mathtt{u}}_p \notin term(\phi), \\
\dot\forall\, \dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])[f,g]^U_\infty & \leftrightarrow & \forall z(\phi[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q][f,g]^U_\infty) \wedge \dot{\mathtt{u}}_p \notin term(\phi), \\
\dot\exists\, \dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])[f,g]^U_\infty & \leftrightarrow & \exists z(\phi[\dot{\mathtt{C}}_z/\dot{\mathtt{V}}_q][f,g]^U_\infty) \wedge \dot{\mathtt{U}}_p \notin term(\phi), \\
\dot\forall\, \dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])[f,g]^U_\infty & \leftrightarrow & \forall z(\phi[\dot{\mathtt{C}}_z/\dot{\mathtt{V}}_q][f,g]^U_\infty) \wedge \dot{\mathtt{U}}_p \notin term(\phi).
\end{array}
$$

Proof. By the definition of $\phi[f,g]_\infty$ and $\phi[f,g]^U_\infty$, and because if we have any Y with $Tr[rk(\phi), U, Y]$ then we get Z with $Tr[rk(\phi)', U, Z]$ by elementary comprehension. □

Corollary 89.

(1) $\mathsf{NBG} \vdash \forall\phi\forall p\forall q(\dot{\forall}\,\dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])[f,g]^U_\infty \to \phi[f,g]^U_\infty)$,

(2) $\mathsf{NBG} \vdash \forall\phi\forall p\forall q(\dot{\mathtt{u}}_p \notin \mathit{term}(\phi) \wedge \phi[f,g]^U_\infty \to \dot{\exists}\,\dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])[f,g]^U_\infty)$,

(3) $\mathsf{NBG} \vdash \forall\phi\forall p\forall q(\dot{\forall}\,\dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])[f,g]^U_\infty \to \phi[f,g]^U_\infty)$,

(4) $\mathsf{NBG} \vdash \forall\phi\forall p\forall q(\dot{\mathtt{U}}_p \notin \mathit{term}(\phi) \wedge \phi[f,g]^U_\infty \to \dot{\exists}\,\dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])[f,g]^U_\infty)$.

Proof.

(1) If we have $\dot{\forall}\,\dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])[f,g]^U_\infty$ then $\forall z(\phi[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q][f,g]^U_\infty)$ by Lemma 88, hence $\phi[\dot{\mathtt{c}}_{f(q)}/\dot{\mathtt{v}}_q][f,g]^U_\infty$, and because of $\phi[f,g]_\infty = \phi[\dot{\mathtt{c}}_{f(q)}/\dot{\mathtt{v}}_q][f,g]_\infty$ we get $\phi[f,g]^U_\infty$.

(2) If we have $\phi[f,g]^U_\infty$ then $\phi[\dot{\mathtt{c}}_{f(q)}/\dot{\mathtt{v}}_q][f,g]^U_\infty$, that is $\exists z(\phi[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q][f,g]^U_\infty)$, and hence $\dot{\exists}\,\dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])[f,g]^U_\infty$ by Lemma 88.

(3) Analogous to (1).

(4) Analogous to (2). $\square$

The most important property of the formalized truth predicate is that it reflects the truth of the original formulas, at least for an appropriate class of formulas.

Lemma 90. (Truth Reflection)

Let $\sharp_D[f,g,U]$ be the conjunction of $\top$, and of all formulas $f(\overline{\sharp x}) = x$, and $(U)_{g(\overline{\sharp Y})} = Y$, for all free variables $x, Y \in D$.

(1) If $A \in \Sigma^1(\Pi^1_0)$ then $\mathsf{NBG} \vdash \sharp_A[f,g,U] \to (\ulcorner A \urcorner[f,g]^U_\infty \to A)$.

(2) If $B \in \Pi^1(\Sigma^1_0)$ then $\mathsf{NBG} \vdash \sharp_B[f,g,U] \to (B \to \ulcorner B \urcorner[f,g]^U_\infty)$.

(3) If $C \in \Pi^1_0$ then $\mathsf{NBG} \vdash \sharp_C[f,g,U] \to (C \leftrightarrow \ulcorner C \urcorner[f,g]^U_\infty)$.

For the reverse implications $A \to \ulcorner A \urcorner[f,g]^U_\infty$ and $\ulcorner B \urcorner[f,g]^U_\infty \to B$ there are easy counter-examples, i.e., for $A := \exists X\forall y(y \in X)$ and $U = \emptyset$ we have $\mathsf{NBG} \vdash A \wedge \neg\exists x\forall y(y \in (\emptyset)_x)$, and for $B := \forall X\forall y(y \in X)$ and $U = \mathcal{V}$ we have $\mathsf{NBG} \vdash \forall x\forall y(y \in (\mathcal{V})_x) \wedge \neg B$.

Proof.

(1) By induction on the structure of A. If A is an atomic formula then the claim follows directly by Lemma 88. If $A = \exists X B[X/V]$ then by i.h. we

have $\ulcorner B \urcorner [f,g]^U_\infty \wedge \sharp_B[f,g,U] \to B$, and each of the following statements is a consequence of the preceding one

$$
\begin{aligned}
&\ulcorner B\urcorner[f,g]^U_\infty \wedge (U)_{g(\overline{\sharp V})} = V \wedge \sharp_A[f,g,U] \to B,\\
&\ulcorner B\urcorner[f,g]^U_\infty \wedge g(\overline{\sharp V}) = x \wedge (U)_x = V \wedge \sharp_A[f,g,U] \to B,\\
&\ulcorner B\urcorner[f,g]^U_\infty \wedge g(\overline{\sharp V}) = x \wedge \exists Y((U)_x = Y) \wedge \sharp_A[f,g,U] \to A,\\
&\ulcorner B\urcorner[f,g]^U_\infty \wedge g(\overline{\sharp V}) = x \wedge \sharp_A[f,g,U] \to A,\\
&\ulcorner B\urcorner[\dot{\mathsf{c}}_x/\dot{V}][f,g]^U_\infty \wedge g(\overline{\sharp V}) = x \wedge \sharp_A[f,g,U] \to A,\\
&\exists h(\exists z(\ulcorner B\urcorner[\dot{\mathsf{c}}_z/\dot{V}][f,h]^U_\infty \wedge h(\overline{\sharp V}) = z) \wedge \sharp_A[f,h,U]) \to A,\\
&\exists h(\exists z(\ulcorner B\urcorner[\dot{\mathsf{c}}_z/\dot{V}][f,h]^U_\infty) \wedge \sharp_A[f,h,U]) \to A,\\
&\ulcorner A\urcorner[f,g]^U_\infty \wedge \sharp_A[f,g,U] \to A.
\end{aligned}
$$

If $A = \forall x B[x/v]$ then we have $\ulcorner B\urcorner[f,g]^U_\infty \wedge \sharp_B[f,g,U] \to B$ by i.h., hence $\ulcorner B\urcorner[f,g]^U_\infty \wedge \sharp_A[f,g,U] \to \forall z(f(\overline{\sharp v}) = z \to B[z/v])$, that is $\forall h(\ulcorner A\urcorner[h,g]^U_\infty \wedge \sharp_A[h,g,U] \to B[h(\overline{\sharp v})/v])$ by Corollary 89(1). If we assume there is some f such that $\ulcorner A\urcorner[f,g]^U_\infty \wedge \sharp_A[f,g,U]$ then for any f_0 with $\forall p(p = \overline{\sharp v} \vee f_0(p) = f(p))$ we have that $\ulcorner A\urcorner[f_0,g]^U_\infty \wedge \sharp_A[f_0,g,U]$, hence $B[f_0(\overline{\sharp v})/v])$. And because $f_0(\overline{\sharp v})$ can be any set, we have shown $\exists h(\ulcorner A\urcorner[h,g]^U_\infty \wedge \sharp_A[h,g,U]) \to A$, that is $\ulcorner A\urcorner[f,g]^U_\infty \wedge \sharp_A[f,g,U] \to A$. Other cases of A are shown similarly and mainly by using Lemma 88.

(2) By (1) and Lemma 87(3), and because $\Sigma^1(\Pi^1_0) = \neg\Pi^1(\Sigma^1_0)$.

(3) By (1) and (2), and because $\Pi^1_0 \subseteq \Sigma^1(\Pi^1_0) \cap \Pi^1(\Sigma^1_0)$. □

The definition of the infinitary proof system and the corresponding proof predicate (within NBG) is in analogy to the formal proofs on the meta-level, i.e., Definition 12. We formalize Definition 12 within NBG, in almost complete analogy, except for the infinitary rule for the universal set quantifiers. By infinitary proofs we derive finite sets of Gödel-codes $z \subseteq \mathcal{G}^-_{\mathcal{L}^1}$. The inference rules for infinitary proofs consist of the common rules for classical logic, and the rules for some axioms $X \subseteq \mathcal{G}^-_{\mathcal{L}^1}$, and additional inference rules specified by some class $Y \subseteq \mathcal{G}^-_{\mathcal{L}^1}$. The rules can be depicted as follows (where $z \subseteq \mathcal{G}^-_{\mathcal{L}^1}$ is any finite set)

$$z \cup \{\phi, \neg\phi\} \quad \phi \in \mathcal{G}^-_{\mathcal{L}^1} \text{ atomic},$$

$$\frac{z \cup \{\phi\}}{z \cup \{\phi \mathbin{\dot\vee} \psi\}}, \qquad \frac{z \cup \{\psi\}}{z \cup \{\phi \mathbin{\dot\vee} \psi\}},$$

$$\frac{z \cup \{\phi\} \qquad z \cup \{\psi\}}{z \cup \{\phi \mathbin{\dot\wedge} \psi\}},$$

$$\frac{z \cup \{\phi[\dot{\mathrm{v}}_t/\dot{\mathrm{v}}_q]\}}{z \cup \{\dot\exists\, \dot{\mathrm{u}}_p(\phi[\dot{\mathrm{u}}_p/\dot{\mathrm{v}}_q])\}}, \qquad \frac{z \cup \{\phi\}}{z \cup \{\dot\forall\, \dot{\mathrm{u}}_p(\phi[\dot{\mathrm{u}}_p/\dot{\mathrm{v}}_q])\}} \quad \dot{\mathrm{v}}_q \notin \bigcup_{\xi \in z} term(\xi),$$

$$\frac{z \cup \{\phi[\dot{\mathtt{V}}_t/\dot{\mathtt{V}}_q]\}}{z \cup \{\dot{\exists}\,\dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])\}}, \qquad \frac{z \cup \{\phi\}}{z \cup \{\dot{\forall}\,\dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])\}} \quad \dot{\mathtt{V}}_q \notin \bigcup_{\xi \in z} term(\xi),$$

$$\frac{z \cup \{\phi\} \qquad z \cup \{\neg\phi\}}{z},$$

$$z \cup \{\phi\} \quad \text{with } \phi \in X,$$

$$\frac{z \cup \{\psi\}}{z \cup \{\phi\}} \quad \text{with } (\psi \dot{\to} \phi) \in Y,$$

$$z \cup \{\dot{\mathtt{c}}_u \mathrel{\dot{\in}} \dot{\mathtt{c}}_v\} \quad \text{with } u \in v, \qquad z \cup \{\dot{\sim}(\dot{\mathtt{c}}_u \mathrel{\dot{\in}} \dot{\mathtt{c}}_v)\} \quad \text{with } u \notin v,$$

$$\frac{z \cup \{\phi[\dot{\mathtt{c}}_w/\dot{\mathtt{v}}_q]\}}{z \cup \{\dot{\exists}\,\dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\}}, \qquad \frac{z \cup \{\phi[\dot{\mathtt{c}}_w/\dot{\mathtt{v}}_q]\} \text{ for all sets } w}{z \cup \{\dot{\forall}\,\dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\}}.$$

In the following, i.e., by the formula $Pr^{+}_{\Omega_n}[U,X,Y,Z]$, we define the provability relation U, which captures provability by such infinitary proofs. In addition to the axioms $X \subseteq \mathcal{G}_{\mathcal{L}^1}$ and the additional inference rules $Y \subseteq \mathcal{G}_{\mathcal{L}^1}$, the relation also has control parameters $\mathfrak{a}$ and r, such that $z \in (U)_{\mathfrak{a},r}$ essentially means that

(1) there is a derivation of the finite set $z \subseteq \mathcal{G}^{-}_{\mathcal{L}^1}$, which possibly uses axioms in X and additional inference rules from Y,

(2) this derivation takes at most $\mathfrak{a} \lhd \Omega_n$ steps (by definition we have $\emptyset \lhd \mathfrak{a}$),

(3) the rank of any cut in this derivation is at most r with respect to $Z \subseteq \mathcal{G}_{\mathcal{L}^1}$, that is, any "cut-formula" ϕ is such that $rk_Z(\phi) \leq r$. If $r = \emptyset$ then $\phi \in Z$ or $\neg\phi \in Z$.

The elementary formulas A and A^{+} in the following definition, are such that $Pr^{+}_{\Omega_n}[U,X,Y,Z] = Hier^{\Omega_n}_{\lhd}[A^{+}[U,x,\mathfrak{b},X,Y,Z]]$ and $Pr_{\Omega_n}[U,X,Y,Z] = Hier^{\Omega_n}_{\lhd}[A[U,x,\mathfrak{b},X,Y,Z]]$, where Pr_{Ω_n} is just $Pr^{+}_{\Omega_n}$ without the common inference rules for *set quantifiers*, e.g., in Pr_{Ω_n} only the infinitary rule for the universal set quantifers is included. E.g., if $Pr_{\Omega_n}[U,X,Y,Z]$ then for any $\mathfrak{b} \lhd \Omega_n$ we have

$$(U)_{\mathfrak{b}} = \{x \mid A[(U)^{\lhd,\mathfrak{b}}, x, \mathfrak{b}, X, Y, Z]\}.$$

$Pr^{\emptyset}_{\Omega_n}[U]$, $Pr^{DC}_{\Omega_n}[U]$, and $Pr^{\emptyset+}_{\Omega_n}[U]$, $Pr^{DC+}_{\Omega_n}[U]$, in the following definition, are instances of Pr_{Ω_n} and $Pr^{+}_{\Omega_n}$, respectively, with specified classes X, Y, Z.

Definition 91. (Formalized Proof)

Let $A[U,x,\mathfrak{b},X,Y,Z]$ be the formula

$$\mathfrak{b} \neq \emptyset \wedge \exists\phi\exists\psi\exists y\exists z\exists u\exists v\exists p\exists q\exists r\exists s\exists t(y \subseteq_{\mathsf{f}} \mathcal{G}^{-}_{\mathcal{L}^1} \wedge r \in s' \wedge x = \langle y, s\rangle \wedge B),$$

where $y \subseteq_{\mathsf{f}} V$ means that y is a *finite* subset of V, i.e., $y \subseteq_{\mathsf{f}} V$ stands for $y \subseteq V \wedge \exists g(Fun[g] \wedge dom(g) \in \omega \wedge ran(g) = y)$, and where B is the formula

$$\begin{array}{l}
\{\phi, \dot{\sim}\phi\} \subseteq y \vee \\
(\phi \in y \wedge \phi \in X) \vee \\
(y = z \cup \{\phi\} \wedge \psi \dot{\rightarrow} \phi \in Y \wedge (\exists \mathfrak{a} \lhd \mathfrak{b}) z \cup \{\psi\} \in (U)_{\mathfrak{a},r}) \vee \\
(y = z \cup \{\phi \dot{\vee} \psi\} \wedge (\exists \mathfrak{a} \lhd \mathfrak{b})(z \cup \{\phi\} \in (U)_{\mathfrak{a},r} \vee z \cup \{\psi\} \in (U)_{\mathfrak{a},r})) \vee \\
(y = z \cup \{\phi \dot{\wedge} \psi\} \wedge (\exists \mathfrak{a} \lhd \mathfrak{b})(z \cup \{\phi\} \in (U)_{\mathfrak{a},r} \wedge z \cup \{\psi\} \in (U)_{\mathfrak{a},r})) \vee \\
(y = z \cup \{\dot{\exists}\, \dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])\} \wedge (\exists \mathfrak{a} \lhd \mathfrak{b}) z \cup \{\phi[\dot{\mathtt{V}}_t/\dot{\mathtt{V}}_q]\} \in (U)_{\mathfrak{a},r}) \vee \\
(y = z \cup \{\dot{\forall}\, \dot{\mathtt{U}}_p(\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])\} \wedge (\forall \xi \in z) \dot{\mathtt{V}}_q \notin term(\xi) \wedge \\
\hfill (\exists \mathfrak{a} \lhd \mathfrak{b}) z \cup \{\phi\} \in (U)_{\mathfrak{a},r}) \vee \\
((\exists \mathfrak{a} \lhd \mathfrak{b})(y \cup \{\phi\} \in (U)_{\mathfrak{a},r} \wedge y \cup \{\neg\phi\} \in (U)_{\mathfrak{a},r}) \wedge rk_Z(\phi) \in s') \vee \\
((\dot{\mathtt{c}}_u \mathrel{\dot{\in}} \dot{\mathtt{c}}_v) \in y \wedge u \in v) \vee \\
(\dot{\sim}(\dot{\mathtt{c}}_u \mathrel{\dot{\in}} \dot{\mathtt{c}}_v) \in y \wedge u \notin v) \vee \\
(y = z \cup \{\dot{\exists}\, \dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\} \wedge \exists w (\exists \mathfrak{a} \lhd \mathfrak{b}) z \cup \{\phi[\dot{\mathtt{c}}_w/\dot{\mathtt{v}}_q]\} \in (U)_{\mathfrak{a},r}) \vee \\
(y = z \cup \{\dot{\forall}\, \dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\} \wedge \forall w (\exists \mathfrak{a} \lhd \mathfrak{b}) z \cup \{\phi[\dot{\mathtt{c}}_w/\dot{\mathtt{v}}_q]\} \in (U)_{\mathfrak{a},r}).
\end{array}$$

We define the formula $Pr_{\Omega_n}[U, X, Y, Z] := Hier_{\lhd}^{\Omega_n}[A[U, x, \mathfrak{b}, X, Y, Z]]$, and we define $Pr^{\emptyset}_{\Omega_n}$, and $Pr^{DC}_{\Omega_n}$, with specific X, Y, Z, such that

$$\begin{array}{rcl}
Pr^{\emptyset}_{\Omega_n}[U] & := & Pr_{\Omega_n}[U, \emptyset, \emptyset, \emptyset], \\
Pr^{DC}_{\Omega_n}[U] & := & Pr_{\Omega_n}[U, \mathcal{G}_{CA[\Pi^1_0]}, \mathcal{G}_{DC_{\mathcal{O}n}[\Pi^1_0]}, \mathcal{G}_{CA[\Pi^1_0]} \cup \mathcal{G}^{\rightarrow}_{DC_{\mathcal{O}n}[\Pi^1_0]}],
\end{array}$$

where the expression $\mathcal{G}^{\rightarrow}_{\mathcal{C}}$ is defined (analogous Definition 18) such that

$$z \in \mathcal{G}^{\rightarrow}_{\mathcal{C}} \quad := \quad \exists \phi((\phi \dot{\rightarrow} z) \in \mathcal{G}_{\mathcal{C}}).$$

We define A^+ to be the formula A, with B replaced by

$$\begin{array}{l}
B \vee \\
y = z \cup \{\dot{\exists}\, \dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\} \wedge (\exists \mathfrak{a} \lhd \mathfrak{b}) z \cup \{\phi[\dot{\mathtt{v}}_t/\dot{\mathtt{v}}_q]\} \in (U)_{\mathfrak{a},r} \vee \\
y = z \cup \{\dot{\forall}\, \dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\} \wedge (\forall \xi \in z) \dot{\mathtt{v}}_q \notin term(\xi) \wedge \\
\hfill (\exists \mathfrak{a} \lhd \mathfrak{b}) z \cup \{\phi\} \in (U)_{\mathfrak{a},r}.
\end{array}$$

Finally, the formulas $Pr^{+}_{\Omega_n}$, $Pr^{\emptyset+}_{\Omega_n}$, $Pr^{DC+}_{\Omega_n}$, are defined analogous to Pr_{Ω_n}, $Pr^{\emptyset}_{\Omega_n}$, and $Pr^{DC}_{\Omega_n}$, respectively, with A replaced by A^+.

The proof for the existence of such proof predicates is straight forward, because of the inductive definition of the predicates.

Lemma 92.

If A is the formula Pr_{Ω_n} or $Pr^{+}_{\Omega_n}$ then

$$\mathsf{NBG} \cup \exists\, Hier_{\lhd}^{\Omega_n}[\Pi^1_0] \vdash \exists U A[U, X, Y, Z].$$

Clearly, the proof predicates $Pr^{\emptyset+}_{\Omega_n}$ and $Pr^{DC+}_{\Omega_n}$ are extensions of $Pr^{\emptyset}_{\Omega_n}$ and $Pr^{DC}_{\Omega_n}$, respectively, because we just added the common finitary versions of the quantifier rules for sets to the infinitary versions. The following lemma also shows that these extensions prove essentially the same as the basic systems, i.e. with respect to set-closed formulas.

Lemma 93.

If A and B are the formulas $Pr^{\emptyset}_{\Omega_n}$ and $Pr^{\emptyset+}_{\Omega_n}$, or $Pr^{DC}_{\Omega_n}$ and $Pr^{DC+}_{\Omega_n}$, respectively, then

(1) $\mathsf{NBG} \vdash A[U] \wedge B[V] \rightarrow (\forall \mathfrak{a} \lhd \Omega_n)(U)_{\mathfrak{a}} \subseteq (V)_{\mathfrak{a}}$,

(2) $\mathsf{NBG} \vdash B[U] \wedge A[V] \rightarrow (\forall \mathfrak{a} \lhd \Omega_n)\forall r(y \in (U)_{\mathfrak{a},r} \rightarrow y[\![f]\!]_\infty \in (V)_{\mathfrak{a},r})$.

Proof.

(1) The statement follows by induction on $\mathfrak{a}$, i.e. we use $TI^{\Omega_n}_{\lhd}[\Pi^1_0]$ by Theorem 74. For the induction we distinguish the cases of y in Definition 91, i.e. $y \in (U)_{\mathfrak{a},r}$. The "embedding" is trivial because all cases of y in $Pr^{\emptyset}_{\Omega_n}$, $Pr^{DC}_{\Omega_n}$, also occur in $Pr^{\emptyset+}_{\Omega_n}$, $Pr^{DC+}_{\Omega_n}$, respectively.

(2) Analogous to Part 1, but here we need to "embed" $Pr^{\emptyset+}_{\Omega_n}$, $Pr^{DC+}_{\Omega_n}$, into $Pr^{\emptyset}_{\Omega_n}$, $Pr^{DC}_{\Omega_n}$, respectively. The only nontrivial cases are the "quantifier rules" for "set variables". E.g. if we have $y = z \cup \{\dot{\forall}\, \dot{\mathtt{u}}_p(\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\}$ and $z \cup \{\phi\} \in (U)_{\mathfrak{b},s}$ with $\mathfrak{b} \lhd \mathfrak{a}$ and $s \in r'$ and $\dot{\mathtt{v}}_q \notin \psi$ for all $\psi \in z$ then by i.h. we have that $z[\![g]\!]_\infty \cup \{\phi[g]_\infty\} \in (V)_{\mathfrak{b},s}$ for all g with $t \neq q \rightarrow f(t) = g(t)$, hence $z[\![f]\!]_\infty \cup \{\phi[\dot{\mathsf{c}}_{g(q)}/\dot{\mathtt{v}}_q][f]_\infty\} \in (V)_{\mathfrak{b},s}$ for any such g, that is $\forall w(z[\![f]\!]_\infty \cup \{\phi[\dot{\mathsf{c}}_w/\dot{\mathtt{v}}_q][f]_\infty\} \in (V)_{\mathfrak{b},s})$, and finally we have $y[\![f]\!]_\infty \in (V)_{\mathfrak{a},r}$. $\square$

Structural properties analogous to the properties of formal proofs in Lemma 14 can be formalized for the proof predicates in NBG, as it is shown in the next lemma.

Lemma 94. (Structural Properties)

(1) If A is the formula Pr_{Ω_n} or $Pr^{+}_{\Omega_n}$ then

$$\begin{aligned}\mathsf{NBG} \vdash (\forall \mathfrak{a} \lhd \Omega_n)\forall r(&X_0 \subseteq X_1 \wedge Y_0 \subseteq Y_1 \wedge Z_0 \subseteq Z_1 \wedge \\ &A[U, X_0, Y_0, Z_0] \wedge A[V, X_1, Y_1, Z_1] \wedge \\ &\mathfrak{b} \unlhd \mathfrak{a} \wedge s \in r' \wedge y \subseteq_{\mathsf{f}} \mathcal{G}^{-}_{\mathcal{L}^1} \wedge x \in (U)_{\mathfrak{b},s} \rightarrow x \cup y \in (V)_{\mathfrak{a},r}),\end{aligned}$$

(2) If A is the formula $Pr^{\emptyset}_{\Omega_n}$, $Pr^{\emptyset+}_{\Omega_n}$, $Pr^{DC}_{\Omega_n}$ or $Pr^{DC+}_{\Omega_n}$ then

$$\begin{aligned}\mathsf{NBG} \vdash (\forall \mathfrak{a} \lhd \Omega_n)\forall p \forall q \forall r(&A[U] \wedge x \in (U)_{\mathfrak{a},r} \rightarrow \\ &\{x[\![\dot{\mathtt{v}}_p/\dot{\mathtt{v}}_q]\!], x[\![\dot{\mathtt{V}}_p/\dot{\mathtt{V}}_q]\!], x[\![\dot{\mathsf{c}}_z/\dot{\mathtt{v}}_q]\!], x[\![f]\!]_\infty\} \subseteq (U)_{\mathfrak{a},r}).\end{aligned}$$

Proof. The proofs of the two parts of this lemma are by induction on $\mathfrak{a}$, i.e. we use $TI_{\lhd}^{\Omega_n}[\Pi^1_0]$ by Theorem 74. The proofs are analogous to formal versions of the proofs of Lemma 14(1) and 14(2). $\square$

For the proof predicates we also have the complete analogon of Theorem 16 about partial cut-elimination.

Theorem 95. (Partial Cut-Elimination)

If A is the formula $Pr^{\emptyset}_{\Omega_{n+m}}$ or $Pr^{DC}_{\Omega_{n+m}}$, and ω^x_0, ω^x_{k+1} stand for x, $\widehat{\omega}^{\omega^x_k}$, respectively, then

$$\mathsf{NBG} \vdash A[U] \to (\forall \mathfrak{a} \lhd \Omega_n)(U)_{\mathfrak{a},\overline{m}} \subseteq (U)_{\omega^{\mathfrak{a}}_m,\emptyset}.$$

Proof. The statement is proved analogous to Corollary 17, i.e. by a formal "infinitary" version of the proof of Corollary 17. We need appropriate versions of Theorem 16 and of some parts of Lemma 14 for this proof, too, but the proofs for these statements are in complete analogy to the ones already given for Theorem 16 and Lemma 14. The infinitary rule for quantification over sets is handled as usual for such infinitary systems. The tedious technical details are left to the reader. $\square$

The following lemma shows that full induction $TI_{\in}[\mathcal{L}^1]$ is provable without using any cuts in the infinitary system, that is, the Gödel-codes of all formulas in $TI_{\in}[\mathcal{L}^1]$ are derivable in $Pr^{\emptyset}_{\Omega_n}$.

Lemma 96.

$$\begin{aligned}\mathsf{NBG} \vdash Pr^{\emptyset}_{\Omega_n}[U] \to (\forall \phi \in \mathcal{G}^-_{\mathcal{L}^1})\forall\psi\forall r(&\\ &\psi = (\dot\forall\, \dot{\mathrm{u}}_p(\dot\forall\, \dot{\mathrm{u}}_q \mathrel{\dot\in} \dot{\mathrm{u}}_p)\phi[\dot{\mathrm{u}}_q/\dot{\mathrm{v}}_r] \mathrel{\dot\to} \phi[\dot{\mathrm{u}}_p/\dot{\mathrm{v}}_r]) \to\\ &\{\neg\psi, \dot\forall\, \dot{\mathrm{u}}_p\phi[\dot{\mathrm{u}}_p/\dot{\mathrm{v}}_r]\} \in (U)_{\Omega,\emptyset}).\end{aligned}$$

Proof. We assume $Pr^{\emptyset}_{\Omega_n}[U]$ and $\psi = \dot\forall\, \dot{\mathrm{u}}_p(\dot\forall\, \dot{\mathrm{u}}_q \mathrel{\dot\in} \dot{\mathrm{u}}_p)\phi[\dot{\mathrm{u}}_q/\dot{\mathrm{v}}_r] \mathrel{\dot\to} \phi[\dot{\mathrm{u}}_p/\dot{\mathrm{v}}_r]$ and we show $\forall z\{\neg\psi, \phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r]\} \in (U)_{\widehat{\omega\cdot\rho(z)+\overline{2}},\emptyset}$ by set induction on z, where $\rho(z)$ is the same set-theoretic rank as in the proof of Lemma 29. We have $\{\neg\phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r], \phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r]\} \in (U)_{\widehat{\omega},\emptyset}$ for any z, and for any $z \neq \emptyset$ we have that

$$\begin{aligned}
&(\forall x \in z)\{\neg\psi, \phi[\dot{\mathrm{c}}_x/\dot{\mathrm{v}}_r], \phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r]\} \in (U)_{\widehat{\omega\cdot\rho(x)+\overline{2}},\emptyset}, \qquad \text{by i.h.}\\
&(\forall x \in z)\{\neg\psi, \dot\sim(\dot{\mathrm{c}}_x \mathrel{\dot\in} \dot{\mathrm{c}}_z) \mathrel{\dot\vee} \phi[\dot{\mathrm{c}}_x/\dot{\mathrm{v}}_r], \phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r]\} \in (U)_{\widehat{\omega\cdot\rho(x)+\overline{3}},\emptyset},\\
&(\forall x \notin z)\{\neg\psi, \dot\sim(\dot{\mathrm{c}}_x \mathrel{\dot\in} \dot{\mathrm{c}}_z), \phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r]\} \in (U)_{\widehat{1},\emptyset},\\
&(\forall x \notin z)\{\neg\psi, \dot\sim(\dot{\mathrm{c}}_x \mathrel{\dot\in} \dot{\mathrm{c}}_z) \mathrel{\dot\vee} \phi[\dot{\mathrm{c}}_x/\dot{\mathrm{v}}_r], \phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r]\} \in (U)_{\widehat{2},\emptyset},\\
&\{\neg\psi, (\dot\forall\, \dot{\mathrm{u}}_q \mathrel{\dot\in} \dot{\mathrm{c}}_z)\phi[\dot{\mathrm{u}}_q/\dot{\mathrm{v}}_r], \phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r]\} \in (U)_{\widehat{\omega\cdot\rho(z)},\emptyset},\\
&\{\neg\psi, (\dot\forall\, \dot{\mathrm{u}}_q \mathrel{\dot\in} \dot{\mathrm{c}}_z)\phi[\dot{\mathrm{u}}_q/\dot{\mathrm{v}}_r] \mathrel{\dot\wedge} \neg\phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r], \phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r]\} \in (U)_{\widehat{\omega\cdot\rho(z)+\overline{1}},\emptyset},\\
&\{\neg\psi, \phi[\dot{\mathrm{c}}_z/\dot{\mathrm{v}}_r]\} \in (U)_{\widehat{\omega\cdot\rho(z)+\overline{2}},\emptyset}.
\end{aligned}$$

A similar argument works for $z = \emptyset$ because of

$$\forall x\{\neg\psi, \dot{\sim}(\dot{\mathsf{c}}_x \mathrel{\dot{\in}} \dot{\mathsf{c}}_\emptyset), \phi[\dot{\mathsf{c}}_\emptyset/\dot{\mathsf{v}}_r]\} \in (U)_{\widehat{1},\emptyset}.$$

Finally we get $\{\neg\psi, \dot{\forall}\,\dot{\mathsf{u}}_p\phi[\dot{\mathsf{u}}_p/\dot{\mathsf{v}}_r]\} \in (U)_{\Omega,\emptyset}$ because of $\widehat{\alpha} \lhd \Omega$ for any α. $\square$

We end this section by showing that the defined proof predicates really correspond to the formal proofs on the meta-level, i.e. we show that the formal proofs are easily embedded into the proof predicates. The next lemma is used in the proof of the subsequent theorem; together with Lemma 90, the next lemma shows that all Gödel-codes of sentences in $\mathcal{L}^0$ that are provable in NBG, are also derivable in $Pr^{\emptyset+}_{\Omega_k}$, e.g. all axioms of ZFC.

Lemma 97.

$$\mathsf{NBG} \vdash Pr^{\emptyset}_{\Omega_n}[U] \to (\forall\phi \in \mathcal{G}_{\mathcal{L}^0})\forall p(rk(\phi) \in p \wedge \phi[f,g]^V_\infty \to \{\phi[f]_\infty\} \in (U)_{\widehat{p},\emptyset}).$$

Proof. The claim is proved by induction on p. By considering all different cases of ϕ, we have that the base cases of the induction and the induction step follow almost immediately by applying Lemma 88 and the induction hypothesis. E.g. if $\phi = \dot{\forall}\,\dot{\mathsf{u}}_q(\psi[\dot{\mathsf{u}}_q/\dot{\mathsf{v}}_r])$ and $\phi[f,g]^V_\infty$ then $\forall z(\psi[\dot{\mathsf{c}}_z/\dot{\mathsf{v}}_r][f,g]^V_\infty)$ by Lemma 88, hence we get $\forall z\{\psi[\dot{\mathsf{c}}_z/\dot{\mathsf{v}}_r][f]_\infty\} \in (U)_{\widehat{rk(\phi)},\emptyset}$ by i.h., and further $\{\phi[f]_\infty\} \in (U)_{\widehat{p},\emptyset}$. $\square$

Theorem 98.

(1) If $\mathsf{ZFC} \,|\!\frac{n,m}{i}\, \Gamma$ then $\mathsf{NBG} \vdash Pr^{\emptyset+}_{\Omega_k}[U] \to \ulcorner\Gamma\urcorner \in (U)_{\widehat{n+m},\bar{i}}$.

(2) If $CA[\Pi^1_0] \cup TI_\in[\mathcal{L}^1]\ [DC_{\mathcal{O}n}[\Pi^1_0]] |\!\frac{n}{i}\, \Gamma$ then
$$\mathsf{NBG} \vdash Pr^{DC+}_{\Omega_{k+1}}[U] \to \ulcorner\Gamma\urcorner \in (U)_{\Omega\hat{+}\widehat{n+1},\bar{i}}.$$

Proof.

(1) By induction on n, considering all cases in Definition 12. We have that $Pr^{\emptyset+}_{\Omega_k}$ directly implements all cases from Definition 12, hence the proof by induction is straightforward. The only nontrivial case is $\Gamma = \Gamma_0, A$ with $A \in \mathsf{ZFC} \subseteq \mathcal{L}^0$. In this case we have $\mathsf{rk}(A) \leq m$, and $\ulcorner A\urcorner[f]_\infty = \ulcorner A\urcorner$ because A is a closed formula, and $\ulcorner A\urcorner[f,g]^V_\infty$ by Lemma 90(3), hence $\{\ulcorner A\urcorner\} \in (U)_{\widehat{m+1},\emptyset}$ by Lemma 97, that is $\ulcorner\Gamma\urcorner \in (U)_{\widehat{n+m},\bar{i}}$ by Lemma 94(1) because of $n > 0$.

(2) Analogously to Part 1. The only nontrivial case here is $\Gamma = \Gamma_0, A$ with $A \in TI_\in[\mathcal{L}^1]$. In this case we have $\{\ulcorner A\urcorner\} \in (U)_{\Omega\hat{+}\hat{2},\emptyset}$ by Lemma 96, that is $\ulcorner\Gamma\urcorner \in (U)_{\Omega\hat{+}\widehat{n+1},\bar{i}}$ by Lemma 94(1) because of $n > 0$. $\square$

9. Cumulative Hierarchies of Classes

In this section we introduce the cumulative hierarchy of classes which is used for the asymmetric interpretation in Section 10. The hierarchy is defined by induction, and any level of the hierarchy contains all classes that can be built by elementary comprehension from classes of lower levels of the hierarchy. Further, any level of the hierarchy contains all previous stages of the hierarchy as classes, i.e. the cumulative hierarchy of classes is some sort of constructible universe.

Definition 99. (Abbreviations)

We define the following shorthand notations

$$\begin{array}{rcl} X \mathrel{\mathring{\in}} Y & := & \exists u(X = (Y)_u), \\ X \mathrel{\mathring{\subseteq}} Y & := & \forall u((X)_u \mathrel{\mathring{\in}} Y). \end{array}$$

We write $(\forall X \mathrel{\mathring{\in}} Y)A[X]$ and $(\exists X \mathrel{\mathring{\in}} Y)A[X]$ for $\forall X(X \mathrel{\mathring{\in}} Y \to A[X])$ and $\exists X(X \mathrel{\mathring{\in}} Y \wedge A[X])$, respectively.

The class hierarchy U, e.g., defined by the formula $Cl_{\Omega_n}[U, V, W]$ below, is such that $(U)_\emptyset = W$, and for all levels $(U)_\mathfrak{a}$, $(U)_\mathfrak{b}$, of the hierarchy, with $\mathfrak{b} \lhd \mathfrak{a} \lhd \Omega_n$, we have that, $(U)_\mathfrak{b} \mathrel{\mathring{\in}} (U)_\mathfrak{a}$ (i.e., $(U)_\mathfrak{b} = ((U)_\mathfrak{a})_\mathfrak{b}$), and $(U)_\mathfrak{b} \mathrel{\mathring{\subseteq}} (U)_\mathfrak{a}$ (i.e., $((U)_\mathfrak{b})_x = ((U)_\mathfrak{a})_{\langle\!\langle \mathfrak{b}, x \rangle\!\rangle}$). That is, all classes contained in lower levels of the hierarchy can be accessed in a uniform way in higher levels of the hierarchy. Further, we have $\{z \mid \phi[\dot{\mathsf{c}}_z/\dot{\mathsf{v}}_p][f,g]_\infty^{(U)_\mathfrak{b}}\} \mathrel{\mathring{\in}} (U)_\mathfrak{a}$ for every $p \in \omega$, f, g, and Gödel-code $\phi \in \mathcal{G}_{\mathcal{L}^1}$ (i.e., $\{z \mid \phi[\dot{\mathsf{c}}_z/\dot{\mathsf{v}}_p][f,g]_\infty^{(U)_\mathfrak{b}}\} = ((U)_{\mathfrak{b} \hat{+} \hat{1}})_{\langle\!\langle \phi, p, f, g \rangle\!\rangle}$). Clearly, to be able to inductively build such a hierarchy U, we need to simultaneously build some truth predicates $(V)_\mathfrak{b}$ with respect to the class universes $(U)_\mathfrak{b}$, that is, U and V are such that $Tr[\omega, (U)_\mathfrak{b}, (V)_\mathfrak{b}]$ for all $\mathfrak{b} \lhd \Omega_n$. For the construction of the level $(U)_\mathfrak{a}$ of the hierarchy we actually use $\widehat{\omega} \mathbin{\hat{\cdot}} (\widehat{2} \mathbin{\hat{\cdot}} \mathfrak{a})$ stages (see the proof of Lemma 101), because we need at least ω extra steps for the construction of each truth predicate $(V)_\mathfrak{b}$ with $\mathfrak{b} \lhd \mathfrak{a}$. In case of $Cl_n[\overline{m}, U, V, W]$, we restrict the truth predicates to Gödel-codes with rank at most m, hence we only need $m + 1$ extra steps for the construction of each truth predicate, and in this case the existence of the hierarchy U can be proved in NBG (by Theorem 79).

The formula A in the following definition is such that for the class hierarchy U with $Cl_{\Omega_n}[U,V,W]$, and for all $\mathfrak{b} \lhd \Omega_n$, we have that

$$(U)_{\mathfrak{b}} = \{x \mid A[(U)^{\lhd,\mathfrak{b}}, x, \mathfrak{b}, V, W]\},$$

where $(V)_{\mathfrak{b}}$ is the truth predicate with respect to the class universe $(U)_{\mathfrak{b}}$.

Definition 100.

Let $A[U,x,\mathfrak{b},V,W]$ be the formula

$$\begin{array}{l}(\mathfrak{b} = \emptyset \wedge x \in W) \vee \\ (\exists \mathfrak{a} \lhd \mathfrak{b}) \exists z (z \in (U)_{\mathfrak{a}} \wedge x = \langle z, \mathfrak{a} \rangle) \vee \\ (\exists \mathfrak{a} \lhd \mathfrak{b}) \exists y \exists z (\langle z, y \rangle \in (U)_{\mathfrak{a}} \wedge x = \langle z, \langle \mathfrak{a}, y \rangle \rangle) \vee \\ \exists \mathfrak{a} \exists z \exists f \exists g \exists \phi \exists p (\mathfrak{a} \mathbin{\hat{+}} \widehat{1} = \mathfrak{b} \wedge \\ \qquad\qquad x = \langle z, \langle\!\langle \phi, p, f, g \rangle\!\rangle \rangle \wedge \phi[\dot{\mathsf{c}}_z / \dot{\mathsf{v}}_p][f,g]_\infty \in (V)_{\mathfrak{a}}),\end{array}$$

and let $B[U,x,r,V,W]$ be the formula A with $\mathfrak{a}, \mathfrak{b}, \lhd, \mathfrak{a} \mathbin{\hat{+}} \widehat{1}$, replaced by $q, r, \in, q'$, respectively. We define the formulas Cl_n, Cl_{Ω^n} and Cl_{Ω_n} such that

$$\begin{array}{rcl} Cl_n[\alpha, U, V, W] & := & \mathit{Hier}_{\in}^{\overline{n}}[B[U,x,r,V,W]] \wedge \\ & & (\forall r \in \overline{n})\, \mathit{Tr}[\alpha, (U)_r, (V)_r], \\ Cl_{\Omega^n}[U,V,W] & := & \mathit{Hier}_{\lhd}^{\Omega^{\widehat{n}}}[A[U,x,\mathfrak{b},V,W]] \wedge \\ & & (\forall \mathfrak{b} \lhd \Omega^{\widehat{n}})\, \mathit{Tr}[\omega, (U)_{\mathfrak{b}}, (V)_{\mathfrak{b}}], \\ Cl_{\Omega_n}[U,V,W] & := & \mathit{Hier}_{\lhd}^{\Omega_n}[A[U,x,\mathfrak{b},V,W]] \wedge \\ & & (\forall \mathfrak{b} \lhd \Omega_n)\, \mathit{Tr}[\omega, (U)_{\mathfrak{b}}, (V)_{\mathfrak{b}}]. \end{array}$$

By the following lemma we get the existence of such class hierarchies in the appropriate set theory, i.e. in NBG extended by iterated comprehension.

Lemma 101.

(1) $\mathsf{NBG} \vdash \exists X \exists Y\, Cl_n[\overline{m}, X, Y, W]$,

(2) $\mathsf{NBG} \cup \exists\, \mathit{Hier}_{\lhd}^{\Omega^{\widehat{n}}}[\Pi^1_0] \vdash \exists X \exists Y\, Cl_{\Omega^n}[X,Y,W]$,

(3) $\mathsf{NBG} \cup \exists\, \mathit{Hier}_{\lhd}^{\Omega_n}[\Pi^1_0] \vdash \exists X \exists Y\, Cl_{\Omega_n}[X,Y,W]$.

Proof.

(1) We use that $\mathsf{NBG} \vdash \exists\, \mathit{Hier}_{\in}^{\overline{k}}[\Pi^1_0]$ for any k by Theorem 79, and we proceed analogous to Part 3, taking into account that the truth predicates $(V)_r$, in $(\forall r \in \overline{n})\, \mathit{Tr}[\overline{m}, (U)_r, (V)_r]$, are for formulas of rank at most m, hence the construction of any $(V)_r$ takes only $m+1$ stages.

(2) Analogous to Part 3.

(3) To get U, V with $Cl_{\Omega_n}[U,V,W]$, the simultaneous inductive definition of U and V is replaced by the construction of a single class Z such that for

$$\begin{array}{rcl} U & = & \{\langle x,\mathfrak{c}\rangle \mid \mathfrak{c} \lhd \Omega_n \wedge x \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c})}\}, \\ V & = & \{\langle x,\mathfrak{c}\rangle \mid \mathfrak{c} \lhd \Omega_n \wedge x \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c} \hat{+} \widehat{1})}\}, \end{array}$$

we get that $Cl_{\Omega_n}[U,V,W]$. We build Z such that $Hier_{\lhd}^{\Omega_n}[A[Z,x,\mathfrak{b},W]]$ where $A[Z,x,\mathfrak{b},W]$ is the disjunction of the following formulas (i.e., A is a modification of the formula A in Definition 100, and the formula A in the proof of Lemma 85)

$$\begin{array}{l} \exists \mathfrak{c}(\mathfrak{b} = \widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \wedge (\\ \mathfrak{c} = \emptyset \wedge z \in W \vee \\ (\exists \mathfrak{a} \lhd \mathfrak{c}) \exists z (z \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{a})} \wedge x = \langle z, \mathfrak{a}\rangle) \vee \\ (\exists \mathfrak{a} \lhd \mathfrak{c}) \exists y \exists z (\langle z, y\rangle \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{a})} \wedge x = \langle z, \langle \mathfrak{a}, y\rangle\rangle) \vee \\ \exists \mathfrak{a} \exists z \exists f \exists g \exists \phi \exists p (\mathfrak{a} \hat{+} \widehat{1} = \mathfrak{c} \wedge \\ \qquad\qquad x = \langle z, \langle\!\langle \phi, p, f, g\rangle\!\rangle\rangle \wedge \phi[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_p][f,g]_\infty \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{a} \hat{+} \widehat{1})}))), \end{array}$$

$$\begin{array}{l} \exists \mathfrak{c} \exists r (\mathfrak{b} = \widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{r} \wedge \\ \exists \phi \exists \psi \exists y \exists z \exists p \exists q \exists s \exists t (x \in \mathcal{G}_{\mathcal{L}^1} \wedge rk(x) = r \wedge rk(\phi) = s \wedge rk(\psi) = t \wedge (\\ x = (\dot{\mathtt{c}}_y \mathrel{\dot{\in}} \dot{\mathtt{c}}_z) \wedge y \in z \vee \\ x = \dot{\sim}(\dot{\mathtt{c}}_y \mathrel{\dot{\in}} \dot{\mathtt{c}}_z) \wedge y \notin z \vee \\ x = (\dot{\mathtt{c}}_y \mathrel{\dot{\in}} \dot{\mathtt{C}}_z) \wedge y \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}), z} \vee \\ x = \dot{\sim}(\dot{\mathtt{c}}_y \mathrel{\dot{\in}} \dot{\mathtt{C}}_z) \wedge y \notin (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}), z} \vee \\ x = (\phi \mathbin{\dot{\vee}} \psi) \wedge (\phi \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{s}} \vee \psi \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{t}}) \vee \\ x = (\phi \mathbin{\dot{\wedge}} \psi) \wedge \phi \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{s}} \wedge \psi \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{t}} \vee \\ x = \dot{\exists}\, \dot{\mathtt{u}}_p (\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q]) \wedge \exists u (\phi[\dot{\mathtt{c}}_u/\dot{\mathtt{v}}_q] \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{s}}) \wedge \dot{\mathtt{u}}_p \notin term(\phi) \vee \\ x = \dot{\forall}\, \dot{\mathtt{u}}_p (\phi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q]) \wedge \forall u (\phi[\dot{\mathtt{c}}_u/\dot{\mathtt{v}}_q] \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{s}}) \wedge \dot{\mathtt{u}}_p \notin term(\phi) \vee \\ x = \dot{\exists}\, \dot{\mathtt{U}}_p (\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q]) \wedge \exists u (\phi[\dot{\mathtt{C}}_u/\dot{\mathtt{V}}_q] \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{s}}) \wedge \dot{\mathtt{U}}_p \notin term(\phi) \vee \\ x = \dot{\forall}\, \dot{\mathtt{U}}_p (\phi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q]) \wedge \forall u (\phi[\dot{\mathtt{C}}_u/\dot{\mathtt{V}}_q] \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{s}}) \wedge \dot{\mathtt{U}}_p \notin term(\phi)))), \end{array}$$

$$\exists \mathfrak{c} \exists r (\mathfrak{b} = \widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c} \hat{+} \widehat{1}) \wedge x \in (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \hat{+} \widehat{r}}).$$

The stages $(Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c})}$ all belong to the cumulative hierarchy of classes; the levels between $\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c})$ and $\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c} \hat{+} \widehat{1})$ are used to build the truth predicate with respect to the class universe $(Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c})}$; and the stages $(Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c} \hat{+} \widehat{1})}$ are truth predicates such that $Tr[\omega, (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c})}, (Z)_{\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c} \hat{+} \widehat{1})}]$ (see also Jäger and Krähenbühl [10] for a similar construction). The levels between $\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c} \hat{+} \widehat{1})$ and $\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c} \hat{+} \widehat{2})$ are not used. Finally, we get Z by $\exists\, Hier_{\lhd}^{\Omega_n}[\Pi_0^1]$, because of $\widehat{\omega} \hat{\cdot} (\widehat{2} \hat{\cdot} \mathfrak{c}) \lhd \Omega_n$ for any $\mathfrak{c} \lhd \Omega_n$. $\square$

The following theorem and the subsequent corollary are formally stating the properties of the hierarchy, i.e. they are a direct consequence of the definition of the class hierarchy.

Theorem 102. (Class Hierarchy)

(1) $\mathsf{NBG} \vdash \forall\alpha\forall\phi\forall r(Cl_n[\alpha, U, V, W] \wedge rk(\phi) \in \alpha' \to (\forall p \in \overline{n})(\forall q \in p)$
$W = (U)_\emptyset \wedge ((U)_p)_q = (U)_q \wedge \forall z((U)_p)_{\langle\!\langle q,z\rangle\!\rangle} = ((U)_q)_z \wedge$
$\forall z(z \in ((U)_{q'})_{\langle\!\langle\phi,r,f,g\rangle\!\rangle} \leftrightarrow \phi[\dot{\mathsf{c}}_z/\dot{\mathrm{v}}_r][f,g]_\infty^{(U)_q})),$

(2) $\mathsf{NBG} \vdash \forall\phi\forall r(Cl_{\Omega^n}[U, V, W] \to (\forall\mathfrak{a} \lhd \Omega^{\widehat{n}})(\forall\mathfrak{b} \lhd \mathfrak{a})$
$W = (U)_\emptyset \wedge ((U)_\mathfrak{a})_\mathfrak{b} = (U)_\mathfrak{b} \wedge \forall z((U)_\mathfrak{a})_{\langle\!\langle \mathfrak{b},z\rangle\!\rangle} = ((U)_\mathfrak{b})_z \wedge$
$\forall z(z \in ((U)_{\mathfrak{b}\hat{+}\hat{1}})_{\langle\!\langle\phi,r,f,g\rangle\!\rangle} \leftrightarrow \phi[\dot{\mathsf{c}}_z/\dot{\mathrm{v}}_r][f,g]_\infty^{(U)_\mathfrak{b}})),$

(3) $\mathsf{NBG} \vdash \forall\phi\forall r(Cl_{\Omega_n}[U, V, W] \to (\forall\mathfrak{a} \lhd \Omega_n)(\forall\mathfrak{b} \lhd \mathfrak{a})$
$W = (U)_\emptyset \wedge ((U)_\mathfrak{a})_\mathfrak{b} = (U)_\mathfrak{b} \wedge \forall z((U)_\mathfrak{a})_{\langle\!\langle \mathfrak{b},z\rangle\!\rangle} = ((U)_\mathfrak{b})_z \wedge$
$\forall z(z \in ((U)_{\mathfrak{b}\hat{+}\hat{1}})_{\langle\!\langle\phi,r,f,g\rangle\!\rangle} \leftrightarrow \phi[\dot{\mathsf{c}}_z/\dot{\mathrm{v}}_r][f,g]_\infty^{(U)_\mathfrak{b}})).$

Proof.

(1) The formula $\phi[\dot{\mathsf{c}}_z/\dot{\mathrm{v}}_r][f,g]_\infty^{(U)_q}$ is equivalent to the elementary formula $\phi[\dot{\mathsf{c}}_z/\dot{\mathrm{v}}_r][f,g]_\infty \in (V)_q$, because of $Cl_n[\alpha, U, V, W]$. Hence the claim follows directly by Definition 100 and by elementary induction on p.

(2) Analogous to Part 1, using that $\Omega^{\widehat{n}} \lhd \Omega_1$ and $TI_\lhd^{\Omega_k}[\Pi_0^1]$ by Theorem 74.

(3) Analogous to Part 2. □

Corollary 103.

(1) $\mathsf{NBG} \vdash \forall\alpha\forall\phi\forall r(Cl_{\overline{n}}[\alpha, U, V, W] \wedge rk(\phi) \in \alpha' \to (\forall p \in \overline{n})(\forall q \in p)$
$W \mathrel{\mathring{\in}} (U)_p \wedge (U)_q \mathrel{\mathring{\in}} (U)_p \wedge (U)_q \mathrel{\mathring{\subsetneq}} (U)_p \wedge$
$(\exists X \mathrel{\mathring{\in}} (U)_{q'})\forall z(z \in X \leftrightarrow \phi[\dot{\mathsf{c}}_z/\dot{\mathrm{v}}_r][f,g]_\infty^{(U)_q})),$

(2) $\mathsf{NBG} \vdash \forall\phi\forall r(Cl_{\Omega^n}[U, V, W] \to (\forall\mathfrak{a} \lhd \Omega^{\widehat{n}})(\forall\mathfrak{b} \lhd \mathfrak{a})$
$W \mathrel{\mathring{\in}} (U)_\mathfrak{a} \wedge (U)_\mathfrak{b} \mathrel{\mathring{\in}} (U)_\mathfrak{a} \wedge (U)_\mathfrak{b} \mathrel{\mathring{\subsetneq}} (U)_\mathfrak{a} \wedge$
$(\exists X \mathrel{\mathring{\in}} (U)_{\mathfrak{b}\hat{+}\hat{1}})\forall z(z \in X \leftrightarrow \phi[\dot{\mathsf{c}}_z/\dot{\mathrm{v}}_r][f,g]_\infty^{(U)_\mathfrak{b}})),$

(3) $\mathsf{NBG} \vdash \forall\phi\forall r(Cl_{\Omega_n}[U, V, W] \to (\forall\mathfrak{a} \lhd \Omega_n)(\forall\mathfrak{b} \lhd \mathfrak{a})$
$W \mathrel{\mathring{\in}} (U)_\mathfrak{a} \wedge (U)_\mathfrak{b} \mathrel{\mathring{\in}} (U)_\mathfrak{a} \wedge (U)_\mathfrak{b} \mathrel{\mathring{\subsetneq}} (U)_\mathfrak{a} \wedge$
$(\exists X \mathrel{\mathring{\in}} (U)_{\mathfrak{b}\hat{+}\hat{1}})\forall z(z \in X \leftrightarrow \phi[\dot{\mathsf{c}}_z/\dot{\mathrm{v}}_r][f,g]_\infty^{(U)_\mathfrak{b}})).$

10. Reduction by Asymmetric Interpretation

In contrast to Cantini [1], where the asymmetric interpretation is into hierarchies of sets of numbers, which are built by using fixed standard Π^1_n-complete predicates, in this section, the asymmetric interpretation is into cumulative hierarchies of classes, which are built by using truth predicates for each level of the hierarchy, i.e., the truth predicates and the stages of the hierarchy are built simultaneously (see Section 9). An asymmetric interpretation into hierarchies of classes of this sort is also used in Jäger and Krähenbühl [10].

The majorizing functions (see Cantini [1]) for the asymmetric interpretations in this section are of common exponential form, see Lemma 109. Clearly the role of ω, as e.g. in Cantini [1], is taken over by Ω (i.e., the class of all ordinals) in context of this thesis.

As already described in the Introduction, the aim of the asymmetric interpretation is the reduction of the systems with choice $\mathcal{T}_{\mathsf{ch}}$ (i.e., $CA[\Pi^1_0] \cup AC[\Pi^1_0]$, $CA[\Pi^1_0] \cup DC_{\mathcal{O}n}[\Pi^1_0]$, and $CA[\Pi^1_0] \cup DC_{\mathcal{O}n}[\Pi^1_0] \cup TI_{\in}[\mathcal{L}^1]$), to the corresponding systems with iterated comprehension $\mathcal{T}_{\mathsf{it}}$ (i.e., NBG, $\mathsf{NBG} \cup (CA[\Pi^1_0])_{<\Omega^\omega}$, and $\mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}$). Roughly depicted and just in a nutshell, the asymmetric interpretation is used in the following way

(1) After the asymmetric interpretation of the formula A into the hierarchy U, denoted by $A\langle \mathfrak{a}, \mathfrak{b}\rangle^U$, every quantified class variable of A ranges over some specific level of the class hierarchy, i.e. generally all existential quantifiers range over some level $(U)_{\mathfrak{b}}$ higher than the level of the universal quantifiers $(U)_{\mathfrak{a}}$ (see Definition 104).

(2) All formulas *provable* (cut-reduced) in the system $\mathcal{T}_{\mathsf{ch}}$ hold true after asymmetric interpretation into the class hierarchy U, i.e.,

$$\mathcal{T}_{\mathsf{ch}} \vdash_0 A \quad \Rightarrow \quad \mathcal{T}_{\mathsf{it}} \vdash Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \to \exists \mathfrak{b}(A\langle \emptyset, \mathfrak{b}\rangle^U).$$

We write $Cl_{\mathcal{T}_{\mathsf{ch}}}[U]$ to emphasize the dependence of the class hierarchy U on the specific theory $\mathcal{T}_{\mathsf{ch}}$ (actually, the hierarchy U also depends on the formula A, see Lemma 109).

(3) For any formula up to some logical complexity, i.e. essentially Σ^1_1, the asymmetric interpretation of the formula reflects the truth of the original

formula, that is,

$$A \text{ essentially } \Sigma^1_1 \quad \Rightarrow \quad \mathcal{T}_{\mathsf{it}} \vdash Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \to (A\langle \mathfrak{a}, \mathfrak{b}\rangle^U \to A)$$

(see Lemma 107).

(4) The class hierarchy exists in the system $\mathcal{T}_{\mathsf{it}}$, that is, $\mathcal{T}_{\mathsf{it}} \vdash \exists U\, Cl_{\mathcal{T}_{\mathsf{ch}}}[U]$ (see Lemma 101).

By putting (2) to (4) together, we have that if a formula A is essentially Σ^1_1 and $\mathcal{T}_{\mathsf{ch}} \vdash A$ then we also have $\mathcal{T}_{\mathsf{it}} \vdash A$. For the system $\mathcal{T}_{\mathsf{ch}}$ with full induction, we are going to use a formalized version of the asymmetric interpretation with truth and proof predicates involved, such that steps (2) and (3) become

(2')

$$\mathcal{T}_{\mathsf{it}} \vdash Pr_{\mathcal{T}_{\mathsf{ch}}}[\ulcorner A \urcorner] \to (Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \to \exists \mathfrak{b}(\ulcorner A \urcorner\langle \dot{\mathsf{c}}_{\emptyset}, \dot{\mathsf{c}}_{\mathfrak{b}}\rangle[f,g]^U_\infty)),$$

where $Pr_{\mathcal{T}_{\mathsf{ch}}}[\ulcorner A \urcorner]$ stands for $\exists Z(Pr^{DC+}_{\Omega_{n+3}}[Z] \wedge \{\ulcorner A \urcorner\} \in (Z)_{\Omega \hat{+} \widehat{\omega}, \overline{n}})$ (n actually depends on the formula A, see Lemma 109).

(3') If A is essentially Σ^1_1 then

$$\mathcal{T}_{\mathsf{it}} \vdash Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \to (\ulcorner A \urcorner\langle \dot{\mathsf{c}}_{\mathfrak{a}}, \dot{\mathsf{c}}_{\mathfrak{b}}\rangle[f,g]^U_\infty \to \ulcorner A \urcorner[f,g]^U_\infty)$$

(see Lemma 107).

The proof predicate and the truth definition are such that

$$\mathcal{T}_{\mathsf{ch}} \vdash A \quad \Rightarrow \quad \mathcal{T}_{\mathsf{it}} \vdash Pr_{\mathcal{T}_{\mathsf{ch}}}[\ulcorner A \urcorner],$$

$$A \text{ essentially } \Sigma^1_1 \quad \Rightarrow \quad \mathcal{T}_{\mathsf{it}} \vdash \sharp_A[f,g,U] \to (\ulcorner A \urcorner[f,g]^U_\infty \to A)$$

(see Theorem 98 and Lemma 90), where $\sharp_A$ stands for the proper assignment of sets and classes to the free variables in A. Together with (4), and because of $\mathcal{T}_{\mathsf{it}} \vdash Cl_{\mathcal{T}_{\mathsf{ch}}}[U] \to \exists f \exists g(\sharp_A[f,g,U])$, we have that if a formula A is essentially Σ^1_1 and $\mathcal{T}_{\mathsf{ch}} \vdash A$ then we also have $\mathcal{T}_{\mathsf{it}} \vdash A$.

Finally, by using these reductions at the end of this section, we are able to prove the desired proof-theoretic equivalences between choice schemes and iterated comprehension.

First, we need to define the asymmetric translation of formulas and Gödel-codes, which is the basis for the asymmetric interpretation.

Definition 104.

(1) For variables $\mathtt{v}_i$, $\mathtt{u}_i$, $\mathtt{V}_j$, $\mathtt{U}_j$ we define ${\mathtt{v}_i}^* := \mathtt{v}_{2\cdot i}$, ${\mathtt{u}_i}^* := \mathtt{u}_{2\cdot i}$, and ${\mathtt{V}_j}^* := \mathtt{V}_{2\cdot j}$, ${\mathtt{U}_j}^* := \mathtt{U}_{2\cdot j}$, and ${\mathtt{V}_j}^+ := \mathtt{v}_{j+1}$, ${\mathtt{U}_j}^+ := \mathtt{u}_{j+1}$.

(2) $A^* \in \mathcal{L}^1$ is the formula $A \in \mathcal{L}^1$ with all symbols $\mathtt{v}_i$, $\mathtt{u}_i$, $\mathtt{V}_j$, $\mathtt{U}_j$ in A replaced by $\mathtt{v}_{2\cdot i}$, $\mathtt{u}_{2\cdot i}$, $\mathtt{V}_{2\cdot j}$, $\mathtt{U}_{2\cdot j}$ respectively. For $\mathcal{T} \subseteq \mathcal{L}^1$ we define $\mathcal{T}^* := \{A^* \mid A \in \mathcal{T}\}$.

(3) For $A \in \mathcal{L}^1$ we define

$$A\langle U, V\rangle := \begin{cases} A & A \text{ atomic}, \\ (B\langle U,V\rangle \circ C\langle U,V\rangle) & A = (B \circ C), \\ \sigma x B\langle U,V\rangle[x] & A = \sigma x B[x], \\ \forall X^+(B\langle U,V\rangle[(U)_{X^+}]) & A = \forall X B[X], \\ \exists X^+(B\langle U,V\rangle[(V)_{X^+}]) & A = \exists X B[X], \end{cases}$$

such that $A\langle U,V\rangle \in \mathcal{L}^1$ for $A \in \mathcal{L}^{1^*}$.

We write $A\langle x,y\rangle^U$ for $A\langle (U)_x, (U)_y\rangle$.

(4) Analogously to Part 1 and 2 we define the expressions $\dot{\mathtt{v}}_p^*$, $\dot{\mathtt{u}}_p^*$, $\dot{\mathtt{V}}_p^*$, $\dot{\mathtt{U}}_p^*$, $\dot{\mathtt{V}}_p^+$, $\dot{\mathtt{U}}_p^+$, and ϕ^*. And similar to Part 3 we define the expression $\phi\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_y\rangle$ such that the following holds (in NBG)

$$\phi\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_y\rangle = \begin{cases} \phi & rk(\phi) = \overline{1}, \\ \psi\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_y\rangle \,\dot{\circ}\, \xi\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_y\rangle & \phi = \psi \,\dot{\circ}\, \xi, \\ \dot{\sigma}\, \dot{\mathtt{u}}_p(\psi\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_y\rangle[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q]) & \phi = \dot{\sigma}\, \dot{\mathtt{u}}_p(\psi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q]), \\ \dot{\forall}\, \dot{\mathtt{U}}_p^+(\psi\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_y\rangle[(\dot{\mathtt{C}}_x)_{\dot{\mathtt{U}}_p^+}/\dot{\mathtt{V}}_q]) & \phi = \dot{\forall}\, \dot{\mathtt{U}}_p(\psi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q]), \\ \dot{\exists}\, \dot{\mathtt{U}}_p^+(\psi\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_y\rangle[(\dot{\mathtt{C}}_y)_{\dot{\mathtt{U}}_p^+}/\dot{\mathtt{V}}_q]) & \phi = \dot{\exists}\, \dot{\mathtt{U}}_p(\psi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q]), \end{cases}$$

where $\xi[(\dot{\mathtt{C}}_x)_{\dot{\mathtt{v}}_p}/\dot{\mathtt{V}}_q]$ is ξ with all "subformulas" $(\dot{\mathtt{v}}_r \mathrel{\dot{\in}} \dot{\mathtt{V}}_q)$ in ξ replaced by $(\ulcorner z \in (U)_y \urcorner[\dot{\mathtt{v}}_r/\dot{z}][\dot{\mathtt{v}}_p/\dot{y}][\dot{\mathtt{C}}_x/\dot{U}])$. Analogously $\xi[(\dot{\mathtt{C}}_x)^{\dot{\mathtt{v}}_p}/\dot{\mathtt{V}}_q]$.

The following lemma and its corollary are about some general technical properties of class constants within the class hierarchy.

Lemma 105.

(1) $\mathsf{NBG} \vdash Cl_{\Omega_{k+1}}[U,V,W] \wedge ((U)_{\Omega_k})_x = ((U)_{\Omega_k})_y \to$
$$\forall\phi\forall q(\phi[\dot{\mathtt{C}}_x/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow \phi[\dot{\mathtt{C}}_y/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}}),$$

(2) $\mathsf{NBG} \vdash Cl_{\Omega_{k+1}}[U,V,W] \wedge (((U)_{\Omega_k})_x)_z = ((U)_{\Omega_k})_y \to$
$$\forall\phi\forall q(\phi[(\dot{\mathtt{C}}_x)_{\dot{\mathtt{c}}_z}/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow \phi[\dot{\mathtt{C}}_y/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}}),$$

(3) $\mathsf{NBG} \vdash Cl_{\Omega_{k+1}}[U,V,W] \wedge (((U)_{\Omega_k})_x)^z = ((U)_{\Omega_k})_y \to$
$$\forall\phi\forall q(\phi[(\dot{\mathtt{C}}_x)^{\dot{\mathtt{c}}_z}/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow \phi[\dot{\mathtt{C}}_y/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}}).$$

Proof.

(1) Analogous to Part 2.

(2) We have that $\psi[f,g]_\infty^{(U)_{\Omega_k}}$ is equivalent to $\psi[f,g]_\infty \in (V)_{\Omega_k}$, hence the claim can be proved by elementary induction on the rank $rk(\phi)$. The only difficult case is $\phi = (\dot{\mathtt{c}}_v \mathbin{\dot\in} \dot{\mathtt{V}}_q)$. All other cases are either trivial, or follow from this case, or follow directly by the i.h. and Lemma 88. By definition we have

$$(\dot{\mathtt{c}}_v \mathbin{\dot\in} \dot{\mathtt{V}}_q)[(\dot{\mathtt{C}}_x)_{\dot{\mathtt{c}}_z}/\dot{\mathtt{V}}_q] = \ulcorner v \in (X)_z \urcorner[\dot{\mathtt{c}}_v/\dot{v}][\dot{\mathtt{c}}_z/\dot{z}][\dot{\mathtt{C}}_x/\dot{X}],$$

and by Lemma 90(3) we have that

$$f(\overline{\sharp v}) \in (((U)_{\Omega_k})_{g(\overline{\sharp X})})_{f(\overline{\sharp z})} \leftrightarrow \ulcorner v \in (X)_z \urcorner[f,g]_\infty^{(U)_{\Omega_k}},$$

hence we get $v \in (((U)_{\Omega_k})_x)_z \leftrightarrow (\dot{\mathtt{c}}_v \mathbin{\dot\in} \dot{\mathtt{V}}_q)[(\dot{\mathtt{C}}_x)_{\dot{\mathtt{c}}_z}/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}}$. On the other hand we have $v \in ((U)_{\Omega_k})_y \leftrightarrow (\dot{\mathtt{c}}_v \mathbin{\dot\in} \dot{\mathtt{V}}_q)[\dot{\mathtt{C}}_y/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}}$, hence because of $(((U)_{\Omega_k})_x)_z = ((U)_{\Omega_k})_y$ we get

$$(\dot{\mathtt{c}}_v \mathbin{\dot\in} \dot{\mathtt{V}}_q)[(\dot{\mathtt{C}}_x)_{\dot{\mathtt{c}}_z}/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow (\dot{\mathtt{c}}_v \mathbin{\dot\in} \dot{\mathtt{V}}_q)[\dot{\mathtt{C}}_y/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}}.$$

(3) Analogous to Part 2. □

Corollary 106.

(1) $\mathsf{NBG} \vdash Cl_{\Omega_{k+1}}[U,V,W] \to \forall\phi(\forall\mathfrak{a} \lhd \Omega_k)(\forall\mathfrak{b} \lhd \mathfrak{a})\forall q($
$$\phi[\dot{\mathtt{C}}_{\langle\mathfrak{a},\langle\mathfrak{b},x\rangle\rangle}/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow \phi[\dot{\mathtt{C}}_{\langle\mathfrak{b},x\rangle}/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}}),$$

(2) $\mathsf{NBG} \vdash Cl_{\Omega_{k+1}}[U,V,W] \to \forall\phi(\forall\mathfrak{a} \lhd \Omega_k)\forall q($
$$\phi[(\dot{\mathtt{C}}_{\mathfrak{a}})_{\dot{\mathtt{c}}_x}/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow \phi[\dot{\mathtt{C}}_{\langle\mathfrak{a},x\rangle}/\dot{\mathtt{V}}_q][f,g]_\infty^{(U)_{\Omega_k}}).$$

Proof.

(1) $((U)_{\Omega_k})_{\langle\mathfrak{a},\langle\mathfrak{b},x\rangle\rangle} = ((U)_{\mathfrak{a}})_{\langle\mathfrak{b},x\rangle} = ((U)_{\mathfrak{b}})_x = ((U)_{\Omega_k})_{\langle\mathfrak{b},x\rangle}$ by Theorem 102.

(2) $(((U)_{\Omega_k})_{\mathfrak{a}})_x = ((U)_{\mathfrak{a}})_x = ((U)_{\Omega_k})_{\langle\mathfrak{a},x\rangle}$ by Theorem 102. □

The following lemma covers Step 3 of the asymmetric interpretation, which was described at the beginning of this section. It shows that validity of formulas in $\Sigma^1(\Pi^1_0)$ is reflected by the asymmetric translation.

Lemma 107.

(1) If $A \in \Sigma^1(\Pi^1_0)$ then $\mathsf{NBG} \vdash A^*\langle U,V\rangle \to A^*$,

(2) $\mathsf{NBG} \vdash Cl_{\Omega_{k+1}}[U,V,W] \to$
$$(\forall\phi \in \mathcal{G}_{\Sigma^1(\Pi^1_0)})(\forall\mathfrak{a} \lhd \Omega_k)(\phi^*\langle\dot{\mathtt{C}}_x,\dot{\mathtt{C}}_{\mathfrak{a}}\rangle[f,g]_\infty^{(U)_{\Omega_k}} \to \phi^*[f,g]_\infty^{(U)_{\Omega_k}}).$$

Proof.

(1) By induction on the structure of the formula A. If A is atomic then the claim follows trivially. If $A = B \circ C$ then the claim follows by i.h.. If $A = \forall x B[x/u]$ then $A^* = \forall x^* B^*[x^*/u^*]$, and by i.h. we have $B^*\langle U, V\rangle \to B^*$, hence we get $\forall x^* B^*\langle U, V\rangle[x^*/u^*] \to \forall x^* B^*[x^*/u^*]$, that is $A^*\langle U, V\rangle \to A^*$ because of $\forall x^*(B^*\langle U, V\rangle[x^*/u^*]) = A^*\langle U, V\rangle$. Similar for $A = \exists x B[x/u]$. If $A = \exists X B[X/Y]$ then we have

$$Y^* = (V)_z \to (B^*\langle U, V\rangle[(V)_z/Y^*] \to B^*\langle U, V\rangle)$$

and because of $B^* \to A^*$ and $B^*\langle U, V\rangle \to B^*$ by i..h., we get

$$Y^* = (V)_z \to (B^*\langle U, V\rangle[(V)_z/Y^*] \to A^*).$$

We have $\mathsf{NBG} \vdash \exists Y^*(Y^* = (V)_z)$ hence $B^*\langle U, V\rangle[(V)_z/Y^*] \to A^*$, that is $A^*\langle U, V\rangle \to A^*$ because of $\exists x(B^*\langle U, V\rangle[(V)_x/Y^*]) = A^*\langle U, V\rangle$ for $x = X^{*+}$.

(2) We have that $\psi[f, g]_\infty^{(U)_{\Omega_k}}$ is equivalent to $\psi[f, g]_\infty \in (V)_{\Omega_k}$, hence the claim can be proved by elementary induction on the rank $rk(\phi)$. If $rk(\phi) = \overline{1}$ then $\phi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle = \phi^*$. If $\phi = \psi \mathbin{\dot\vee} \xi$ and $\phi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[f, g]_\infty^{(U)_{\Omega_k}}$ then we get $\psi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[f, g]_\infty^{(U)_{\Omega_k}} \vee \xi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[f, g]_\infty^{(U)_{\Omega_k}}$ by Lemma 88 and $\psi^*[f, g]_\infty^{(U)_{\Omega_k}} \vee \xi^*[f, g]_\infty^{(U)_{\Omega_k}}$ by i.h., hence $\phi^*[f, g]_\infty^{(U)_{\Omega_k}}$ by Lemma 88 again. Analogous for $\phi = \psi \mathbin{\dot\wedge} \xi$. If $\phi = \dot\forall\, \dot{\mathtt{u}}_p(\psi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])$ and if we assume $\phi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[f, g]_\infty^{(U)_{\Omega_k}}$ then $\phi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle = \dot\forall\, \dot{\mathtt{u}}_p^*(\psi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[\dot{\mathtt{u}}_p^*/\dot{\mathtt{v}}_q^*])$, hence $\forall z(\psi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q^*][f, g]_\infty^{(U)_{\Omega_k}})$ by Lemma 88, and because of $\psi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q^*] = \psi^*[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q^*]\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle$ we have $\forall z(\psi^*[\dot{\mathtt{c}}_z/\dot{\mathtt{v}}_q^*][f, g]_\infty^{(U)_{\Omega_k}})$ by i.h., hence $\phi^*[f, g]_\infty^{(U)_{\Omega_k}}$ by Lemma 88. Similar for $\phi = \dot\exists\, \dot{\mathtt{u}}_p(\psi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])$. If $\phi = \dot\exists\, \dot{\mathtt{U}}_p(\psi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])$ and $\phi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[f, g]_\infty^{(U)_{\Omega_k}}$ then by Lemma 88 we get

$$\exists z(\psi^*\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[(\dot{\mathtt{C}}_\mathfrak{a})_{\dot{\mathtt{c}}_z}/\dot{\mathtt{V}}_q^*][f, g]_\infty^{(U)_{\Omega_k}}),$$

hence $\exists z(\psi^*[\dot{\mathtt{C}}_{\langle \mathfrak{a}, z\rangle}/\dot{\mathtt{V}}_q^*]\langle \dot{\mathtt{C}}_x, \dot{\mathtt{C}}_\mathfrak{a}\rangle[f, g]_\infty^{(U)_{\Omega_k}})$ by Corollary 106, and because of $rk(\psi[\dot{\mathtt{C}}_{\langle \mathfrak{a}, z\rangle}/\dot{\mathtt{V}}_q]) \in rk(\phi)$ and by i.h. we have $\exists z(\psi^*[\dot{\mathtt{C}}_{\langle \mathfrak{a}, z\rangle}/\dot{\mathtt{V}}_q^*][f, g]_\infty^{(U)_{\Omega_k}})$, that is $\exists z(\psi^*[\dot{\mathtt{C}}_z/\dot{\mathtt{V}}_q^*][f, g]_\infty^{(U)_{\Omega_k}})$, and finally $\phi^*[f, g]_\infty^{(U)_{\Omega_k}}$ by Lemma 88. □

Upward and downward persistency of the asymmetric translation, with respect to the cumulative class hierarchies, is essential in the proof of the asymmetric interpretation following on the next page.

Lemma 108. (Persistency)

(1) $\mathsf{NBG} \vdash Cl_k[\alpha, U, V, W] \to (\forall s \in \overline{k})(\forall q \in s')(\forall p \in q')(\forall r \in p')($
$$A^*\langle p, q\rangle^U \to A^*\langle r, s\rangle^U),$$

(2) $\mathsf{NBG} \vdash Cl_{\Omega^k}[U, V, W] \to (\forall \mathfrak{d} \lhd \Omega^{\widehat{k}})(\forall \mathfrak{b} \unlhd \mathfrak{d})(\forall \mathfrak{a} \unlhd \mathfrak{b})(\forall \mathfrak{c} \unlhd \mathfrak{a})($
$$A^*\langle \mathfrak{a}, \mathfrak{b}\rangle^U \to A^*\langle \mathfrak{c}, \mathfrak{d}\rangle^U),$$

(3) $\mathsf{NBG} \vdash Cl_{\Omega_{k+1}}[U, V, W] \to (\forall \mathfrak{d} \lhd \Omega_k)(\forall \mathfrak{b} \unlhd \mathfrak{d})(\forall \mathfrak{a} \unlhd \mathfrak{b})(\forall \mathfrak{c} \unlhd \mathfrak{a})\forall \phi($
$$\phi^*\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{b}}\rangle[f, g]_\infty^{(U)_{\Omega_k}} \to \phi^*\langle \dot{\mathsf{C}}_{\mathfrak{c}}, \dot{\mathsf{C}}_{\mathfrak{d}}\rangle[f, g]_\infty^{(U)_{\Omega_k}}).$$

Proof.

(1) By induction on the structure of the formula A. If A is atomic then the claim is trivial. If $A = \forall X B[X/Z]$ and $y = X^{*+}$ then

$$A^*\langle U, p, q\rangle = \forall y(B^*\langle p, q\rangle^U[((U)_p)_y/Z^*]),$$

and we get $\forall y(B^*\langle p, q\rangle^U[((U)_p)_y/Z^*]) \to \forall y(B^*\langle r, s\rangle^U[((U)_p)_y/Z^*])$ by i.h.. Clearly $\forall y(B^*\langle r, s\rangle^U[((U)_p)_y/Z^*]) \to \forall y(B^*\langle r, s\rangle^U[((U)_p)_{\langle r,y\rangle}/Z^*])$, hence $A^*\langle p, q\rangle^U \to A^*\langle r, s\rangle^U$ because of $((U)_p)_{\langle r,y\rangle} = ((U)_r)_y$ by Theorem 102, or because of $p = r$. Analogous for $A = \exists X B[X]$. The other cases follow directly by i.h. and logic.

(2) Analogous to Part 1.

(3) We have that $\psi[f, g]_\infty^{(U)_{\Omega_k}}$ is equivalent to $\psi[f, g]_\infty \in (V)_{\Omega_k}$, hence the claim can be proved by elementary induction on the rank $rk(\phi)$. If $rk(\phi) = \overline{1}$ then $\phi^*\langle \dot{\mathsf{C}}_x, \dot{\mathsf{C}}_y\rangle = \phi^*$. If $\phi = \dot{\exists}\, \dot{\mathsf{U}}_p(\psi[\dot{\mathsf{U}}_p/\dot{\mathsf{V}}_q])$ then by i.h. we have

$$\psi^*[\dot{\mathsf{C}}_{\langle \mathfrak{d}, z\rangle}/\dot{\mathsf{V}}_q^*]\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{b}}\rangle[f, g]_\infty^{(U)_{\Omega_k}} \to \psi^*[\dot{\mathsf{C}}_{\langle \mathfrak{d}, z\rangle}/\dot{\mathsf{V}}_q^*]\langle \dot{\mathsf{C}}_{\mathfrak{c}}, \dot{\mathsf{C}}_{\mathfrak{d}}\rangle[f, g]_\infty^{(U)_{\Omega_k}}.$$

And clearly we have

$$\exists z(\psi^*\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{b}}\rangle[\dot{\mathsf{C}}_{\langle \mathfrak{d}, \langle \mathfrak{b}, z\rangle\rangle}/\dot{\mathsf{V}}_q^*][f, g]_\infty^{(U)_{\Omega_k}}) \to \exists z(\psi^*\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{b}}\rangle[\dot{\mathsf{C}}_{\langle \mathfrak{d}, z\rangle}/\dot{\mathsf{V}}_q^*][f, g]_\infty^{(U)_{\Omega_k}}),$$

hence by Corollary 106 and by combining the two implications we get

$$\exists z(\psi^*\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{b}}\rangle[(\dot{\mathsf{C}}_{\mathfrak{b}})_{\dot{\mathsf{c}}_z}/\dot{\mathsf{V}}_q^*][f, g]_\infty^{(U)_{\Omega_k}}) \to \exists z(\psi^*\langle \dot{\mathsf{C}}_{\mathfrak{c}}, \dot{\mathsf{C}}_{\mathfrak{d}}\rangle[(\dot{\mathsf{C}}_{\mathfrak{d}})_{\dot{\mathsf{c}}_z}/\dot{\mathsf{V}}_q^*][f, g]_\infty^{(U)_{\Omega_k}}),$$

that is $\phi^*\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{b}}\rangle[f, g]_\infty^{(U)_{\Omega_k}} \to \phi^*\langle \dot{\mathsf{C}}_{\mathfrak{c}}, \dot{\mathsf{C}}_{\mathfrak{d}}\rangle[f, g]_\infty^{(U)_{\Omega_k}}$ by Lemma 88. Analogous for $\phi = \dot{\forall}\, \dot{\mathsf{U}}_p(\psi[\dot{\mathsf{U}}_p/\dot{\mathsf{V}}_q])$. The other cases follow directly by i.h. and Lemma 88. □

The next technical lemma is the main part of the reduction, i.e., the asymmetric interpretation, and its proof spans over many pages. The proofs of Part 1 to 3 are quite similar, e.g., the proof of Part 3 is more or less a formalized version (within NBG) of the proof of Part 2.

Lemma 109. (Asymmetric Interpretation)

For $\{\mathtt{V}_i \mid \mathtt{V}_i \in \Gamma^{\vee *}\} \subseteq \{X_0, ..., X_l\}$ we have that

(1) If $\mathcal{C} = AC[\Pi^1_0]^{\rightarrow} \cup CA[\Pi^1_0]$ and $CA[\Pi^1_0]\ [AC[\Pi^1_0]]|\!\!\frac{n,i}{\mathcal{C},0}\ \Gamma$ then there is some j such that

$$\mathsf{NBG} \vdash \begin{array}{c} Cl_k[\overline{i+j}, U, V, W] \wedge Gl[(W)_\emptyset] \rightarrow \\ \forall p(p + \overline{2}^{\overline{n}} \in \overline{k} \wedge \overrightarrow{X} \mathring{\in} (U)_p \rightarrow \Gamma^{\vee *}\langle p, p + \overline{2}^{\overline{n}}\rangle^U), \end{array}$$

(2) If $\mathcal{D} = DC_{\mathcal{O}n}[\Pi^1_0]^{\rightarrow} \cup CA[\Pi^1_0]$ and $CA[\Pi^1_0]\ [DC_{\mathcal{O}n}[\Pi^1_0]]|\!\!\frac{n}{\mathcal{D},0}\ \Gamma$ then

$$\mathsf{NBG} \vdash \begin{array}{c} Cl_{\Omega^k}[U, V, W] \wedge Gl[(W)_\emptyset] \rightarrow \\ \forall \mathfrak{a}(\mathfrak{a} \hat{+} \Omega^{\widehat{n}} \lhd \Omega^{\widehat{k}} \wedge \overrightarrow{X} \mathring{\in} (U)_{\mathfrak{a}} \rightarrow \Gamma^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U), \end{array}$$

(3)
$$\mathsf{NBG} \vdash \begin{array}{c} Cl_{\Omega_{k+1}}[U, V, W] \wedge Gl[(W)_\emptyset] \wedge Pr^{DC+}_{\Omega_k}[Z] \rightarrow \\ \forall \mathfrak{a} \forall \mathfrak{b}(\mathfrak{a} \hat{+} \Omega^{\mathfrak{b}} \lhd \Omega_k \wedge y \in (Z)_{\mathfrak{b},\emptyset} \wedge g|\mathfrak{a} \rightarrow \\ (\exists \phi \in y)\phi^*\langle \dot{\mathtt{c}}_{\mathfrak{a}}, \dot{\mathtt{c}}_{\mathfrak{a} \hat{+} \Omega^{\mathfrak{b}} \hat{+} \widehat{2}}\rangle[f, g]^{(U)_{\Omega_k}}_\infty). \end{array}$$

where $\overrightarrow{X} \mathring{\in} Y$ stands for $X_0 \mathring{\in} Y \wedge ... \wedge X_l \mathring{\in} Y$, and $g|\mathfrak{a}$ stands for the formula $\forall p \exists y(\exists \mathfrak{b} \unlhd \mathfrak{a})g(p) = \langle \mathfrak{b}, y\rangle$.

Proof.

(1) Actually the claim is true for some constant number j. We define $j := \mathsf{rk}(A_0) + \mathsf{rk}(A_1) + \mathsf{rk}(A_2) + 8$, where $A_0 := (x = \langle w, u\rangle)$, and $A_1 := (w \in (Z)_v)$, and $A_2 := (\forall \gamma \forall \delta(\langle v, \gamma\rangle \in Z \wedge \langle w, \delta\rangle \in Z \rightarrow \gamma \in \delta')$, and we show the claim for the constant j by induction on n, considering all cases in Definition 12. The proof is analogous to Part 2, and we show only the following two cases (see Part 2 for all other cases):

If $\Gamma = \Phi, A$ and $A \in CA[\Pi^1_0]$ then $A^* = \exists Z \forall y(y \in Z \leftrightarrow B[y/z])$ for some $B \in \Pi^1_0$, and the asymmetric interpretation of A^* into the hierarchy U is such that

$$A^*\langle p, p + \overline{2}^{\overline{n}}\rangle^U = \exists x \forall y(y \in ((U)_{p + \overline{2}^{\overline{n}}})_x \leftrightarrow B[y/z]).$$

By Theorem 102 and because of $\mathsf{rk}(B) < i$, i.e. $rk(\ulcorner B \urcorner) \in \overline{i}$, we can find a class contained in the stage $(U)_{p'}$ such that

$$z \in ((U)_{p'})_{\langle\!\langle \ulcorner B \urcorner, \overline{\sharp z}, f, g\rangle\!\rangle} \leftrightarrow \ulcorner B \urcorner[\dot{\mathtt{c}}_z/\dot{z}][f, g]^{(U)_p}_\infty,$$

and by Lemma 90, for the elementary formula B, we have that

$$\sharp_B[f, g, (U)_p] \rightarrow (B \leftrightarrow \ulcorner B \urcorner[f, g]^{(U)_p}_\infty),$$

hence because of $\ulcorner B \urcorner[f,g]_\infty = \ulcorner B \urcorner[\dot{\mathsf{c}}_{f(\overline{\sharp z})}/\dot{z}][f,g]_\infty$, and because of

$$\sharp_B[f,g,(U)_p] \to (\ulcorner B \urcorner[\dot{\mathsf{c}}_{f(\overline{\sharp z})}/\dot{z}][f,g]_\infty^{(U)_p} \leftrightarrow \ulcorner B \urcorner[\dot{\mathsf{c}}_z/\dot{z}][f,g]_\infty^{(U)_p}),$$

we get $(\sharp_B[f,g,(U)_p] \to (z \in ((U)_{p'})_{\langle\!\langle \ulcorner B \urcorner, \overline{\sharp z}, f, g \rangle\!\rangle} \leftrightarrow B))$. Because of $p' \in p + \overline{2}^{\overline{n}}$ (i.e., $n > 0$ by definition) and by Theorem 102, the class $((U)_{p'})_{\langle\!\langle \ulcorner B \urcorner, \overline{\sharp z}, f, g \rangle\!\rangle}$ is contained in the stage $(U)_{p+\overline{2}^{\overline{n}}}$, hence we have

$$\sharp_B[f,g,(U)_p] \to (\exists x \forall y (y \in ((U)_{p+\overline{2}^{\overline{n}}})_x \leftrightarrow B[y/z])).$$

We finally get $A^*\langle p, p+\overline{2}^{\overline{n}} \rangle^U$ because of $\vec{X} \mathrel{\mathring{\in}} (U)_p$, i.e., because there exist f, g, such that $\sharp_B[f,g,(U)_p]$.

If $\Gamma = \Phi, A$ and $(B \to A) \in AC[\Pi^1_0]$ and $\mathcal{T}\ [\mathcal{R}]\!\vdash^{m,i}_{\mathcal{C},0} \Phi, B$ for some $m < n$ then there is some $C \in \Pi^1_0$, such that the asymmetric interpretations of A^* and B^* into the hierarchy U are

$$\begin{aligned} B^*\langle p, p+\overline{2}^{\overline{m}} \rangle^U &= \forall u \exists v C[u, ((U)_{p+\overline{2}^{\overline{m}}})_v], \\ A^*\langle p, p+\overline{2}^{\overline{n}} \rangle^U &= \exists y \forall u C[u, (((U)_{p+\overline{2}^{\overline{n}}})_y)_u]. \end{aligned}$$

W.l.o.g., we may assume $m+1 = n$. By i.h. we have

$$\Phi^{\vee *}\langle p, p+\overline{2}^{\overline{m}} \rangle^U \vee B^*\langle p, p+\overline{2}^{\overline{m}} \rangle^U,$$

and we need to show

$$\Phi^{\vee *}\langle p, p+\overline{2}^{\overline{n}} \rangle^U \vee A^*\langle p, p+\overline{2}^{\overline{n}} \rangle^U).$$

We assume $\neg(\Phi^{\vee *}\langle p, p+\overline{2}^{\overline{n}} \rangle^U)$, because otherwise we are done. By Lemma 108 we have $\neg(\Phi^{\vee *}\langle p, p+\overline{2}^{\overline{m}} \rangle^U)$, hence by i.h. we get $B^*\langle p, p+\overline{2}^{\overline{m}} \rangle^U$, i.e.,

$$\forall u \exists v C[u, ((U)_{p+\overline{2}^{\overline{m}}})_v].$$

We fix $q := p + \overline{2}^{\overline{m}}$ and we use the global wellordering $(W)_\emptyset$ (we have $Gl[(W)_\emptyset]$) to define $C_1 \in \Pi^1_0$, such that

$$\exists v C[u, ((U)_q)_v] \leftrightarrow \exists! v C_1[u, v, (U)_q, (W)_\emptyset],$$

that is, we define

$$C_1[u,v,Z_1,Z_2] \quad := \quad \begin{array}{c} C[u,(Z_1)_v] \wedge \forall w (C[u,(Z_1)_w] \to \\ \forall \gamma \forall \delta (\langle v, \gamma \rangle \in Z_2 \wedge \langle w, \delta \rangle \in Z_2 \to \gamma \in \delta'). \end{array}$$

Because of $\forall u \exists v C[u, ((U)_{p+\overline{2}^{\overline{m}}})_v]$ we get $\forall u \exists! v C_1[u, v, (U)_q, (W)_\emptyset]$, and based on C_1 we can now define the class function

$$F \quad := \quad \{\langle u, v\rangle \mid C_1[u, v, (U)_q, (W)_\emptyset]\}.$$

We define $C_2 \in \Pi^1_0$ such that

$$C_2[\langle w, u\rangle, (U)_q, (W)_\emptyset] \leftrightarrow C_1[u, F(u), (U)_q, (W)_\emptyset] \wedge w \in ((U)_q)_{F(u)},$$

that is, we define

$$C_2[x, Z_1, Z_2] \quad := \quad \exists w \exists u \exists v(x = \langle w, u\rangle \wedge C_1[u, v, Z_1, Z_2] \wedge w \in (Z_1)_v),$$

and we define $\theta_2 \in \mathcal{G}_{\Pi^1_0}$ such that $\theta_2[\dot{x}, \dot{Z}_1, \dot{Z}_2] := \ulcorner C_2[x, Z_1, Z_2]\urcorner$. Clearly, we can fix two functions f, g, such that $\sharp_C[f, g, (U)_{q'}]$, because of $\overrightarrow{X} \mathrel{\mathring{\in}} (U)_p$ and $p \in q'$. Having θ_2, and f, g, we now define the class

$$Z \quad := \quad ((U)_{p+\overline{2}^{\overline{n}}})_{\langle\!\langle \theta_2[\dot{x}, \dot{\mathsf{c}}_q, \dot{\mathsf{c}}_{\langle\emptyset,\emptyset\rangle}], \overline{\sharp x}, f, g\rangle\!\rangle},$$

hence by Theorem 102, because of $\mathsf{rk}(C_2) < i + j$, and $q' + \overline{1} = p + \overline{2}^{\overline{n}}$, we get

$$x \in Z \leftrightarrow \theta_2[\dot{x}, \dot{\mathsf{c}}_q, \dot{\mathsf{c}}_{\langle\emptyset,\emptyset\rangle}][f, g]^{(U)_{q'}}_\infty,$$

and by Lemma 88 and 90, and Theorem 102, we have

$$x \in Z \leftrightarrow C_2[x, (U)_q, (W)_\emptyset],$$

$$w \in (Z)_u \leftrightarrow \exists v(C_1[u, v, (U)_q, (W)_\emptyset] \wedge w \in ((U)_q)_v),$$

$$w \in (Z)_u \leftrightarrow w \in ((U)_q)_{F(u)}.$$

We have $C_1[u, F(u), (U)_q, (W)_\emptyset]$ by definition of F, and by definition of C_1 we get

$$\forall u(C[u, ((U)_q)_{F(u)}]),$$

hence $\forall u(C[u, (Z)_u])$ because of $(Z)_u = ((U)_q)_{F(u)}$. Finally, because of the definition of Z, we get $\exists y \forall u C[u, (((U)_{p+\overline{2}^{\overline{n}}})_y)_u]$, that is, $A^*\langle p, p + \overline{2}^{\overline{n}}\rangle^U$.

(2) By induction on n, considering all cases in Definition 12.

a) If $\Gamma = \Phi, A, \neg A$ and A is atomic, then we have $(A \vee \neg A)^*$ and because of $(A \vee \neg A)^* = (A \vee \neg A)^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U$ we get $\Gamma^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U$.

b) If $\Gamma = \Phi, A$ and $A \in CA[\Pi^1_0]$ then $A^* = \exists Z \forall y(y \in Z \leftrightarrow B[y/z])$ for some $B \in \Pi^1_0$, and the asymmetric interpretation of A^* into the hierarchy U is such that

$$A^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U = \exists x \forall y(y \in ((U)_{\mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}})_x \leftrightarrow B[y/z]).$$

By Theorem 102 we can find a class contained in the stage $(U)_{\mathfrak{a}\hat{+}\hat{1}}$, such that

$$z \in ((U)_{\mathfrak{a}\hat{+}\hat{1}})_{\langle\!\langle \ulcorner B \urcorner, \overline{\sharp z}, f, g \rangle\!\rangle} \leftrightarrow \ulcorner B \urcorner[\dot{\mathsf{c}}_z/\dot{z}][f,g]_\infty^{(U)_\mathfrak{a}},$$

and by Lemma 90, we have for the elementary formula B that

$$\sharp_B[f,g,(U)_\mathfrak{a}] \to (B \leftrightarrow \ulcorner B \urcorner[f,g]_\infty^{(U)_\mathfrak{a}}),$$

hence because of $\ulcorner B \urcorner[f,g]_\infty = \ulcorner B \urcorner[\dot{\mathsf{c}}_{f(\overline{\sharp z})}/\dot{z}][f,g]_\infty$, and because of

$$\sharp_B[f,g,(U)_\mathfrak{a}] \to (\ulcorner B \urcorner[\dot{\mathsf{c}}_{f(\overline{\sharp z})}/\dot{z}][f,g]_\infty^{(U)_\mathfrak{a}} \leftrightarrow \ulcorner B \urcorner[\dot{\mathsf{c}}_z/\dot{z}][f,g]_\infty^{(U)_\mathfrak{a}}),$$

we get $(\sharp_B[f,g,(U)_\mathfrak{a}] \to (z \in ((U)_{\mathfrak{a}\hat{+}\hat{1}})_{\langle\!\langle \ulcorner B \urcorner, \overline{\sharp z}, f, g \rangle\!\rangle} \leftrightarrow B))$. By Theorem 102 the class $((U)_{\mathfrak{a}\hat{+}\hat{1}})_{\langle\!\langle \ulcorner B \urcorner, \overline{\sharp z}, f, g \rangle\!\rangle}$ is contained in the stage $(U)_{\mathfrak{a}\hat{+}\Omega^{\hat{n}}\hat{+}\hat{2}}$, hence we have

$$\sharp_B[f,g,(U)_\mathfrak{a}] \to (\exists x \forall y (y \in ((U)_{\mathfrak{a}\hat{+}\Omega^{\hat{n}}\hat{+}\hat{2}})_x \leftrightarrow B[y/z])),$$

and we finally get $A^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U$ because of $\overrightarrow{X} \mathrel{\mathring{\in}} (U)_\mathfrak{a}$, i.e., because there exist f, g, such that $\sharp_B[f,g,(U)_\mathfrak{a}]$.

c) If $\Gamma = \Phi, A$ and $(B \to A) \in DC_{\mathcal{O}n}[\Pi^1_0]$ and $\mathcal{T}\ [\mathcal{R}] \vdash^m_{\mathcal{D},0} \Phi, B$ for some $m < n$ then there is some $C \in \Pi^1_0$ such that the asymmetric interpretations of A^* and B^* into the hierarchy U are

$$\begin{aligned} B^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U &= \forall \alpha \forall u \exists v C[\alpha, ((U)_\mathfrak{a})_u, ((U)_{\mathfrak{a}\hat{+}\Omega^{\hat{m}}\hat{+}\hat{2}})_v], \\ A^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U &= \exists y \forall \alpha C[\alpha, (((U)_{\mathfrak{a}\hat{+}\Omega^{\hat{n}}\hat{+}\hat{2}})_y)^\alpha, (((U)_{\mathfrak{a}\hat{+}\Omega^{\hat{n}}\hat{+}\hat{2}})_y)_\alpha]. \end{aligned}$$

By i.h. we have that

$$\forall \mathfrak{a} (\mathfrak{a} \hat{+} \Omega^{\widehat{m}} \lhd \Omega^{\widehat{k}} \wedge \overrightarrow{X} \mathrel{\mathring{\in}} (U)_\mathfrak{a} \to \\ \Phi^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U \vee B^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U),$$

and we need to show

$$\forall \mathfrak{a} (\mathfrak{a} \hat{+} \Omega^{\widehat{n}} \lhd \Omega^{\widehat{k}} \wedge \overrightarrow{X} \mathrel{\mathring{\in}} (U)_\mathfrak{a} \to \\ \Phi^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U \vee A^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U).$$

We fix some $\mathfrak{a}$ and we assume $\mathfrak{a} \hat{+} \Omega^{\widehat{n}} \lhd \Omega^{\widehat{k}}$, and $\overrightarrow{X} \mathrel{\mathring{\in}} (U)_\mathfrak{a}$, and

$$\neg(\Phi^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U),$$

because otherwise we are done. For any $\mathfrak{d}$ with $\mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\widehat{m}} \unlhd \mathfrak{a} \hat{+} \Omega^{\widehat{n}}$ we get $\mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\widehat{m}} \lhd \Omega^{\widehat{k}}$, and $\overrightarrow{X} \mathrel{\mathring{\in}} (U)_{(\mathfrak{a}\hat{+}\mathfrak{d})}$ by Theorem 102. Hence by i.h. we have

$$B^*\langle \mathfrak{a} \hat{+} \mathfrak{d}, \mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U,$$

because otherwise we must have $\Phi^{\vee *}\langle \mathfrak{a} \hat{+} \mathfrak{d}, \mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2}\rangle^{U}$, and in contradiction to our assumption we get $\Phi^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^{U}$ by Lemma 108. Hence for any β and $\mathfrak{d} = \Omega^{\widehat{m}} \hat{\cdot} \widehat{\beta} \hat{+} \widehat{5}$ (i.e., $\mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\widehat{m}} \lhd \mathfrak{a} \hat{+} \Omega^{\widehat{n}}$), we have $B^{*}\langle \mathfrak{a} \hat{+} \mathfrak{d}, \mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2}\rangle^{U}$, that is,

$$\forall\alpha\forall\beta\forall u\exists v C[\alpha, ((U)_{\mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}\widehat{\beta}\hat{+}\widehat{5}})_u, ((U)_{\mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}(\widehat{\beta}\hat{+}\widehat{1})\hat{+}\widehat{2}})_v].$$

We fix $\mathfrak{e} := \mathfrak{a} \hat{+} \Omega^{\widehat{n}}$ and we define $C_1 \in \Pi^1_0$ such that

$$\begin{gathered} C_1[\alpha, \beta, u, v, (U)_{\mathfrak{e}}] \leftrightarrow \\ C[\alpha, ((U)_{\mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}\widehat{\beta}\hat{+}\widehat{5}})_u, ((U)_{\mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}(\widehat{\beta}\hat{+}\widehat{1})\hat{+}\widehat{2}})_v], \end{gathered}$$

that is, we define

$$C_1[\alpha, \beta, u, v, Z_1] \quad := \quad C[\alpha, (Z_1)_{\langle \mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}\widehat{\beta}\hat{+}\widehat{5}, u\rangle}, (Z_1)_{\langle \mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}(\widehat{\beta}\hat{+}\widehat{1})\hat{+}\widehat{2}, v\rangle}]$$

(e.g., $((U)_{\mathfrak{e}})_{\langle \mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}\widehat{\beta}\hat{+}\widehat{5}, u\rangle} = ((U)_{\mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}\widehat{\beta}\hat{+}\widehat{5}})_u$ by Theorem 102). We further use the global wellordering $(W)_{\emptyset}$ (we have $Gl[(W)_{\emptyset}]$) to define $C_2 \in \Pi^1_0$ such that

$$\exists v C_1[\alpha, \beta, u, v, (U)_{\mathfrak{e}}] \leftrightarrow \exists! v C_2[\alpha, \beta, u, v, (U)_{\mathfrak{e}}, (W)_{\emptyset}],$$

that is, we define

$$\begin{gathered} C_2[\alpha, \beta, u, v, Z_1, Z_2] := \\ C_1[\alpha, \beta, u, v, Z_1] \wedge \forall w(C_1[\alpha, \beta, u, w, Z_1] \rightarrow \\ \forall\gamma\forall\delta(\langle v, \gamma\rangle \in Z_2 \wedge \langle w, \delta\rangle \in Z_2 \rightarrow \gamma \in \delta'). \end{gathered}$$

In the following we also need $C_3 \in \Pi^1_0$ such that

$$C_3[\langle w, \gamma\rangle, \beta, h, Z_3] \leftrightarrow w \in (Z_3)_{\langle \mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}(\widehat{\gamma}\hat{+}\widehat{1})\hat{+}\widehat{2}, h(\gamma)\rangle} \wedge \gamma \in \beta,$$

that is, we define

$$C_3[x, \beta, h, Z_3] \quad := \quad \begin{gathered} \exists w(\exists\gamma \in \beta)(x = \langle w, \gamma\rangle \wedge \\ w \in (Z_3)_{\langle \mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}(\widehat{\gamma}\hat{+}\widehat{1})\hat{+}\widehat{2}, h(\gamma)\rangle}). \end{gathered}$$

Based on C_2 and C_3 we now define a formula $C_4 \in \Pi^1_0$ such that, for the class function

$$F \quad = \quad \{\langle \alpha, h\rangle \mid C_4[\alpha, h, (U)_{\mathfrak{e}}, (W)_{\emptyset}]\},$$

and for the class $Z = \{\langle x, \alpha\rangle \mid x \in ((U)_{\mathfrak{e}})_{\langle \mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}(\widehat{\alpha}\hat{+}\widehat{1})\hat{+}\widehat{2}, F(\alpha)(\alpha)\rangle}\}$, we can show that $\forall\alpha(C[\alpha, (Z)^{\alpha}, (Z)_{\alpha}])$ and $Z \mathrel{\mathring{\in}} (U)_{\mathfrak{a}\hat{+}\Omega^{\widehat{n}}\hat{+}\widehat{2}}$. For the definition of C_4 we fix two arbitrary sets f_0, g_0, and we define $\theta_3 \in \mathcal{G}_{\Pi^1_0}$ such that $\theta_3[\dot{x}, \dot{\beta}, \dot{h}, \dot{Z}_3] := \ulcorner C_3[x, \beta, h, Z_3] \urcorner[\dot{\mathsf{c}}_{\mathfrak{a}}/\dot{\mathfrak{a}}]$, and finally we define

$$\begin{gathered} C_4[\alpha, h, Z_1, Z_2] := \\ Fun[h] \wedge dom(h) = \alpha' \wedge (\forall \beta \in \alpha')(\\ C_2[\beta, \beta, \langle\!\langle \theta_3[\dot{x}, \dot{\mathsf{c}}_\beta, \dot{\mathsf{c}}_{h\restriction\beta}, \dot{\mathsf{C}}_{\mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}\widehat{\beta}\hat{+}\widehat{3}}], \overline{\sharp x}, f_0, g_0 \rangle\!\rangle, h(\beta), Z_1, Z_2]). \end{gathered}$$

Hence we have $C_4[\alpha, h, (U)_{\mathfrak{e}}, (W)_\emptyset] \to (\forall \gamma \in \alpha) C_4[\gamma, h\restriction\gamma', (U)_{\mathfrak{e}}, (W)_\emptyset]$, and by elementary induction on α we get

$$\forall \alpha \exists! h C_4[\alpha, h, (U)_{\mathfrak{e}}, (W)_\emptyset].$$

We define the class function

$$F \quad := \quad \{\langle \alpha, h \rangle \mid C_4[\alpha, h, (U)_{\mathfrak{e}}, (W)_\emptyset]\},$$

hence $\forall \alpha (\forall \beta \in \alpha) F(\alpha)(\beta) = F(\beta)(\beta)$. We further define $C_5 \in \Pi^1_0$ such that

$$C_5[\langle w, \alpha \rangle, (U)_{\mathfrak{e}}, (W)_\emptyset] \leftrightarrow w \in ((U)_{\mathfrak{e}})_{\langle \mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}(\widehat{\alpha}\hat{+}\widehat{1})\hat{+}\widehat{2}, F(\alpha)(\alpha) \rangle},$$

that is, we define

$$C_5[x, Z_1, Z_2] \quad := \quad \begin{gathered} \exists \alpha \exists w \exists h (x = \langle w, \alpha \rangle \wedge C_4[\alpha, h, Z_1, Z_2] \wedge \\ w \in (Z_1)_{\langle \mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}(\widehat{\alpha}\hat{+}\widehat{1})\hat{+}\widehat{2}, h(\alpha) \rangle}), \end{gathered}$$

and $\theta_5[\dot{x}, \dot{Z}_1, \dot{Z}_2] := \ulcorner C_5[x, Z_1, Z_2] \urcorner[\dot{\mathsf{c}}_{\mathfrak{a}}/\dot{\mathfrak{a}}][\dot{\mathsf{c}}_{f_0}/\dot{f}_0][\dot{\mathsf{c}}_{g_0}/\dot{g}_0]$. Clearly we can find f_1, g_1, such that $\sharp_C[f_1, g_1, (U)_{\mathfrak{e}\hat{+}\hat{1}}]$, because of $\overrightarrow{X} \mathring{\in} (U)_{\mathfrak{a}}$. Based on θ_5, f_1, g_1, we define the class

$$Z \quad := \quad ((U)_{\mathfrak{a}\hat{+}\Omega^{\hat{n}}\hat{+}\widehat{2}})_{\langle\!\langle \theta_5[\dot{x}, \dot{\mathsf{C}}_{\mathfrak{e}}, \dot{\mathsf{C}}_{\langle \emptyset, \emptyset \rangle}], \overline{\sharp x}, f_1, g_1 \rangle\!\rangle},$$

hence by Theorem 102 we get $x \in Z \leftrightarrow \theta_5[\dot{x}, \dot{\mathsf{C}}_{\mathfrak{e}}, \dot{\mathsf{C}}_{\langle \emptyset, \emptyset \rangle}][f_1, g_1]_\infty^{(U)_{\mathfrak{e}\hat{+}\hat{1}}}$, and by Lemma 88 and 90 and Theorem 102 we have

$$x \in Z \leftrightarrow C_5[x, (U)_{\mathfrak{e}}, (W)_\emptyset],$$

that is, we have $(Z)_\alpha = ((U)_{\mathfrak{e}})_{\langle \mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}(\widehat{\alpha}\hat{+}\widehat{1})\hat{+}\widehat{2}, F(\alpha)(\alpha) \rangle}$ for any α. By definition of F we have $C_4[\alpha, F(\alpha), (U)_{\mathfrak{e}}, (W)_\emptyset]$ for any α, that is,

$$C_2[\alpha, \alpha, \langle\!\langle \theta_3[\dot{x}, \dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_{F(\alpha)\restriction\alpha}, \dot{\mathsf{C}}_{\mathfrak{a}\hat{+}\Omega^{\widehat{m}}\hat{\cdot}\widehat{\alpha}\hat{+}\widehat{3}}], \overline{\sharp x}, f_0, g_0 \rangle\!\rangle, F(\alpha)(\alpha), (U)_{\mathfrak{e}}, (W)_\emptyset],$$

hence by definition of C_2 and by Theorem 102 we get

$$\forall \alpha (C[\alpha, (Y)_\alpha, (Z)_\alpha])$$

for any Y such that

$$(Y)_\alpha \quad = \quad ((U)_{\mathfrak{e}})_{\langle \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{:} \widehat{\alpha} \hat{+} \widehat{5}, \langle\!\langle \theta_3[\dot{x}, \dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_{F(\alpha)\restriction\alpha}, \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{:} \widehat{\alpha} \hat{+} \widehat{3}}], \overline{\sharp x}, f_0, g_0 \rangle\!\rangle \rangle},$$

and for such Y we further have by Theorem 102 and Lemma 88 that

$$x \in (Y)_\alpha \leftrightarrow \theta_3[\dot{\mathsf{c}}_x, \dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_{F(\alpha)\restriction\alpha}, \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{:} \widehat{\alpha} \hat{+} \widehat{3}}][f_0, g_0]_\infty^{(U)_{\mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{:} \widehat{\alpha} \hat{+} \widehat{4}}},$$

$$x \in (Y)_\alpha \leftrightarrow C_3[x, \alpha, F(\alpha)\restriction\alpha, (U)_{\mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{:} \widehat{\alpha} \hat{+} \widehat{3}}],$$

$$x \in (Y)_\alpha \leftrightarrow \exists w (\exists \gamma \in \alpha)(x = \langle w, \gamma \rangle \wedge w \in ((U)_{\mathfrak{e}})_{\langle \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{:} (\widehat{\gamma} \hat{+} \widehat{1}) \hat{+} \widehat{2}, F(\alpha)(\gamma) \rangle}),$$

that is, $(Y)_\alpha = (Z)^\alpha$. Hence we have $\forall \alpha (C[\alpha, (Z)^\alpha, (Z)_\alpha])$, and by definition of Z we finally get

$$\exists y \forall \alpha C[\alpha, (((U)_{\mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}})_y)^\alpha, (((U)_{\mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}})_y)_\alpha],$$

that is $A^* \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U$.

d) If $\Gamma = \Phi, A \wedge B$ and $\mathcal{T}\ [\mathcal{R}] \Vdash^m_{\mathcal{D},0} \Phi, A$ and $\mathcal{T}\ [\mathcal{R}] \Vdash^m_{\mathcal{D},0} \Phi, B$ for some $m < n$, then by i.h. and logic we have

$$\Phi^{\vee *} \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U \vee$$
$$(A^* \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U \wedge B^* \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U),$$

i.e. $\Gamma^{\vee *} \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U$, and $\Gamma^{\vee *} \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U$ by Lemma 108.

e) If $\Gamma = \Phi, A \vee B$, and $\mathcal{T}\ [\mathcal{R}] \Vdash^m_{\mathcal{D},0} \Phi, A$ or $\mathcal{T}\ [\mathcal{R}] \Vdash^m_{\mathcal{D},0} \Phi, B$ for some $m < n$, then we proceed analogously to the previous case.

f) If $\Gamma = \Phi, \exists x A[x/v]$ and $\mathcal{T}\ [\mathcal{R}] \Vdash^m_{\mathcal{D},0} \Phi, A[w/v]$ for some $m < n$, then we proceed analogously to the next case.

g) If $\Gamma = \Phi, \forall x A[x/u]$ and $\mathcal{T}\ [\mathcal{R}] \Vdash^m_{\mathcal{D},0} \Phi, A$ for some $m < n$ and $u \notin \Gamma$ then by i.h. we have

$$\Phi^{\vee *} \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U \vee A^* \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U,$$

and by Lemma 108 and because of $u \notin \Gamma$ we get

$$\Phi^{\vee *} \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U \vee \forall y (A^* \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U [y/u^*]),$$

that is $\Gamma^{\vee *} \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2} \rangle^U$.

h) If $\Gamma = \Phi, \exists X A[X/Y]$ and $\mathcal{T}\ [\mathcal{R}] \Vdash^m_{\mathcal{D},0} \Phi, A[Z/Y]$ for some $m < n$ then by i.h. we have

$$Z^* \mathrel{\mathring{\in}} (U)_{\mathfrak{a}} \to (\Phi^{\vee *} \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U \vee A[Z/Y]^* \langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2} \rangle^U),$$

hence by Lemma 108, and Theorem 102, and by logic we get

$$Z^* \mathrel{\mathring{\in}} (U)_{\mathfrak{a}} \to (\Phi^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U \vee \\ \exists z(A^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U [((U)_{\mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}})_z / Y^*])),$$

that is $Z^* \mathrel{\mathring{\in}} (U)_{\mathfrak{a}} \to \Gamma^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U$. We either have $Z^* \in \{X_0, \ldots, X_l\}$, or $Z^* \notin \Gamma^{\vee *}$ and $\exists Z(Z \mathrel{\mathring{\in}} (U)_{\mathfrak{a}})$, hence in both cases

$$\vec{X} \mathrel{\mathring{\in}} (U)_{\mathfrak{a}} \to \Gamma^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U.$$

i) If $\Gamma = \Phi, \forall X A[X/Z]$ and $\mathcal{T}\ [\mathcal{R}]|\!\!\frac{m}{\mathcal{D},0}\ \Phi, A$ for some $m < n$ and $Z \notin \Gamma$ then by i.h. we have

$$Z^* \mathrel{\mathring{\in}} (U)_{\mathfrak{a}} \to (\Phi^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2}\rangle^U \vee A^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2}\rangle^U),$$

and by Lemma 108 and because of $Z \notin \Gamma$ we get

$$\Phi^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U \vee \\ \forall Y(Y \mathrel{\mathring{\in}} (U)_{\mathfrak{a}} \to A^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U [Y/Z^*]),$$

hence by logic we have

$$\Phi^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2}\rangle^U \vee \forall y(A^*\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U [((U)_{\mathfrak{a}})_y / Z^*]),$$

that is $\Gamma^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U$.

j) If $\mathcal{T}\ [\mathcal{R}]|\!\!\frac{m}{\mathcal{D},0}\ \Gamma, A$ and $\mathcal{T}\ [\mathcal{R}]|\!\!\frac{m}{\mathcal{D},0}\ \Gamma, \neg A$ for some $m < n$ and $A \in \mathcal{D}$, then $A = \exists X B[X/Y]$ for some $B \in \Pi^1_0$. We define $\mathfrak{a}_0 := \mathfrak{a} \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2}$, hence by i.h. we get

$$\Gamma^{\vee *}\langle \mathfrak{a}, \mathfrak{a}_0\rangle^U \vee A^*\langle \mathfrak{a}, \mathfrak{a}_0\rangle^U$$

and

$$\Gamma^{\vee *}\langle \mathfrak{a}_0, \mathfrak{a}_0 \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2}\rangle^U \vee (\neg A^*)\langle \mathfrak{a}_0, \mathfrak{a}_0 \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2}\rangle^U,$$

hence by Lemma 108 we get

$$\Gamma^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U \vee \\ (A^*\langle \mathfrak{a}, \mathfrak{a}_0\rangle^U \wedge (\neg A^*)\langle \mathfrak{a}_0, \mathfrak{a}_0 \hat{+} \Omega^{\widehat{m}} \hat{+} \widehat{2}\rangle^U),$$

that is

$$\Gamma^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U \vee \\ (\exists x(B^*[((U)_{\mathfrak{a}_0})_x / Y^*]) \wedge \forall x((\neg B^*)[((U)_{\mathfrak{a}_0})_x / Y^*])),$$

and we finally have $\Gamma^{\vee *}\langle \mathfrak{a}, \mathfrak{a} \hat{+} \Omega^{\widehat{n}} \hat{+} \widehat{2}\rangle^U$.

(3) We have that $\psi[f,g]_\infty^{(U)_{\Omega_k}}$ is equivalent to $\psi[f,g]_\infty \in (V)_{\Omega_k}$, hence the claim can be proved by elementary induction on $\mathfrak{b}$, considering all cases (i.e. disjuncts) in Definition 91.

a) If $\{\psi, \dot{\sim}\psi\} \subseteq y$ then $(\dot{\sim}\psi)^* = \neg(\psi^*)$ and $rk(\psi^*) = \overline{1}$ and because of $Cl_{\Omega_{k+1}}[U,V,W]$ we have $\psi^*[f,g]_\infty^{(U)_{\Omega_k}} \vee (\neg(\psi^*))[f,g]_\infty^{(U)_{\Omega_k}}$ by Lemma 87.

b) If $\psi \in y$ for some $\psi \in \mathcal{G}_{CA[\Pi^1_0]}$ then $\psi^* = \dot{\exists}\dot{\mathtt{U}}_p \dot{\forall}\dot{\mathtt{u}}_q(\dot{\mathtt{u}}_q \dot{\in} \dot{\mathtt{U}}_p \dot{\leftrightarrow} \theta[\dot{\mathtt{u}}_q/\dot{\mathtt{v}}_r])$ for some $\theta \in \mathcal{G}_{\Pi^1_0}$, and by Theorem 102 we have

$$x \in (((U)_{\Omega_k})_{\mathfrak{a}\hat{+}\hat{2}})_{\langle\!\langle\theta,r,f,g\rangle\!\rangle} \leftrightarrow \theta[\dot{\mathtt{c}}_x/\dot{\mathtt{v}}_r][f,g]_\infty^{((U)_{\Omega_k})_{\mathfrak{a}\hat{+}\hat{1}}}.$$

Because of $g|\mathfrak{a}$ we have

$$\theta[\dot{\mathtt{c}}_x/\dot{\mathtt{v}}_r][f,g]_\infty^{((U)_{\Omega_k})_{\mathfrak{a}\hat{+}\hat{1}}} \leftrightarrow \theta[\dot{\mathtt{c}}_x/\dot{\mathtt{v}}_r][f,g]_\infty^{(U)_{\Omega_k}},$$

hence

$$x \in ((U)_{\Omega_k})_{\langle\mathfrak{a}\hat{+}\hat{2},\langle\!\langle\theta,r,f,g\rangle\!\rangle\rangle} \leftrightarrow \theta[\dot{\mathtt{c}}_x/\dot{\mathtt{v}}_r][f,g]_\infty^{(U)_{\Omega_k}}$$

for any x. By Lemma 87 and 88 we get

$$(\dot{\mathtt{c}}_x \dot{\in} \dot{\mathtt{C}}_{\langle\mathfrak{a}\hat{+}\hat{2},\langle\!\langle\theta,r,f,g\rangle\!\rangle\rangle})[f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow \theta[\dot{\mathtt{c}}_x/\dot{\mathtt{v}}_r][f,g]_\infty^{(U)_{\Omega_k}},$$

$$\forall x((\dot{\mathtt{c}}_x \dot{\in} \dot{\mathtt{C}}_{\langle\mathfrak{a}\hat{+}\hat{2},\langle\!\langle\theta,r,f,g\rangle\!\rangle\rangle} \dot{\leftrightarrow} \theta[\dot{\mathtt{c}}_x/\dot{\mathtt{v}}_r])[f,g]_\infty^{(U)_{\Omega_k}}),$$

$$\dot{\forall}\dot{\mathtt{u}}_q(\dot{\mathtt{u}}_q \dot{\in} \dot{\mathtt{C}}_{\langle\mathfrak{a}\hat{+}\hat{2},\langle\!\langle\theta,r,f,g\rangle\!\rangle\rangle} \dot{\leftrightarrow} \theta[\dot{\mathtt{u}}_q/\dot{\mathtt{v}}_r])[f,g]_\infty^{(U)_{\Omega_k}}.$$

By Corollary 106 and for $\dot{\mathtt{V}}_t$ with $\dot{\mathtt{V}}_t \notin term(\theta)$ we have

$$\exists y(\dot{\forall}\dot{\mathtt{u}}_q(\dot{\mathtt{u}}_q \dot{\in} \dot{\mathtt{V}}_t \dot{\leftrightarrow} \theta[\dot{\mathtt{u}}_q/\dot{\mathtt{v}}_r])[(\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}})_{\dot{\mathtt{c}}_y}/\dot{\mathtt{V}}_t][f,g]_\infty^{(U)_{\Omega_k}}),$$

because there is some y such that

$$((U)_{\Omega_k})_{\langle\mathfrak{a}\hat{+}\hat{2},\langle\!\langle\theta,r,f,g\rangle\!\rangle\rangle} = (((U)_{\Omega_k})_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}})_y.$$

Hence by Lemma 88 we finally get $\psi^*\langle\dot{\mathtt{C}}_\mathfrak{a}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}$.

c) If $y = z \cup \{\psi\}$ and $(\xi \dot{\rightarrow} \psi) \in \mathcal{G}_{DC_{\mathcal{O}n}[\Pi^1_0]}$, and $z \cup \{\xi\} \in (Z)_{\mathfrak{c},\emptyset}$ for some $\mathfrak{c} \lhd \mathfrak{b}$ then there is some $\theta \in \mathcal{G}_{\Pi^1_0}$ such that

$$\begin{array}{rcl} \xi & = & \dot{\forall}\dot{\mathtt{u}}_p(\ulcorner x \in \mathcal{O}n\urcorner[\dot{\mathtt{u}}_p/\dot{x}] \dot{\rightarrow} \dot{\forall}\dot{\mathtt{U}}_q \dot{\exists}\dot{\mathtt{U}}_r \theta[\dot{\mathtt{u}}_p, \dot{\mathtt{U}}_q, \dot{\mathtt{U}}_r]), \\ \psi & = & \dot{\exists}\dot{\mathtt{U}}_r \dot{\forall}\dot{\mathtt{u}}_p(\ulcorner x \in \mathcal{O}n\urcorner[\dot{\mathtt{u}}_p/\dot{x}] \dot{\rightarrow} \theta[\dot{\mathtt{u}}_p, (\dot{\mathtt{U}}_r)^{\dot{\mathtt{u}}_p}, (\dot{\mathtt{U}}_r)_{\dot{\mathtt{u}}_p}]). \end{array}$$

By i.h. we have

$$\forall\mathfrak{a}(\mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \lhd \Omega_k \wedge g|\mathfrak{a} \rightarrow$$
$$(\exists\phi \in z \cup \{\xi\})\phi^*\langle\dot{\mathtt{C}}_\mathfrak{a}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}),$$

and we need to show

$$\forall \mathfrak{a}(\mathfrak{a} \hat{+} \Omega^{\mathfrak{b}} \lhd \Omega_k \wedge g|\mathfrak{a} \rightarrow \\ (\exists \phi \in z \cup \{\psi\})\phi^*\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \Omega^{\mathfrak{b}} \hat{+} \widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}).$$

We fix some $\mathfrak{a}, f, g$ with $\mathfrak{a} \hat{+} \Omega^{\mathfrak{b}} \lhd \Omega_k$ and $g|\mathfrak{a}$, and we assume

$$(\forall \phi \in z)\neg(\phi^*\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \Omega^{\mathfrak{b}} \hat{+} \widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}),$$

because otherwise $(\exists \phi \in z)\phi^*\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \Omega^{\mathfrak{b}} \hat{+} \widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}$, and we are done. If $\mathfrak{d}$ is such that $\mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\mathfrak{c}} \trianglelefteq \mathfrak{a} \hat{+} \Omega^{\mathfrak{b}}$ then $\mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\mathfrak{c}} \lhd \Omega_k$, and $g|(\mathfrak{a} \hat{+} \mathfrak{d})$, and by i.h. we get

$$\xi^*\langle \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \mathfrak{d}}, \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\mathfrak{c}} \hat{+} \widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}},$$

because otherwise $\phi^*\langle \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \mathfrak{d}}, \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\mathfrak{c}} \hat{+} \widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}$ for some $\phi \in z$, that is $\phi^*\langle \dot{\mathsf{C}}_{\mathfrak{a}}, \dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \Omega^{\mathfrak{b}} \hat{+} \widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}$ by Lemma 108, in contradiction to our assumption. Hence by Lemma 88, 90 and Corollary 106 we have

$$\forall \alpha \forall u \exists v(\theta^*[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{C}}_{\langle \mathfrak{a} \hat{+} \mathfrak{d}, u\rangle}, \dot{\mathsf{C}}_{\langle \mathfrak{a} \hat{+} \mathfrak{d} \hat{+} \Omega^{\mathfrak{c}} \hat{+} \widehat{2}, v\rangle}][f,g]_\infty^{(U)_{\Omega_k}}),$$

and hence, because of $\mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} \widehat{\beta} \hat{+} \widehat{5} \hat{+} \Omega^{\mathfrak{c}} \lhd \mathfrak{a} \hat{+} \Omega^{\mathfrak{b}}$ for any β, we get

$$\forall \alpha \forall \beta \forall u \exists v(\theta^*[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{C}}_{\langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} \widehat{\beta} \hat{+} \widehat{5}, u\rangle}, \dot{\mathsf{C}}_{\langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} (\widehat{\beta} \hat{+} \widehat{1}) \hat{+} \widehat{2}, v\rangle}][f,g]_\infty^{(U)_{\Omega_k}}).$$

We fix $\mathfrak{e} := \mathfrak{a} \hat{+} \Omega^{\mathfrak{b}}$ and we define $\theta_1 \in \mathcal{G}_{\Pi^1_0}$ such that

$$\theta_1[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_\beta, \dot{\mathsf{c}}_u, \dot{\mathsf{c}}_v][f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow \\ \theta^*[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{C}}_{\langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} \widehat{\beta} \hat{+} \widehat{5}, u\rangle}, \dot{\mathsf{C}}_{\langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} (\widehat{\beta} \hat{+} \widehat{1}) \hat{+} \widehat{2}, v\rangle}][f,g]_\infty^{(U)_{\Omega_k}},$$

that is, we define

$$\theta_1[\dot{\alpha}, \dot{\beta}, \dot{u}, \dot{v}] := \begin{array}{c} \dot{\forall} \dot{z_0} \dot{\forall} \dot{z_1}(\ulcorner z_0 = \langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} \widehat{\beta} \hat{+} \widehat{5}, u\rangle \urcorner[\dot{\mathsf{c}}_{\mathfrak{a}}/\dot{\mathfrak{a}}][\dot{\mathsf{c}}_{\mathfrak{c}}/\dot{\mathfrak{c}}] \dot{\wedge} \\ \ulcorner z_1 = \langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} (\widehat{\beta} \hat{+} \widehat{1}) \hat{+} \widehat{2}, v\rangle \urcorner[\dot{\mathsf{c}}_{\mathfrak{a}}/\dot{\mathfrak{a}}][\dot{\mathsf{c}}_{\mathfrak{c}}/\dot{\mathfrak{c}}] \dot{\rightarrow} \\ \theta^*[\dot{\alpha}, (\dot{\mathsf{C}}_{\mathfrak{e}})_{\dot{z_0}}, (\dot{\mathsf{C}}_{\mathfrak{e}})_{\dot{z_1}}]), \end{array}$$

hence θ_1 has the desired property, by Lemma 88, 90, and Corollary 106. We further use the global wellordering $((U)_{\Omega_k})_{\langle \emptyset, \emptyset\rangle} = (W)_\emptyset$ (we have $Gl[(W)_\emptyset]$) to define $\theta_2 \in \mathcal{G}_{\Pi^1_0}$ such that

$$\exists v(\theta_1[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_\beta, \dot{\mathsf{c}}_u, \dot{\mathsf{c}}_v][f,g]_\infty^{(U)_{\Omega_k}}) \leftrightarrow \exists! v(\theta_2[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_\beta, \dot{\mathsf{c}}_u, \dot{\mathsf{c}}_v][f,g]_\infty^{(U)_{\Omega_k}}),$$

that is, we define

$$\theta_2[\dot\alpha,\dot\beta,\dot u,\dot v] \quad := \quad \begin{array}{c}\theta_1[\dot\alpha,\dot\beta,\dot u,\dot v]\mathbin{\dot\wedge}\dot\forall\,\dot w\,\dot\forall\,\dot\gamma\,\dot\forall\,\dot\delta(\theta_1[\dot\alpha,\dot\beta,\dot u,\dot w]\mathbin{\dot\to}\\ \ulcorner\langle v,\gamma\rangle\in X\wedge\langle w,\delta\rangle\in X\to\gamma\in\delta'\urcorner[\dot{\mathsf{C}}_{\langle\emptyset,\emptyset\rangle}/\dot X]).\end{array}$$

Based on θ_2 we now define $\theta_6\in\mathcal{G}_{\Pi^1_0}$ (and $\theta_3,\theta_4,\theta_5\in\mathcal{G}_{\Pi^1_0}$ as parts of θ_6) such that, for the class function

$$F \quad = \quad \{\langle\alpha,h\rangle \mid \theta_6[\dot{\mathsf{c}}_\alpha,\dot{\mathsf{c}}_h][f,g]_\infty\in(V)_{\Omega_k}\},$$

and for the class $Z=\{\langle x,\alpha\rangle\mid x\in((U)_{\mathfrak{c}})_{\langle\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}(\widehat{\alpha}\hat{+}\widehat{1})\hat{+}\widehat{2},F(\alpha)(\alpha)\rangle}\}$, we can show that there is some y such that $Z=((U)_{\Omega_k})_{\langle\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2},y\rangle}$ and

$$\forall\alpha(\theta^*[\dot{\mathsf{c}}_\alpha,(\dot{\mathsf{C}}_{\langle\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2},y\rangle})^{\dot{\mathsf{c}}_\alpha},(\dot{\mathsf{C}}_{\langle\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2},y\rangle})_{\dot{\mathsf{c}}_\alpha}][f,g]_\infty^{(U)_{\Omega_k}}).$$

We define

$$\theta_3[\dot x,\dot\beta,\dot h,\dot X] \quad := \quad \begin{array}{c}\dot\exists\,\dot w\,\dot\exists\,\dot\gamma(\ulcorner\gamma\in\beta\wedge x=\langle w,\gamma\rangle\urcorner\mathbin{\dot\wedge}\\ \ulcorner w\in(X)_{\langle\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}(\widehat{\gamma}\hat{+}\widehat{1})\hat{+}\widehat{2},h(\gamma)\rangle}\urcorner[\dot{\mathsf{c}}_{\mathfrak{a}}/\dot{\mathfrak{a}}][\dot{\mathsf{c}}_{\mathfrak{c}}/\dot{\mathfrak{c}}],\end{array}$$

$$\theta_4[\dot f_0,\dot\beta,\dot h] \quad := \quad \begin{array}{c}\ulcorner f_0(\overline{\sharp\beta})=\beta\wedge f_0(\overline{\sharp h})=h{\restriction}\beta\urcorner\mathbin{\dot\wedge}\\ \ulcorner\forall p(p=\overline{\sharp\beta}\vee p=\overline{\sharp h}\vee f_0(p)=f(p))\urcorner[\dot{\mathsf{c}}_f/\dot f],\end{array}$$

$$\theta_5[\dot g_0,\dot\beta] \quad := \quad \begin{array}{c}\ulcorner g_0(\overline{\sharp X})=\mathfrak{a}\mathbin{\hat{+}}\Omega^{\mathfrak{c}}\mathbin{\hat{\cdot}}\widehat{\beta}\mathbin{\hat{+}}\widehat{3}\urcorner[\dot{\mathsf{c}}_{\mathfrak{a}}/\dot{\mathfrak{a}}][\dot{\mathsf{c}}_{\mathfrak{c}}/\dot{\mathfrak{c}}]\mathbin{\dot\wedge}\\ \ulcorner\forall p(p=\overline{\sharp X}\vee g_0(p)=g(p))\urcorner[\dot{\mathsf{c}}_g/\dot g],\end{array}$$

$$\theta_6[\dot\alpha,\dot h] \quad := \quad \begin{array}{c}\ulcorner Fun[h]\wedge dom(h)=\alpha'\urcorner\mathbin{\dot\wedge}\dot\forall\,\dot\beta\,\dot\forall\,\dot u\,\dot\forall\,\dot v\,\dot\forall\,\dot f_0\,\dot\forall\,\dot g_0(\\ \ulcorner\beta\in\alpha'\urcorner\mathbin{\dot\wedge}\theta_4[\dot f_0,\dot\beta,\dot h]\mathbin{\dot\wedge}\theta_5[\dot g_0,\dot\beta]\mathbin{\dot\wedge}\\ \ulcorner u=\langle\!\langle u_0,u_1,f_0,g_0\rangle\!\rangle\urcorner[\dot{\mathsf{c}}_{\theta_3[\dot x,\dot\beta,\dot h,\dot X]}/\dot u_0][\dot{\mathsf{c}}_{\overline{\sharp x}}/\dot u_1]\mathbin{\dot\wedge}\\ \ulcorner v=h(\beta)\urcorner\mathbin{\dot\to}\theta_2[\dot\beta,\dot\beta,\dot u,\dot v]).\end{array}$$

In a first step we show the following equivalence for θ_6,

$$\begin{array}{c}\theta_6[\dot{\mathsf{c}}_\alpha,\dot{\mathsf{c}}_h][f,g]_\infty^{(U)_{\Omega_k}}\leftrightarrow Fun[h]\wedge dom(h)=\alpha'\wedge\\ (\forall\beta\in\alpha')(\theta_2[\dot{\mathsf{c}}_\beta,\dot{\mathsf{c}}_\beta,\dot{\mathsf{c}}_{\langle\!\langle\theta_3[\dot x,\dot{\mathsf{c}}_\beta,\dot{\mathsf{c}}_{h\restriction\beta},\dot{\mathsf{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}\widehat{\beta}\hat{+}\widehat{3}}],\overline{\sharp x},f,g\rangle\!\rangle},\dot{\mathsf{c}}_{h(\beta)}][f,g]_\infty^{(U)_{\Omega_k}}).\end{array}$$

If f_0,g_0 are such that $f_0(\overline{\sharp\beta})=\beta$, and $f_0(\overline{\sharp h})=h{\restriction}\beta$, and $f_0(p)=f(p)$ for $p\notin\{\overline{\sharp\beta},\overline{\sharp h}\}$, and $g_0(\overline{\sharp X})=\mathfrak{a}\mathbin{\hat{+}}\Omega^{\mathfrak{c}}\mathbin{\hat{\cdot}}\widehat{\beta}\mathbin{\hat{+}}\widehat{3}$, and $g_0(p)=g(p)$ for $p\neq\overline{\sharp X}$, then we have that

$$\theta_3[\dot x,\dot\beta,\dot h,\dot X][f_0,g_0]_\infty=\theta_3[\dot x,\dot{\mathsf{c}}_\beta,\dot{\mathsf{c}}_{h\restriction\beta},\dot{\mathsf{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}\widehat{\beta}\hat{+}\widehat{3}}][f,g]_\infty,$$

hence for such f_0, g_0, and any β, and $\mathfrak{a}_0=\mathfrak{a}\mathbin{\hat{+}}\Omega^{\mathfrak{c}}\mathbin{\hat{\cdot}}\widehat{\beta}$, we have that

$$\theta_3[\dot{\mathsf{c}}_x,\dot\beta,\dot h,\dot X][f_0,g_0]_\infty^{(U)_{\mathfrak{a}_0\hat{+}\widehat{4}}}\leftrightarrow\theta_3[\dot{\mathsf{c}}_x,\dot{\mathsf{c}}_\beta,\dot{\mathsf{c}}_{h\restriction\beta},\dot{\mathsf{C}}_{\mathfrak{a}_0\hat{+}\widehat{3}}][f,g]_\infty^{(U)_{\mathfrak{a}_0\hat{+}\widehat{4}}},$$

and by Theorem 102 we get

$$((U)_{\Omega_k})_{\langle \mathfrak{a}_0 \hat{+} \widehat{5}, \langle\!\langle \theta_3[\dot{x},\dot{\beta},\dot{h},\dot{X}],\overline{\sharp x},f_0,g_0 \rangle\!\rangle \rangle} = \\ ((U)_{\Omega_k})_{\langle \mathfrak{a}_0 \hat{+} \widehat{5}, \langle\!\langle \theta_3[\dot{x},\dot{\mathsf{c}}_\beta,\dot{\mathsf{c}}_{h\restriction\beta},\dot{\mathsf{C}}_{\mathfrak{a}_0 \hat{+} \widehat{3}}],\overline{\sharp x},f,g \rangle\!\rangle \rangle},$$

and because of this equality and by Lemma 105 we finally have

$$\theta_6[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_h][f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow \mathit{Fun}[h] \wedge \mathit{dom}(h) = \alpha' \wedge \\ (\forall \beta \in \alpha')(\theta_2[\dot{\mathsf{c}}_\beta, \dot{\mathsf{c}}_\beta, \dot{\mathsf{c}}_{\langle\!\langle \theta_3[\dot{x},\dot{\mathsf{c}}_\beta,\dot{\mathsf{c}}_{h\restriction\beta},\dot{\mathsf{C}}_{\mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} \widehat{\beta} \hat{+} \widehat{3}}],\overline{\sharp x},f,g \rangle\!\rangle}, \dot{\mathsf{c}}_{h(\beta)}][f,g]_\infty^{(U)_{\Omega_k}}).$$

Having this equivalence we get

$$\theta_6[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_h][f,g]_\infty^{(U)_{\Omega_k}} \to (\forall \gamma \in \alpha)(\theta_6[\dot{\mathsf{c}}_\gamma, \dot{\mathsf{c}}_{h\restriction\gamma'}][f,g]_\infty^{(U)_{\Omega_k}}),$$

and by elementary induction on α (because $\theta_6[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_h][f,g]_\infty^{(U)_{\Omega_k}}$ is equivalent to the elementary formula $\theta_6[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_h][f,g]_\infty \in (V)_{\Omega_k}$) we get

$$\forall \alpha \exists ! h(\theta_6[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_h][f,g]_\infty^{(U)_{\Omega_k}}).$$

We define the class function

$$F \quad := \quad \{\langle \alpha, h\rangle \mid \theta_6[\dot{\mathsf{c}}_\alpha, \dot{\mathsf{c}}_h][f,g]_\infty \in (V)_{\Omega_k}\},$$

hence $\forall \alpha (\forall \beta \in \alpha) F(\alpha)(\beta) = F(\beta)(\beta)$. We further define $\theta_7 \in \mathcal{G}_{\Pi^1_0}$ such that

$$\theta_7[\dot{\mathsf{c}}_{\langle w,\alpha\rangle}][f,g]_\infty^{(U)_{\Omega_k}} \leftrightarrow w \in ((U)_{\mathfrak{e}})_{\langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} (\widehat{\alpha} \hat{+} \widehat{1}) \hat{+} \widehat{2}, F(\alpha)(\alpha)\rangle},$$

that is, we define

$$\theta_7[\dot{x}] \quad := \quad \begin{array}{c} \dot{\exists}\, \dot{\alpha}\, \dot{\exists}\, \dot{w}\, \dot{\exists}\, \dot{h}(\ulcorner \alpha \in \mathcal{O}n \wedge x = \langle w, \alpha\rangle \urcorner \,\dot{\wedge}\, \theta_6[\dot{\alpha}, \dot{h}] \,\dot{\wedge} \\ \ulcorner w \in (X)_{\langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} (\widehat{\alpha} \hat{+} \widehat{1}) \hat{+} \widehat{2}, h(\alpha)\rangle} \urcorner [\dot{\mathsf{c}}_{\mathfrak{a}}/\dot{\mathfrak{a}}][\dot{\mathsf{c}}_{\mathfrak{c}}/\dot{\mathfrak{c}}][\dot{\mathsf{C}}_{\mathfrak{e}}/\dot{X}]), \end{array}$$

and based on θ_7 we define the class

$$Z \quad := \quad ((U)_{\Omega_k})_{\langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{b}} \hat{+} \widehat{2}, \langle\!\langle \theta_7[\dot{x}],\overline{\sharp x},f,g \rangle\!\rangle \rangle}.$$

We have $\theta_7[\dot{\mathsf{c}}_x][f,g]_\infty^{(U)_{\mathfrak{e} \hat{+} \widehat{1}}} \leftrightarrow \theta_7[\dot{\mathsf{c}}_x][f,g]_\infty^{(U)_{\Omega_k}}$ because of $g|\mathfrak{a}$, and $w \trianglelefteq \mathfrak{e}$ for all $\dot{\mathsf{C}}_w \in \mathit{term}(\theta_7[\dot{\mathsf{c}}_x])$, hence by Theorem 102 we get

$$x \in Z \leftrightarrow \theta_7[\dot{\mathsf{c}}_x][f,g]_\infty^{(U)_{\Omega_k}},$$

that is, we have $(Z)_\alpha = ((U)_{\mathfrak{e}})_{\langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} (\widehat{\alpha} \hat{+} \widehat{1}) \hat{+} \widehat{2}, F(\alpha)(\alpha)\rangle}$ for any α, and we further have that

$$x \in (Z)^\alpha \leftrightarrow \exists w (\exists \gamma \in \alpha)(x = \langle w, \gamma\rangle \wedge w \in ((U)_{\mathfrak{e}})_{\langle \mathfrak{a} \hat{+} \Omega^{\mathfrak{c}} \hat{\cdot} (\widehat{\gamma} \hat{+} \widehat{1}) \hat{+} \widehat{2}, F(\alpha)(\gamma)\rangle}),$$

hence by Theorem 102 we get

$$x \in (Z)^\alpha \leftrightarrow \theta_3[\dot{\mathtt{c}}_x, \dot{\mathtt{c}}_\alpha, \dot{\mathtt{c}}_{F(\alpha)\restriction\alpha}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}\widehat{\alpha}\hat{+}\widehat{3}}][f,g]_\infty^{(U)_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}\widehat{\alpha}\hat{+}\widehat{4}}},$$

and $(Z)^\alpha = ((U)_{\mathfrak{e}})_{\langle \mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}\widehat{\alpha}\hat{+}\widehat{5}, \langle\!\langle \theta_3[\dot{x}, \dot{\mathtt{c}}_\alpha, \dot{\mathtt{c}}_{F(\alpha)\restriction\alpha}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}\widehat{\alpha}\hat{+}\widehat{3}}], \overline{\sharp x}, f, g\rangle\!\rangle\rangle}$. By definition of F we have $\theta_6[\dot{\mathtt{c}}_\alpha, \dot{\mathtt{c}}_{F(\alpha)}][f,g]_\infty^{(U)_{\Omega_k}}$ for any α, hence

$$\forall\alpha(\theta_2[\dot{\mathtt{c}}_\alpha, \dot{\mathtt{c}}_\alpha, \dot{\mathtt{c}}_{\langle\!\langle \theta_3[\dot{x}, \dot{\mathtt{c}}_\alpha, \dot{\mathtt{c}}_{F(\alpha)\restriction\alpha}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}\widehat{\alpha}\hat{+}\widehat{3}}], \overline{\sharp x}, f, g\rangle\!\rangle}, \dot{\mathtt{c}}_{F(\alpha)(\alpha)}][f,g]_\infty^{(U)_{\Omega_k}}),$$

and by Lemma 105, because of $(Z)_\alpha = ((U)_{\mathfrak{e}})_{\langle \mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}(\widehat{\alpha}\hat{+}\widehat{1})\hat{+}\widehat{2}, F(\alpha)(\alpha)\rangle}$, and $(Z)^\alpha = ((U)_{\mathfrak{e}})_{\langle \mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}\widehat{\alpha}\hat{+}\widehat{5}, \langle\!\langle \theta_3[\dot{x}, \dot{\mathtt{c}}_\alpha, \dot{\mathtt{c}}_{F(\alpha)\restriction\alpha}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{\cdot}\widehat{\alpha}\hat{+}\widehat{3}}], \overline{\sharp x}, f, g\rangle\!\rangle\rangle}$, we get

$$\exists y \forall\alpha(\theta^*[\dot{\mathtt{c}}_\alpha, (\dot{\mathtt{C}}_{\langle \mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}, y\rangle})^{\dot{\mathtt{c}}_\alpha}, (\dot{\mathtt{C}}_{\langle \mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}, y\rangle})_{\dot{\mathtt{c}}_\alpha}][f,g]_\infty^{(U)_{\Omega_k}}),$$

hence we finally have $\psi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}$.

d) If $y = z \cup \{\psi_0 \dot{\vee} \psi_1\}$ and $(z \cup \{\psi_0\} \in (Z)_{\mathfrak{c},\emptyset} \vee z \cup \{\psi_1\} \in (Z)_{\mathfrak{c},\emptyset})$ for some $\mathfrak{c} \lhd \mathfrak{b}$ then by i.h. we have

$$(\exists\phi \in z \cup \{\psi_0\})\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}} \vee$$
$$(\exists\phi \in z \cup \{\psi_1\})\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}},$$

hence by logic we get

$$(\exists\phi \in z)\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}} \vee$$
$$(\psi_0^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}} \vee \psi_1^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}),$$

and by Lemma 88 and Lemma 108 we have

$$(\exists\phi \in z \cup \{\psi_0 \dot{\vee} \psi_1\})\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}.$$

e) If $y = z \cup \{\psi_0 \dot{\wedge} \psi_1\}$ and $(z \cup \{\psi_0\} \in (Z)_{\mathfrak{c},\emptyset} \wedge z \cup \{\psi_1\} \in (Z)_{\mathfrak{c},\emptyset})$ for some $\mathfrak{c} \lhd \mathfrak{b}$, then we proceed analogously to the previous case.

f) If $y = z \cup \{\dot{\exists}\, \dot{\mathtt{U}}_p(\psi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])\}$ and $z \cup \{\psi[\dot{\mathtt{V}}_t/\dot{\mathtt{V}}_q]\} \in (Z)_{\mathfrak{c},\emptyset}$ for some $\mathfrak{c} \lhd \mathfrak{b}$ then by i.h. and Lemma 108 we have

$$(\exists\phi \in z)\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}} \vee$$
$$(\psi^*[\dot{\mathtt{V}}_t^*/\dot{\mathtt{V}}_q^*])\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}},$$

that is

$$(\exists\phi \in z)\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}} \vee$$
$$(\psi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle)[\dot{\mathtt{C}}_{g(\overline{2}\cdot t)}/\dot{\mathtt{V}}_q^*][f,g]_\infty^{(U)_{\Omega_k}},$$

hence because of $g|\mathfrak{a}$ and Corollary 106 we get

$$(\exists\phi\in z)\phi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}\vee$$
$$\exists x(\psi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[(\dot{\mathtt{C}}_{\mathfrak{a}})_{\dot{\mathtt{c}}_x}/\dot{\mathtt{V}}_q^*][f,g]_\infty^{(U)\Omega_k}),$$

and finally $(\exists\phi\in y)\phi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}$ by Lemma 88.

g) If $y=z\cup\{\dot\forall\dot{\mathtt{U}}_p(\psi[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])\}$ and $(\forall\xi\in z)\dot{\mathtt{V}}_q\notin \mathit{term}(\xi)$ and $z\cup\{\psi\}\in(Z)_{\mathfrak{c},\emptyset}$ for some $\mathfrak{c}\lhd\mathfrak{b}$ then for any g_0 and x such that $g_0(\overline{2}\cdot q)=\langle\mathfrak{a},x\rangle$ and $g_0(r)=g(r)$ for $r\neq\overline{2}\cdot q$, i.e. $g_0|\mathfrak{a}$, we have

$$(\exists\phi\in z)\phi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}\vee$$
$$\psi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g_0]_\infty^{(U)\Omega_k}$$

by i.h. and Lemma 108, hence

$$(\exists\phi\in z)\phi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}\vee$$
$$\forall x(\psi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[\dot{\mathtt{C}}_{\langle\mathfrak{a},x\rangle}/\dot{\mathtt{V}}_q^*][f,g]_\infty^{(U)\Omega_k}),$$

and by Corollary 106 and Lemma 88 we finally get

$$(\exists\phi\in y)\phi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}.$$

h) If $y\cup\{\psi\}\in(Z)_{\mathfrak{c},\emptyset}\wedge y\cup\{\neg\psi\}\in(Z)_{\mathfrak{c},\emptyset}$ for some $\mathfrak{c}\lhd\mathfrak{b}$ and some $\psi\in\mathcal{G}_{CA[\Pi^1_0]}\cup\mathcal{G}_{\overrightarrow{DC}_{\mathcal{O}n}[\Pi^1_0]}$ then $\psi=\dot\exists\dot{\mathtt{U}}_p(\theta[\dot{\mathtt{U}}_p/\dot{\mathtt{V}}_q])$ for some $\theta\in\mathcal{G}_{\Pi^1_0}$. We fix some g such that $g|\mathfrak{a}$, and we define $\mathfrak{a}_0:=\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{+}\widehat{2}$, hence we have $g|\mathfrak{a}_0$, and by i.h. we get

$$(\exists\phi\in y)\phi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}_0}\rangle[f,g]_\infty^{(U)\Omega_k}\vee$$
$$\psi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}_0}\rangle[f,g]_\infty^{(U)\Omega_k}$$

and

$$(\exists\phi\in y)\phi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}_0},\dot{\mathtt{C}}_{\mathfrak{a}_0\hat{+}\Omega^{\mathfrak{c}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}\vee$$
$$(\neg\psi)^*\langle\dot{\mathtt{C}}_{\mathfrak{a}_0},\dot{\mathtt{C}}_{\mathfrak{a}_0\hat{+}\Omega^{\mathfrak{c}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k},$$

hence by Lemma 108 we get

$$(\exists\phi\in y)\phi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}\vee$$
$$(\psi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}_0}\rangle[f,g]_\infty^{(U)\Omega_k}\wedge(\neg\psi)^*\langle\dot{\mathtt{C}}_{\mathfrak{a}_0},\dot{\mathtt{C}}_{\mathfrak{a}_0\hat{+}\Omega^{\mathfrak{c}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}),$$

and by Lemma 88 and Corollary 106 we have

$$(\exists\phi\in y)\phi^*\langle\dot{\mathtt{C}}_{\mathfrak{a}},\dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\widehat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}\vee$$
$$\exists x(\theta^*[\dot{\mathtt{C}}_{\langle\mathfrak{a}_0,x\rangle}/\dot{\mathtt{V}}_q^*][f,g]_\infty^{(U)\Omega_k})\wedge\forall x((\neg\theta)^*[\dot{\mathtt{C}}_{\langle\mathfrak{a}_0,x\rangle}/\dot{\mathtt{V}}_q^*][f,g]_\infty^{(U)\Omega_k}).$$

By Lemma 87 we get

$$\neg(\theta^*[\dot{\mathtt{C}}_{\langle \mathfrak{a}_0, x\rangle}/\dot{\mathtt{V}}_q^*][f,g]_\infty^{(U)\Omega_k}) \leftrightarrow ((\neg\theta)^*[\dot{\mathtt{C}}_{\langle \mathfrak{a}_0, x\rangle}/\dot{\mathtt{V}}_q^*][f,g]_\infty^{(U)\Omega_k}),$$

hence we finally have $(\exists\phi \in y)\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}$.

i) If $(\dot{\mathtt{c}}_u \mathrel{\dot{\in}} \dot{\mathtt{c}}_v) \in y \wedge u \in v$ then $(\dot{\mathtt{c}}_u \mathrel{\dot{\in}} \dot{\mathtt{c}}_v)^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}$ follows directly by Lemma 88.

j) If $\dot{\sim}(\dot{\mathtt{c}}_u \mathrel{\dot{\in}} \dot{\mathtt{c}}_v) \in y \wedge u \notin v$ then $\dot{\sim}(\dot{\mathtt{c}}_u \mathrel{\dot{\in}} \dot{\mathtt{c}}_v)^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}$ follows directly by Lemma 88.

k) If $y = z \cup \{\dot{\exists}\,\dot{\mathtt{u}}_p(\psi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\}$ and $\exists w(z \cup \{\psi[\dot{\mathtt{c}}_w/\dot{\mathtt{v}}_q]\} \in (Z)_{\mathfrak{c},\emptyset})$ for some $\mathfrak{c} \lhd \mathfrak{b}$ then by i.h. and Lemma 108 we have

$$(\exists\phi \in z)\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)\Omega_k} \vee$$
$$\exists w(\psi[\dot{\mathtt{c}}_w/\dot{\mathtt{v}}_q]^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g_0]_\infty^{(U)\Omega_k}),$$

hence $(\exists\phi \in y)\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}$ by Lemma 88.

l) If $y = z \cup \{\dot{\forall}\,\dot{\mathtt{u}}_p(\psi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\}$ and $\forall w(\exists\mathfrak{c} \lhd \mathfrak{b})z \cup \{\psi[\dot{\mathtt{c}}_w/\dot{\mathtt{v}}_q]\} \in (Z)_{\mathfrak{c},\emptyset}$ then by i.h. we have

$$\forall w(\exists\mathfrak{c} \lhd \mathfrak{b})((\exists\phi \in z)\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)\Omega_k} \vee$$
$$\psi[\dot{\mathtt{c}}_w/\dot{\mathtt{v}}_q]^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{c}}\hat{+}\hat{2}}\rangle[f,g_0]_\infty^{(U)\Omega_k}),$$

and by Lemma 108 we get

$$(\exists\phi \in z)\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)\Omega_k} \vee$$
$$\forall w(\psi[\dot{\mathtt{c}}_w/\dot{\mathtt{v}}_q]^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g_0]_\infty^{(U)\Omega_k}),$$

hence $(\exists\phi \in y)\phi^*\langle \dot{\mathtt{C}}_{\mathfrak{a}}, \dot{\mathtt{C}}_{\mathfrak{a}\hat{+}\Omega^{\mathfrak{b}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)\Omega_k}$ by Lemma 88.

m) If $y = z \cup \{\dot{\exists}\,\dot{\mathtt{u}}_p(\psi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\}$ and $z \cup \{\psi[\dot{\mathtt{v}}_t/\dot{\mathtt{v}}_q]\} \in (Z)_{\mathfrak{c},\emptyset}$ for some $\mathfrak{c} \lhd \mathfrak{b}$ then we proceed analogous to Case f.

n) If $y = z \cup \{\dot{\forall}\,\dot{\mathtt{u}}_p(\psi[\dot{\mathtt{u}}_p/\dot{\mathtt{v}}_q])\}$ and $(\forall\xi \in z)\dot{\mathtt{v}}_q \notin \mathit{term}(\xi)$ and $z \cup \{\psi\} \in (Z)_{\mathfrak{b},\emptyset}$ for some $\mathfrak{c} \lhd \mathfrak{b}$ then we proceed analogous to Case g.

$\square$

By putting together the pieces and by cancelling out all the irrelevant parts in the statements of the previous lemma (the parts were essential for the proof by induction), we easily get the desired reductions.

Lemma 110. (Reduction)

For $\Gamma \subseteq \Sigma^1(\Pi^1_0)$ we have that

(1) If $\mathcal{C} = AC[\Pi^1_0]^{\rightarrow} \cup CA[\Pi^1_0]$ and $CA[\Pi^1_0]\ [AC[\Pi^1_0]]\vdash_{\mathcal{C},0} \Gamma$ then

$$\mathsf{NBG} \vdash \Gamma.$$

(2) If $\mathcal{D} = DC_{\mathcal{O}n}[\Pi^1_0]^{\rightarrow} \cup CA[\Pi^1_0]$ and $CA[\Pi^1_0]\ [DC_{\mathcal{O}n}[\Pi^1_0]]\vdash^{n}_{\mathcal{D},0} \Gamma$ then

$$\mathsf{NBG} \cup \exists\, Hier^{\Omega^{n+1}}_{\lhd}[\Pi^1_0] \vdash \Gamma.$$

(3) If $CA[\Pi^1_0] \cup TI_{\in}[\mathcal{L}^1]\ [DC_{\mathcal{O}n}[\Pi^1_0]]\vdash_{n} \Gamma$ then

$$\mathsf{NBG} \cup \exists\, Hier^{\Omega_{n+4}}_{\lhd}[\Pi^1_0] \vdash \Gamma.$$

Proof.

(1) If $CA[\Pi^1_0]\ [AC[\Pi^1_0]]\vdash^{n,i}_{\mathcal{C},0} \Gamma$, and $\{\mathtt{V}_j \mid \mathtt{V}_j \in \Gamma^{\vee *}\} \subseteq \{X_0, ..., X_l\}$, and $2^n < k$, then by Lemma 109 there is some m such that we have

$$\mathsf{NBG} \vdash Cl_k[\overline{m}, U, V, W] \wedge Gl[(W)_\emptyset] \wedge \overrightarrow{X} \mathrel{\mathring{\in}} (U)_\emptyset \rightarrow \Gamma^{\vee *}\langle \emptyset, \overline{2}^{\overline{n}} \rangle^{U},$$

and because of $(U)_\emptyset = W$ by Theorem 102, and by Lemma 107 we get

$$\mathsf{NBG} \vdash Cl_k[\overline{m}, U, V, W] \wedge Gl[(W)_\emptyset] \wedge \overrightarrow{X} \mathrel{\mathring{\in}} W \rightarrow \Gamma^{\vee *}.$$

We have $\mathsf{NBG} \vdash Gl[(W)_\emptyset] \wedge \overrightarrow{X} \mathrel{\mathring{\in}} W \rightarrow \Gamma^{\vee *}$ by Lemma 101, and further by Lemma 29 and comprehension we get $\mathsf{NBG} \vdash \exists Z(Gl[(Z)_\emptyset] \wedge \overrightarrow{X} \mathrel{\mathring{\in}} Z)$, hence $\mathsf{NBG} \vdash \Gamma^{\vee *}$, i.e. $\mathsf{NBG} \vdash \Gamma^{\vee}$, and finally $\mathsf{NBG} \vdash \Gamma$ because of $\mathsf{NBG} \vdash \neg(\Gamma^{\vee}), \Gamma$.

(2) Analogous to Part 1.

(3) If $CA[\Pi^1_0] \cup TI_{\in}[\mathcal{L}^1]\ [DC_{\mathcal{O}n}[\Pi^1_0]]\vdash_{n} \Gamma$ then by applying the disjunction rule we easily get $CA[\Pi^1_0] \cup TI_{\in}[\mathcal{L}^1]\ [DC_{\mathcal{O}n}[\Pi^1_0]]\vdash_{n} \Gamma^{\vee}$, and by Theorem 98 and Lemma 94 we have $\mathsf{NBG} \vdash Pr^{DC+}_{\Omega_{m+1}}[Z] \rightarrow \{\ulcorner \Gamma^{\vee} \urcorner\} \in (Z)_{\Omega \hat{+} \widehat{\omega}, \overline{n}}$, hence by Lemma 93, Theorem 95, and Lemma 92, we get

$$\mathsf{NBG} \cup \exists\, Hier^{\Omega_{m+1}}_{\lhd}[\Pi^1_0] \vdash \exists U(Pr^{DC+}_{\Omega_{m+1}}[U] \wedge \forall f(\{\ulcorner \Gamma^{\vee} \urcorner [f]_\infty\} \in (U)_{\omega_n^{\Omega \hat{+} \widehat{\omega}}, \emptyset}))$$

for $m > n$. For $k = n+3$ we have $\Omega^{\omega_n^{\Omega \hat{+} \widehat{\omega}}} \lhd \Omega_k$ because of

$$\Omega^{\omega_n^{\Omega \hat{+} \widehat{\omega}}} = \widehat{\omega}^{\Omega \hat{\cdot} \omega_n^{\Omega \hat{+} \widehat{\omega}}} \lhd \widehat{\omega}^{\Omega \hat{\cdot} \omega_{n+2}^{\Omega \hat{+} \widehat{\omega}}} = \widehat{\omega}^{\omega_{n+2}^{\Omega \hat{+} \widehat{\omega}}} \lhd \widehat{\omega}^{\Omega_{n+2}} = \Omega_{n+3},$$

and by Lemma 109 we get

$$\mathsf{NBG} \vdash \begin{array}{c} Cl_{\Omega_{k+1}}[U,V,W] \wedge Gl[(W)_\emptyset] \wedge \\ Pr^{DC+}_{\Omega_k}[Z] \wedge \{\ulcorner\Gamma^\vee\urcorner[h]_\infty\} \in (Z)_{\omega_n^{\Omega\hat{+}\hat{\omega}},\emptyset} \wedge g|\emptyset \rightarrow \\ \forall f((\ulcorner\Gamma^\vee\urcorner[h]_\infty)^*\langle\dot{\mathsf{C}}_\emptyset, \dot{\mathsf{C}}_{\Omega^{\omega_n^{\Omega\hat{+}\hat{\omega}}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}), \end{array}$$

hence we have

$$\mathsf{NBG} \cup \exists\, Hier^{\Omega_k}_{\triangleleft}[\Pi^1_0] \vdash \begin{array}{c} Cl_{\Omega_{k+1}}[U,V,W] \wedge Gl[(W)_\emptyset] \wedge g|\emptyset \rightarrow \\ \forall h \forall f((\ulcorner\Gamma^\vee\urcorner[h]_\infty)^*\langle\dot{\mathsf{C}}_\emptyset, \dot{\mathsf{C}}_{\Omega^{\omega_n^{\Omega\hat{+}\hat{\omega}}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}), \end{array}$$

that is

$$\mathsf{NBG} \cup \exists\, Hier^{\Omega_k}_{\triangleleft}[\Pi^1_0] \vdash \begin{array}{c} Cl_{\Omega_{k+1}}[U,V,W] \wedge Gl[(W)_\emptyset] \wedge g|\emptyset \rightarrow \\ \forall f(\ulcorner\Gamma^{\vee *}\urcorner\langle\dot{\mathsf{C}}_\emptyset, \dot{\mathsf{C}}_{\Omega^{\omega_n^{\Omega\hat{+}\hat{\omega}}}\hat{+}\hat{2}}\rangle[f,g]_\infty^{(U)_{\Omega_k}}). \end{array}$$

And by Lemma 107 we get

$$\mathsf{NBG} \cup \exists\, Hier^{\Omega_k}_{\triangleleft}[\Pi^1_0] \vdash \begin{array}{c} Cl_{\Omega_{k+1}}[U,V,W] \wedge Gl[(W)_\emptyset] \wedge g|\emptyset \rightarrow \\ \forall f(\ulcorner\Gamma^{\vee *}\urcorner[f,g]_\infty^{(U)_{\Omega_k}}). \end{array}$$

For $\{\mathtt{V}_i \mid \mathtt{V}_i \in \Gamma^{\vee *}\} \subseteq \{X_0, ..., X_m\}$ and $\sharp_{\Gamma^{\vee *}}$ (as in Lemma 90) we have

$$\mathsf{NBG} \cup \exists\, Hier^{\Omega_k}_{\triangleleft}[\Pi^1_0] \vdash \begin{array}{c} Cl_{\Omega_{k+1}}[U,V,W] \wedge Gl[(W)_\emptyset] \wedge \overrightarrow{X} \mathring{\in} W \rightarrow \\ (g_0|\emptyset \wedge \sharp_{\Gamma^{\vee *}}[f_0, g_0, (U)_{\Omega_k}] \rightarrow \ulcorner\Gamma^{\vee *}\urcorner[f_0,g_0]_\infty^{(U)_{\Omega_k}}), \end{array}$$

hence by Lemma 90 we get

$$\mathsf{NBG} \cup \exists\, Hier^{\Omega_k}_{\triangleleft}[\Pi^1_0] \vdash \begin{array}{c} Cl_{\Omega_{k+1}}[U,V,W] \wedge Gl[(W)_\emptyset] \wedge \overrightarrow{X} \mathring{\in} W \rightarrow \\ (g_0|\emptyset \wedge \sharp_{\Gamma^{\vee *}}[f_0, g_0, (U)_{\Omega_k}] \rightarrow \Gamma^{\vee *}), \end{array}$$

If $X_i = (W)_z$ and $g_0(\bar{i}) = \langle\emptyset, z\rangle$ then $((U)_{\Omega_k})_{g_0(\bar{i})} = X_i$ by Theorem 102, that is

$$\mathsf{NBG} \vdash Cl_{\Omega_{k+1}}[U,V,W] \wedge \overrightarrow{X} \mathring{\in} W \rightarrow \exists f_0 \exists g_0 (g_0|\emptyset \wedge \sharp_{\Gamma^{\vee *}}[f_0, g_0, (U)_{\Omega_k}]),$$

and hence we have

$$\mathsf{NBG} \cup \exists\, Hier^{\Omega_k}_{\triangleleft}[\Pi^1_0] \vdash \quad Cl_{\Omega_{k+1}}[U,V,W] \wedge Gl[(W)_\emptyset] \wedge \overrightarrow{X} \mathring{\in} W \rightarrow \Gamma^{\vee *},$$

and we finally get $\mathsf{NBG} \cup \exists\, Hier^{\Omega_{n+4}}_{\triangleleft}[\Pi^1_0] \vdash \Gamma$ analogous to Part 1. $\square$

Finally, we are able to state the proof-theoretic equivalences.

Definition 111.

$$(CA[\Pi^1_0])_{<\Omega^\omega} := \bigcup_{k\in\mathbb{N}}(\exists\, Hier_{\lhd}^{\Omega^{\widehat{k}}}[\Pi^1_0]),$$
$$(CA[\Pi^1_0])_{<E_0} := \bigcup_{k\in\mathbb{N}}(\exists\, Hier_{\lhd}^{\Omega_k}[\Pi^1_0]).$$

We observe that $(CA[\Pi^1_0])_{<\Omega^\omega}$ and $(CA[\Pi^1_0])_{<E_0}$ in the proofs given below, can not be replaced by $\exists\, Hier_{\lhd}^{\Omega^{\widehat{\omega}}}[\Pi^1_0]$ and $\exists\, Hier_{\lhd}^{E_0}[\Pi^1_0]$, respectively.

Definition 112.

For $\mathcal{F}, \mathcal{T}_0, \mathcal{T}_1 \subseteq \mathcal{L}^1$, and if we have $\mathcal{T}_0 \vdash \Gamma$ iff $\mathcal{T}_1 \vdash \Gamma$ for all $\Gamma \subseteq \mathcal{F}$, then we write $\mathcal{T}_0 \overset{\mathcal{F}}{\equiv} \mathcal{T}_1$.

Theorem 113.

If $\mathcal{T} \subseteq \{\neg(\forall)A \mid A \in \Sigma^1(\Pi^1_0)\}$ and $\mathcal{F} = \Sigma^1(\Pi^1_0)$ then

(1) $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG}$,

(2) $\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<\Omega^\omega}$,

(3) $\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \cup Wf_{\in}^{\mathcal{V}}[\mathcal{L}^1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}$.

Proof.

(1) If $\Gamma \subseteq \mathcal{F}$ and $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \vdash \Gamma$ then by Corollary 45 we have $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Pi^1_0] \vdash \Gamma$, and by Lemma 14 there is $\Delta \subseteq \mathsf{ZFC}$, and there are $A_0, ..., A_n \in \Sigma^1(\Pi^1_0)$ (with pairwise disjoint free variables) such that

$$CA[\Pi^1_0]\ [AC[\Pi^1_0]] \vdash (\forall)A_0, ..., (\forall)A_n, \neg\Delta, \neg\forall F(B[F]), \neg\exists F(C[F]), \Gamma$$

where $\forall F(B[F]), \exists F(C[F]) \in \mathsf{NBG}$ are the axiom of replacement and the axiom of global choice, respectively. By Lemma 15 and Corollary 17 we get

$$CA[\Pi^1_0]\ [AC[\Pi^1_0]] \vdash_{\mathcal{C},0} A_0, ..., A_n, \neg\Delta, \neg\forall F(B[F]), \neg C[X], \Gamma$$

for $\mathcal{C} = AC[\Pi^1_0]^{\rightarrow} \cup CA[\Pi^1_0]$, hence by Lemma 110 we have

$$\mathsf{NBG} \vdash A_0, ..., A_n, \neg\Delta, \neg\forall F(B[F]), \neg C[X], \Gamma,$$

and finally $\mathcal{T} \cup \mathsf{NBG} \vdash \Gamma$. The other direction is trivial.

(2) If $\Gamma \subseteq \mathcal{F}$ and $\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \vdash \Gamma$ then analogously to Part 1 we get $\mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<\Omega^\omega} \vdash \Gamma$, and there is $\Delta \subseteq \{(\forall)A \mid A \in (CA[\Pi^1_0])_{<\Omega^\omega}\}$ such that $\mathcal{T} \cup \mathsf{NBG} \vdash \Gamma, \neg\Delta$ by Lemma 14. For any $B \in \Delta$ we further have that $\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \vdash B$ by Theorem 79, hence $\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \vdash \Gamma$.

(3) We have $\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \cup TI_{\in}[\mathcal{L}^1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \cup Wf_{\in}^{\mathcal{V}}[\mathcal{L}^1]$ by Lemma 50. The claim follows analogous to Part 1 and 2. $\square$

Corollary 114.

If $\mathcal{T} \subseteq \{\neg(\forall)A \mid A \in \Sigma^1(\Pi^1_0)\}$ and $\mathcal{F} = \Sigma^1(\Pi^1_0)$ then

(1) $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \cup Wf^{\mathcal{V}}_{\in}[\mathcal{L}^1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}$,

(2) $\mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0} \cup Wf^{\mathcal{V}}_{\in}[\mathcal{L}^1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}$.

Proof.

(1) If $\Gamma \subseteq \mathcal{F}$ and $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \cup Wf^{\mathcal{V}}_{\in}[\mathcal{L}^1] \vdash \Gamma$ then by Corollary 45 we have $\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \cup Wf^{\mathcal{V}}_{\in}[\mathcal{L}^1] \vdash \Gamma$, and hence by Theorem 113 we get $\mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0} \vdash \Gamma$. If $\mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0} \vdash \Gamma$ then there is some $\Delta \subseteq \{(\forall)A \mid A \in (CA[\Pi^1_0])_{<E_0}\}$ such that we have $\mathcal{T} \cup \mathsf{NBG} \vdash \Gamma, \neg\Delta$ by Lemma 14. For any $B \in \Delta$ we have $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \cup Wf^{\mathcal{V}}_{\in}[\mathcal{L}^1] \vdash B$ by Theorem 79, hence we finally get $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \cup Wf^{\mathcal{V}}_{\in}[\mathcal{L}^1] \vdash \Gamma$.

(2) If $\Gamma \subseteq \mathcal{F}$ and $\mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0} \cup Wf^{\mathcal{V}}_{\in}[\mathcal{L}^1] \vdash \Gamma$ then analogous to Part 1 we get $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \cup Wf^{\mathcal{V}}_{\in}[\mathcal{L}^1] \vdash \Gamma$, hence by Part 1 we have $\mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0} \vdash \Gamma$. The other direction is trivial. □

Corollary 115.

If $\mathcal{T}$ and $\mathcal{F}$ are any sets of *sentences* in Σ^1_2 and Π^1_2, respectively, then

(1) $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG}$,

(2) $\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<\Omega^\omega}$,

(3) $\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \cup TI_{\in}[\mathcal{L}^1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}$,

(4) $\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \cup TI_{\in}[\mathcal{L}^1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}$,

(5) $\mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0} \cup TI_{\in}[\mathcal{L}^1] \overset{\mathcal{F}}{\equiv} \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}$.

Conclusion

In this thesis we have considered extensions of Von Neumann–Bernays–Gödel set theory NBG by axiom schemes, i.e., choice $AC[\Sigma^1_1]$, dependent choice $DC[\Sigma^1_1]$, full induction $TI_{\in}[\mathcal{L}^1]$, and iterated elementary comprehension $(CA[\Pi^1_0])_{<\mathsf{c}}$. We have established proof-theoretic equivalences between these schemes, similar to the results for analogous systems of arithmetic. The equivalences are

$$\begin{array}{rcl}
\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] & \equiv & \mathcal{T} \cup \mathsf{NBG}, \\
\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<\Omega^\omega}, \\
\mathcal{T} \cup \mathsf{NBG} \cup DC[\Sigma^1_1] \cup TI_{\in}[\mathcal{L}^1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}, \\
\mathcal{T} \cup \mathsf{NBG} \cup AC[\Sigma^1_1] \cup TI_{\in}[\mathcal{L}^1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0}, \\
\mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0} \cup TI_{\in}[\mathcal{L}^1] & \equiv & \mathcal{T} \cup \mathsf{NBG} \cup (CA[\Pi^1_0])_{<E_0},
\end{array}$$

where $\mathcal{T}$ is any set of axioms with logical complexity essentially Σ^1_2, and the equivalence, $\equiv$, means that any sentence essentially Π^1_2 is either provable in both theories or in none of them. The first equivalence has also been stated (without proof) by Feferman and Sieg [4]. The second last equivalence has been shown in a slightly weaker form by Jäger and Krähenbühl [10].

The main achievements of this thesis are the first three equivalences (the others are easy consequences). The bulk of work was in the reduction of the systems with choice to the systems with iterated comprehension. The reductions are achieved in a uniform way, by using cut-elimination and asymmetric interpretations (Section 10), see also Cantini [1], and Jäger and Krähenbühl [10], for similar reductions. The three different asymmetric interpretations are presented in one sweep to emphasize commonality and genericity of this technique. In contrast to Cantini [1], where the asymmetric interpretation is into hierarchies of sets of numbers, which are built by using fixed standard Π^1_n-complete predicates, in this thesis, the asymmetric interpretation is into cumulative hierarchies of classes (Section 9), which are built by using truth predicates for each level of the hierarchy, i.e., the truth predicates and the stages of the hierarchy are built simultaneously. Hierarchies of classes of this sort are also used in Jäger and Krähenbühl [10].

We can easily come up with further interesting questions in the proof theory of NBG by just translating results from second order arithmetic, and by asking whether these results also hold in the context of NBG, or whether even the same proof-theoretic techniques from arithmetic can be adapted to set theory, e.g., as it is the case for this thesis. An immediate such question with respect to the results of this thesis is whether the following equivalences also hold for $n > 0$, that is,

$$\begin{array}{rcl} \mathsf{NBG} \cup AC[\Sigma^1_{n+1}] & \equiv & \mathsf{NBG} \cup CA[\Pi^1_n], \\ \mathsf{NBG} \cup DC[\Sigma^1_{n+1}] & \equiv & \mathsf{NBG} \cup (CA[\Pi^1_n])_{<\Omega^\omega}, \\ \mathsf{NBG} \cup DC[\Sigma^1_{n+1}] \cup TI_\in[\mathcal{L}^1] & \equiv & \mathsf{NBG} \cup (CA[\Pi^1_n])_{<E_0}, \\ \mathsf{NBG} \cup AC[\Sigma^1_{n+1}] \cup TI_\in[\mathcal{L}^1] & \equiv & \mathsf{NBG} \cup (CA[\Pi^1_n])_{<E_0}. \end{array}$$

It was shown by Friedman [6], Feferman and Sieg [4], and Cantini [1], that analogous statements hold in the context of arithmetic. A first approach to prove these equivalences for NBG could follow the proofs of Cantini [1], again by using asymmetric interpretations, but now with some kind of *constructibility hypothesis* for classes involved (see also Feferman and Sieg [4] for a similar argument).

Another line of questions arises by considering extensions of NBG by restricted forms of choice, AC_ω and DC_ω, i.e.,

$$(\forall \alpha \in \omega)\exists Y A[\alpha, Y] \rightarrow \exists Y (\forall \alpha \in \omega) A[\alpha, (Y)_\alpha],$$

$$(\forall \alpha \in \omega)\forall Y \exists Z A[\alpha, Y, Z] \rightarrow \exists Z (\forall \alpha \in \omega) A[\alpha, (Z)^\alpha, (Z)_\alpha],$$

for formulas A. By Theorem 113 and 40 we already know that the first of the following equivalences holds. And it would be interesting to know whether it is possible to find c_0, c_1, c_2, such that

$$\begin{array}{rcl} \mathsf{NBG} \cup AC_\omega[\Sigma^1_1] & \equiv & \mathsf{NBG}, \\ \mathsf{NBG} \cup DC_\omega[\Sigma^1_1] & \equiv & \mathsf{NBG} \cup (CA[\Pi^1_0])_{<\mathsf{c}_0}, \\ \mathsf{NBG} \cup DC_\omega[\Sigma^1_1] \cup TI^\omega_\in[\mathcal{L}^1] & \equiv & \mathsf{NBG} \cup (CA[\Pi^1_0])_{<\mathsf{c}_1}, \\ \mathsf{NBG} \cup AC_\omega[\Sigma^1_1] \cup TI^\omega_\in[\mathcal{L}^1] & \equiv & \mathsf{NBG} \cup (CA[\Pi^1_0])_{<\mathsf{c}_2}. \end{array}$$

E.g., by adapting the proof of Theorem 79 we can get some possible lower bounds for c_1, c_2.

We have seen that $(E_0, \lhd)$, which corresponds to the wellordering $(\varepsilon_0, \in)$, plays an important role in the characterisation of the choice principles with full induction over NBG. By using the notation system in Appendix C, and an analogous construction as for E_0 in Section 6, we can easily build a linear ordering $(G_0, \lhd)$ corresponding to $(\Gamma_0, \in)$, in the same way as $(E_0, \lhd)$ corresponds to $(\varepsilon_0, \in)$. The ordering $(G_0, \lhd)$ is expected to play an important role if we consider to get analogous results as in Feferman and Jäger [2], where choice principles in presence of the *Bar Rule* are characterised by *autonomously iterated comprehension* in second order arithmetic.

A. Normality and Extensionality

Originally, set theory deals with collections of sets, where sets are themselves collections of sets, and if two sets or classes contain the same sets then the two are equal, i.e. sets and classes are extensional. We show that the notions of normality and extensionality for $\mathcal{L}^1$-structures essentially capture these properties, at least as good as it can be captured in the logic defined in Section 1, that is, modulo isomorphisms between $\mathcal{L}^1$-structures.

Definition 116. (Normal Structures)

Let $\mathcal{M}$ be a $\mathcal{L}^1$-structure.

(1) We define the *extensions* of $a \in |\mathcal{M}|$, $b \in \|\mathcal{M}\|$, such that

$$\begin{array}{rcl} E^0_{\mathcal{M}}(a) & := & \{c \mid c \in^0_{\mathcal{M}} a\}, \\ E^1_{\mathcal{M}}(b) & := & \{c \mid c \in^1_{\mathcal{M}} b\}. \end{array}$$

(2) We define $=^0_{\mathcal{M}} \subseteq |\mathcal{M}| \times |\mathcal{M}|$ and $=^1_{\mathcal{M}} \subseteq \|\mathcal{M}\| \times \|\mathcal{M}\|$, such that

$$\begin{array}{rcl} a =^0_{\mathcal{M}} b & :\Leftrightarrow & E^0_{\mathcal{M}}(a) = E^0_{\mathcal{M}}(b), \\ a =^1_{\mathcal{M}} b & :\Leftrightarrow & E^1_{\mathcal{M}}(a) = E^1_{\mathcal{M}}(b). \end{array}$$

(3) $\mathcal{M}$ is called *normal*, cf. Mendelson [14], if we have for all a, b that

$$a =^0_{\mathcal{M}} b \text{ or } a =^1_{\mathcal{M}} b \;\Rightarrow\; a = b.$$

(4) $\mathcal{M}$ is called *natural*, if we have for all $a \in |\mathcal{M}|$, $b \in \|\mathcal{M}\|$, that

$$a = E^0_{\mathcal{M}}(a) \text{ and } b = E^1_{\mathcal{M}}(b),$$

Clearly, all natural $\mathcal{L}^1$-structures are also normal structures, but generally not the other way around. The following theorem shows that at least we can find isomorphic natural structures for a whole class of normal structures.

Theorem 117.

If the $\mathcal{L}^1$-structure $\mathcal{M}$ is normal, and if we can recursively build the sets $E^\infty_{\mathcal{M}}(a) := \{E^\infty_{\mathcal{M}}(c) \mid c \in^0_{\mathcal{M}} a\}$ for any $a \in |\mathcal{M}|$, e.g. if $\in^0_{\mathcal{M}}$ is well-founded, then there is a natural $\mathcal{L}^1$-structure $\mathcal{N}$ *isomorphic* to $\mathcal{M}$, i.e. such that $\in^0_{\mathcal{N}}$ and $\in^1_{\mathcal{N}}$ are the restrictions of $\in$ to $|\mathcal{N}| \times |\mathcal{N}|$ and $|\mathcal{N}| \times \|\mathcal{N}\|$, respectively.

Proof. The isomorphism is

$$a \mapsto E^{\infty}_{\mathcal{M}}(a), \text{ for } a \in |\mathcal{M}|,$$
$$b \mapsto \{E^{\infty}_{\mathcal{M}}(a) \mid a \in^1_{\mathcal{M}} b\}, \text{ for } b \in \|\mathcal{M}\|.$$ □

Definition 118. (Extensionality)

The $\mathcal{L}^1$-structure $\mathcal{M}$ is called *extensional* if it satisfies the two *extensionality axioms* (also called *equality axioms* or *Leibniz's law*), i.e. if the following holds

$$\mathcal{M} \models \forall x \forall y \forall z (x = y \wedge x \in z \to y \in z),$$
$$\mathcal{M} \models \forall Z \forall x \forall y (x = y \wedge x \in Z \to y \in Z).$$

Lemma 119.

For any $\mathcal{L}^1$-structure $\mathcal{M}$ we have

$$a =^0_{\mathcal{M}} b \;\Leftrightarrow\; \mathcal{M} \models (x = y)[a, b/x, y],$$
$$a =^1_{\mathcal{M}} b \;\Leftrightarrow\; \mathcal{M} \models (X = Y)[a, b/X, Y].$$

Proof. By the definition of $=^0_{\mathcal{M}}$, $=^1_{\mathcal{M}}$, $x = y$, $X = Y$, and $\models$. □

Lemma 120.

For any $\mathcal{L}^1$-structure $\mathcal{M}$ we have

$$\mathcal{M} \models \forall x \forall y \forall z (x = y \wedge z \in x \to z \in y),$$
$$\mathcal{M} \models \forall X \forall Y \forall z (X = Y \wedge z \in X \to z \in Y).$$

Proof. By the definition of $x = y$, $X = Y$, and $\models$. □

Lemma 121.

For any $\mathcal{L}^1$-structure $\mathcal{M}$ we have

(1) $=^1_{\mathcal{M}}$ is a congruence relation for $\in^1_{\mathcal{M}}$, i.e. for all a, b, c we have

$$a =^1_{\mathcal{M}} b \wedge c \in^1_{\mathcal{M}} a \;\Rightarrow\; c \in^1_{\mathcal{M}} b.$$

(2) $\mathcal{M}$ is extensional iff $=^0_{\mathcal{M}}$ is a congruence relation for $\in^0_{\mathcal{M}}$ and $\in^1_{\mathcal{M}}$, i.e. iff for all a, b, c we have that

$$a =^0_{\mathcal{M}} b \wedge a \in^1_{\mathcal{M}} c \;\Rightarrow\; b \in^1_{\mathcal{M}} c,$$
$$a =^0_{\mathcal{M}} b \wedge a \in^0_{\mathcal{M}} c \;\Rightarrow\; b \in^0_{\mathcal{M}} c,$$
$$a =^0_{\mathcal{M}} b \wedge c \in^0_{\mathcal{M}} a \;\Rightarrow\; c \in^0_{\mathcal{M}} b.$$

Proof. By Lemma 119 and 120 and by the definition of $\models$. $\square$

Clearly, extensionality does not imply normality for $\mathcal{L}^1$-structures, but we get the following theorem about the relation between extensionality and normality.

Theorem 122.

> If $\mathcal{T} \subseteq \mathcal{L}^1$ is a set of *sentences* and $\mathcal{T}$ contains the *extensionality axioms*, then $\mathcal{T} \models A$ depends on normal $\mathcal{L}^1$-structures only, i.e. we have $\mathcal{T} \models A$ iff $(\mathcal{V} \models \mathcal{T} \Rightarrow \mathcal{V} \models A)$ for all valuations $\mathcal{V}$ with *normal* $\mathcal{L}^1$-structure $\langle \mathcal{V} \rangle$.

Proof. If $\mathcal{T} \not\models A$ then there is some $\mathcal{V}'$ such that $\mathcal{V}' \models \mathcal{T}$ and $\mathcal{V}' \not\models A$. The $\mathcal{L}^1$-structure $\mathcal{M} := \langle \mathcal{V}' \rangle$ is extensional, hence by Lemma 121 we can build the quotient structure $\mathcal{M}/{=}$ with respect to $=^0_{\mathcal{M}}$ and $=^1_{\mathcal{M}}$, such that $\mathcal{M}/{=}$ is *normal*, and for all $B \in \mathcal{L}^1$ we get that $\mathcal{M} \models B$ iff $\mathcal{M}/{=} \models B$. We have $\mathcal{M} \not\models A$, hence $\mathcal{M}/{=} \not\models A$, i.e. there is some $\mathcal{V} := (\mathcal{M}/{=}, f, g)$ such that $\mathcal{V} \not\models A$. We have $\mathcal{M} \models \mathcal{T}$, hence $\mathcal{M}/{=} \models \mathcal{T}$, and therefore $\mathcal{V} \models \mathcal{T}$. $\square$

B. Another Language for Theories of Sets and Classes

In the literature, there exist different formalizations for theories of sets and classes in classical logic. E.g. Mendelson [14] is using some common formulation of NBG set theory, with just one sort of variables, hence the language is different from the one we use in this thesis. In this section, we want to investigate the exact relationship between such formalizations, that is, in addition to the language $\mathcal{L}^1$ we define the language $\mathcal{L}^0_{\mathsf{C}}$, and in Theorem 132 and Theorem 139 we get some very general conditions, such that formulations of set theories in $\mathcal{L}^1$ and $\mathcal{L}^0_{\mathsf{C}}$ are equivalent, i.e. the theories prove the same theorems up to some fixed translation from one language to the other.

The language $\mathcal{L}^0_{\mathsf{C}}$ extends the language $\mathcal{L}^0$ of ZFC by two unary relation symbols S and C for sets and classes, respectively, and for technical reasons we also include a second binary relation symbol $\in^1$.

Definition 123. (Language $\mathcal{L}^0_{\mathsf{C}}$)

The language $\mathcal{L}^0_{\mathsf{C}}$ extends $\mathcal{L}^0$ and consists of the following

(1) The *logical symbols* of $\mathcal{L}^0_{\mathsf{C}}$ are the same as for $\mathcal{L}^0$.

(2) The *non logical symbols* of $\mathcal{L}^0_{\mathsf{C}}$ are the *element relation* symbols $\in^0$ and $\in^1$, and the unary relation symbols S and C.

(3) The *atomic* $\mathcal{L}^0_{\mathsf{C}}$*-formulas* are $\mathtt{v}_i{\in^0}\mathtt{v}_j$, $\sim\mathtt{v}_i{\in^0}\mathtt{v}_j$, $\mathtt{v}_i{\in^1}\mathtt{v}_j$, $\sim\mathtt{v}_i{\in^1}\mathtt{v}_j$, and $\mathsf{S}(\mathtt{v}_i)$, $\sim\mathsf{S}(\mathtt{v}_i)$, $\mathsf{C}(\mathtt{v}_i)$, $\sim\mathsf{C}(\mathtt{v}_i)$, for any $i, j \in \mathbb{N}$.

(4) The $\mathcal{L}^0_{\mathsf{C}}$*-formulas* are defined analogous to the $\mathcal{L}^0$-formulas.

We use notational conventions analogous to the language $\mathcal{L}^0$.

$\mathcal{L}^0_{\mathsf{C}}$ is a language of pure first order logic, hence its semantic and the adequate notion of formal proof are defined according to common first order logic.

Definition 124. (Semantics and Formal Proof)

(1) A $\mathcal{L}^0_{\mathsf{C}}$*-structure* is a tuple $\mathcal{M} := (|\mathcal{M}|, \mathsf{S}_{\mathcal{M}}, \mathsf{C}_{\mathcal{M}}, \in^0_{\mathcal{M}}, \in^1_{\mathcal{M}})$ with *non-empty* domain $|\mathcal{M}|$, and relations $\mathsf{S}_{\mathcal{M}}, \mathsf{C}_{\mathcal{M}} \subseteq |\mathcal{M}|$, and $\in^0_{\mathcal{M}}, \in^1_{\mathcal{M}} \subseteq |\mathcal{M}| \times |\mathcal{M}|$.

(2) A $\mathcal{L}^0_{\mathsf{C}}$*-valuation* is a tuple $\mathcal{V} := (\langle\mathcal{V}\rangle, f)$ such that $\langle\mathcal{V}\rangle$ is a $\mathcal{L}^0_{\mathsf{C}}$-structure and $f : \mathbb{N} \to |\langle\mathcal{V}\rangle|$.

(3) $\mathcal{V} \models A$ is defined analogous to Definition 9. For atomic $\mathcal{L}^0_{\mathsf{C}}$-formulas A we have $\mathcal{V} \models A$ iff $\mathcal{V} = (\mathcal{M}, f)$ is a $\mathcal{L}^0_{\mathsf{C}}$-valuation and one of the following holds:

$$\begin{array}{lll} A = \mathsf{S}(x) & \text{and} & f(\sharp x) \in \mathsf{S}_{\mathcal{M}}, \\ A = \mathsf{C}(x) & \text{and} & f(\sharp x) \in \mathsf{C}_{\mathcal{M}}, \\ A = x{\in^0}y & \text{and} & f(\sharp x){\in^0_{\mathcal{M}}}f(\sharp y), \\ A = x{\in^1}y & \text{and} & f(\sharp x){\in^1_{\mathcal{M}}}f(\sharp y), \\ A = {\sim}B & \text{and} & \mathcal{V} \not\models B. \end{array}$$

(4) All other semantic notations are analogous to Definition 9.

(5) $\mathcal{T}\ [\mathcal{R}]\!\!\mid\!\!\!\frac{n;j}{\mathcal{C},l}\ \Gamma$ for $\mathcal{T}, \mathcal{R}, \mathcal{C}, \Gamma \subseteq \mathcal{L}^0_{\mathsf{C}}$ is defined analogous to Definition 12.

For the translation of formulas we define two mappings; the first, $A \mapsto A^{\mathsf{c}}$, is a direct translation of $\mathcal{L}^1$ to the *proper* two sorted first order language $\mathcal{L}^0_{\mathsf{C}}$. The second mapping, $A \mapsto A^{\mathsf{s}}$, is such that we also get rid of the relation symbols C and $\in^0$, hence the formulas A^{s} have the form that is most commonly used for theories of sets and classes, i.e. by using just one relation symbol $\in$ and the predicate S for sets.

Definition 125. (Translations)

For formulas $A \in {\mathcal{L}^1}^*$ we define $A^{\mathsf{c}}, A^{\mathsf{s}} \in \mathcal{L}^0_{\mathsf{C}}$ such that

$$A^{\mathsf{c}} \quad := \quad \begin{cases} x{\in^0}y & A = x{\in^0}y, \\ x{\in^1}Y^+ & A = x{\in^1}Y, \\ {\sim}B^{\mathsf{c}} & A = {\sim}B, \\ (B^{\mathsf{c}} \wedge C^{\mathsf{c}}) & A = (B \wedge C), \\ (B^{\mathsf{c}} \vee C^{\mathsf{c}}) & A = (B \vee C), \\ \exists x(\mathsf{S}(x) \wedge B^{\mathsf{c}}[x]) & A = \exists x B[x], \\ \forall x(\mathsf{S}(x) \to B^{\mathsf{c}}[x]) & A = \forall x B[x], \\ \exists X^+(\mathsf{C}(X^+) \wedge B^{\mathsf{c}}[X^+/U^+]) & A = \exists X B[X/U], \\ \forall X^+(\mathsf{C}(X^+) \to B^{\mathsf{c}}[X^+/U^+]) & A = \forall X B[X/U], \end{cases}$$

$$A^{\mathsf{s}} \quad := \quad \begin{cases} x{\in^1}y & A = x{\in^0}y, \\ x{\in^1}Y^+ & A = x{\in^1}Y, \\ {\sim}B^{\mathsf{s}} & A = {\sim}B, \\ (B^{\mathsf{s}} \wedge C^{\mathsf{s}}) & A = (B \wedge C), \\ (B^{\mathsf{s}} \vee C^{\mathsf{s}}) & A = (B \vee C), \\ \exists x(\mathsf{S}(x) \wedge B^{\mathsf{s}}[x]) & A = \exists x B[x], \\ \forall x(\mathsf{S}(x) \to B^{\mathsf{s}}[x]) & A = \forall x B[x], \\ \exists X^+ B^{\mathsf{s}}[X^+/U^+] & A = \exists X B[X/U], \\ \forall X^+ B^{\mathsf{s}}[X^+/U^+] & A = \forall X B[X/U]. \end{cases}$$

We observe that $A^{\mathsf{s}} \in \mathcal{L}^0_{\mathsf{C}}$ contains only the relation symbols $\in^1$ and S.

For $\mathcal{T} \subseteq \mathcal{L}^{1^*}$ we define $\mathcal{T}^{\mathsf{c}} := \{A^{\mathsf{c}} \mid A \in \mathcal{T}\}$, and analogously $\mathcal{T}^{\mathsf{s}}$.

What we are actually looking for, are *minimal* sets of axioms $\mathcal{A} \subseteq \mathcal{L}^0_{\mathsf{C}}$, such that any set theory $\mathcal{T} \subseteq \mathcal{L}^1$ corresponds to the set theory $\mathcal{T}^{\mathsf{c}} \cup \mathcal{A}$ in $\mathcal{L}^0_{\mathsf{C}}$, or $\mathcal{T}^{\mathsf{s}} \cup \mathcal{A}$ in $\mathcal{L}^0_{\mathsf{C}}$ (in case of $\mathcal{T}^{\mathsf{s}}$ we will need to put some restriction on $\mathcal{T}$). As we show in the theorems below, the following definition is a first step in the right direction.

Definition 126. ($\mathcal{A}_{\mathsf{C}}, \mathcal{A}^+_{\mathsf{C}} \subseteq \mathcal{L}^0_{\mathsf{C}}$)

(1) $\mathcal{A}_{\mathsf{C}} \subseteq \mathcal{L}^0_{\mathsf{C}}$ consists of the formulas

$$\begin{array}{l} \exists x \mathsf{S}(x) \wedge \exists x \mathsf{C}(x), \\ \forall x(\mathsf{S}(x) \vee \mathsf{C}(x)), \\ \forall x \forall y(x \in^0 y \to \mathsf{S}(x) \wedge \mathsf{S}(y)), \\ \forall x \forall y(x \in^1 y \to \mathsf{S}(x) \wedge \mathsf{C}(y)). \end{array}$$

(2) $\mathcal{A}^+_{\mathsf{C}} \subseteq \mathcal{L}^0_{\mathsf{C}}$ extends $\mathcal{A}_{\mathsf{C}}$ by the formulas

$$\begin{array}{l} \forall x \mathsf{C}(x), \\ \forall x \forall y(\mathsf{S}(x) \wedge \mathsf{S}(y) \to (x \in^0 y \leftrightarrow x \in^1 y)). \end{array}$$

Almost any argument in this section is purely semantic, hence the following four lemmas and one further definition mainly consist of statements about $\mathcal{L}^0_{\mathsf{C}}$-structures and $\mathcal{L}^1$-structures. The following technical lemmas are essential steps towards the proof of Theorem 132. The next lemma is trivial, but it brings out the role of $\mathcal{A}_{\mathsf{C}}$ on $\mathcal{L}^0_{\mathsf{C}}$-structures.

Lemma 127.

If $\mathcal{N}$ is a $\mathcal{L}^0_{\mathsf{C}}$-structure then we have

$$\begin{array}{rcl} \mathcal{N} \models \exists x \mathsf{S}(x) \wedge \exists x \mathsf{C}(x) & \Leftrightarrow & \mathsf{S}_{\mathcal{N}} \neq \emptyset \wedge \mathsf{C}_{\mathcal{N}} \neq \emptyset, \\ \mathcal{N} \models \forall x(\mathsf{S}(x) \vee \mathsf{C}(x)) & \Leftrightarrow & |\mathcal{N}| = \mathsf{S}_{\mathcal{N}} \cup \mathsf{C}_{\mathcal{N}}, \\ \mathcal{N} \models \forall x \forall y(x \in^0 y \to \mathsf{S}(x) \wedge \mathsf{S}(y)) & \Leftrightarrow & \in^0_{\mathcal{N}} \subseteq \mathsf{S}_{\mathcal{N}} \times \mathsf{S}_{\mathcal{N}}, \\ \mathcal{N} \models \forall x \forall y(x \in^1 y \to \mathsf{S}(x) \wedge \mathsf{C}(y)) & \Leftrightarrow & \in^1_{\mathcal{N}} \subseteq \mathsf{S}_{\mathcal{N}} \times \mathsf{C}_{\mathcal{N}}. \end{array}$$

Proof. By the definition of $\models$. $\square$

Definition 128.

For $\mathcal{L}^1$-structures $\mathcal{M}$ and $\mathcal{L}^0_{\mathsf{C}}$-structures $\mathcal{N}$ we define

$$\begin{array}{rcl}
\mathcal{M}^{\mathsf{c}} & := & (|\mathcal{M}| \cup \|\mathcal{M}\|, |\mathcal{M}|, \|\mathcal{M}\|, \in^0_{\mathcal{M}}, \in^1_{\mathcal{M}}), \\
\mathcal{M}^{\mathsf{s}} & := & (|\mathcal{M}| \cup \|\mathcal{M}\|, |\mathcal{M}|, |\mathcal{M}| \cup \|\mathcal{M}\|, \in^0_{\mathcal{M}}, \in^0_{\mathcal{M}} \cup \in^1_{\mathcal{M}}), \\
\mathcal{N}^{\mathsf{c}} & := & (\mathsf{S}_{\mathcal{N}} \cup \mathsf{C}_{\mathcal{N}}, \mathsf{S}_{\mathcal{N}}, \mathsf{C}_{\mathcal{N}}, \in^0_{\mathcal{N}} \cap (\mathsf{S}_{\mathcal{N}} \times \mathsf{S}_{\mathcal{N}}), \in^1_{\mathcal{N}} \cap (\mathsf{S}_{\mathcal{N}} \times \mathsf{C}_{\mathcal{N}})), \\
\mathcal{N}^{\mathsf{s}} & := & (|\mathcal{N}|, \mathsf{S}_{\mathcal{N}}, |\mathcal{N}|, \in^1_{\mathcal{N}} \cap (\mathsf{S}_{\mathcal{N}} \times \mathsf{S}_{\mathcal{N}}), \in^1_{\mathcal{N}}), \\
\mathcal{N}^{\flat} & := & (\mathsf{S}_{\mathcal{N}} \times \{0\}, \mathsf{C}_{\mathcal{N}} \times \{1\}, \in^0_{\mathcal{N}^\flat}, \in^1_{\mathcal{N}^\flat}),
\end{array}$$

where $\in^0_{\mathcal{N}^\flat} := \{((a,0),(b,0)) \mid a \in^0_{\mathcal{N}} b\}$, $\in^1_{\mathcal{N}^\flat} := \{((a,0),(b,1)) \mid a \in^1_{\mathcal{N}} b\}$.

Lemma 129.

The mapping $\mathcal{M} \mapsto \mathcal{M}^{\mathsf{c}}$ is a bijection onto $\{\mathcal{N} \mid \mathcal{N} \models \mathcal{A}_{\mathsf{C}}\}$.

Proof. $\mathcal{M}^{\mathsf{c}} \models \mathcal{A}_{\mathsf{C}}$ by definition. If $\mathcal{N} \models \mathcal{A}_{\mathsf{C}}$ then $\mathcal{M} = (\mathsf{S}_{\mathcal{N}}, \mathsf{C}_{\mathcal{N}}, \in^0_{\mathcal{N}}, \in^1_{\mathcal{N}})$ is the $\mathcal{L}^1$-structure with $\mathcal{M}^{\mathsf{c}} = \mathcal{N}$. □

Lemma 130.

If $\mathcal{M}$ is a $\mathcal{L}^1$-structure and $A \in \mathcal{L}^{1^*}$ with all *free* variables in $\underline{x}, \underline{Y}$, and $\underline{a} \in |\mathcal{M}|$, $\underline{b} \in \|\mathcal{M}\|$ then we have

$$\mathcal{M} \models A[\underline{a}/\underline{x}][\underline{b}/\underline{Y}] \quad \Leftrightarrow \quad \mathcal{M}^{\mathsf{c}} \models A^{\mathsf{c}}[\underline{a}/\underline{x}][\underline{b}/\underline{Y^+}].$$

Proof. By induction on A. □

Lemma 131.

If $\mathcal{N}$ is a $\mathcal{L}^0_{\mathsf{C}}$-structure with $\mathcal{N} \models \exists x \mathsf{S}(x) \wedge \exists x \mathsf{C}(x)$ then we have

(1) $\mathcal{N}^{\mathsf{c}} \models \mathcal{A}_{\mathsf{C}}$.

(2) If $A \in \mathcal{L}^{1^*}$ with all *free* variables in $\underline{x}, \underline{Y}$, and $\underline{a} \in \mathsf{S}_{\mathcal{N}}$, $\underline{b} \in \mathsf{C}_{\mathcal{N}}$ then

$$\mathcal{N} \models A^{\mathsf{c}}[\underline{a}/\underline{x}][\underline{b}/\underline{Y^+}] \quad \Leftrightarrow \quad \mathcal{N}^{\mathsf{c}} \models A^{\mathsf{c}}[\underline{a}/\underline{x}][\underline{b}/\underline{Y^+}].$$

Proof.

(1) By the definition of $\mathcal{N}^{\mathsf{c}}$ and because of $\mathcal{N} \models \exists x \mathsf{S}(x) \wedge \exists x \mathsf{C}(x)$.

(2) By induction on A, using $\mathsf{S}_{\mathcal{N}} = \mathsf{S}_{\mathcal{N}^{\mathsf{c}}}$ and $\mathsf{C}_{\mathcal{N}} = \mathsf{C}_{\mathcal{N}^{\mathsf{c}}}$. □

The following theorem shows that the set of axioms $\{\exists x \mathsf{S}(x), \exists x \mathsf{C}(x)\}$, and the set $\mathcal{A}_{\mathsf{C}}$, qualify as additional axioms for the theory $\mathcal{T}^{\mathsf{c}}$, to make $\mathcal{T}^{\mathsf{c}}$ equivalent to $\mathcal{T}$ with respect to the mapping $A \mapsto A^{\mathsf{c}}$.

Theorem 132.

If $\mathcal{T} \subseteq \mathcal{L}^{1^*}$ is a set of *sentences* and $A \in \mathcal{L}^{1^*}$ is a *sentence* then the following are equivalent:

(1) $\mathcal{T} \models A$,

(2) $\mathcal{T}^{\mathsf{c}} \cup \mathcal{A}_{\mathsf{C}} \models A^{\mathsf{c}}$,

(3) $\mathcal{T}^{\mathsf{c}} \cup \{\exists x \mathsf{S}(x), \exists x \mathsf{C}(x)\} \models A^{\mathsf{c}}$.

Proof. (2)$\to$(1) If $\mathcal{M} \models \mathcal{T}$ then $\mathcal{M}^{\mathsf{c}} \models \mathcal{T}^{\mathsf{c}} \cup \mathcal{A}_{\mathsf{C}}$ by Lemmas 130 and 129, hence $\mathcal{M}^{\mathsf{c}} \models A^{\mathsf{c}}$ by (2), and $\mathcal{M} \models A$ by Lemma 130. (1)$\to$(2) If $\mathcal{N} \models \mathcal{T}^{\mathsf{c}} \cup \mathcal{A}_{\mathsf{C}}$ then there is $\mathcal{M}$ with $\mathcal{M}^{\mathsf{c}} = \mathcal{N}$ by Lemma 129, hence $\mathcal{M} \models \mathcal{T}$ by Lemma 130, and $\mathcal{M} \models A$ by (1), and $\mathcal{M}^{\mathsf{c}} \models A^{\mathsf{c}}$ by Lemma 130, that is $\mathcal{N} \models A^{\mathsf{c}}$. (3)$\to$(2) is trivial. (2)$\to$(3) If $\mathcal{N} \models \mathcal{T}^{\mathsf{c}} \cup \{\exists x \mathsf{S}(x), \exists x \mathsf{C}(x)\}$ then $\mathcal{N}^{\mathsf{c}} \models \mathcal{T}^{\mathsf{c}} \cup \mathcal{A}_{\mathsf{C}}$ by Lemma 131, and $\mathcal{N}^{\mathsf{c}} \models A^{\mathsf{c}}$ by (2), hence $\mathcal{N} \models A^{\mathsf{c}}$ by Lemma 131. $\square$

We use Theorem 132 to prove completeness of the logic defined in Section 1, that is, for theories $\mathcal{T} \subseteq \mathcal{L}^1$ we show $\mathcal{T} \models A \Rightarrow \mathcal{T} \vdash A$ by applying completeness of *pure* first order predicate logic. For this task, we need a transformation of proofs in $\mathcal{L}^0_{\mathsf{C}}$ to proofs in $\mathcal{L}^1$, and by the following lemma this can be done even recursively for proofs without cuts.

Lemma 133.

If $\Gamma \subseteq \mathcal{L}^{1^*}$, and $\mathsf{S}(x), \mathsf{C}(Y^+) \in \Delta$ for all free variables $x, Y \in \Gamma$, and if Δ contains only formulas of the form $\mathsf{S}(y), \mathsf{C}(y)$, then we have that

$$\vdash^{n}_{0} \neg\Delta, \Gamma^{\mathsf{c}} \quad \Rightarrow \quad \vdash^{n}_{0} \Gamma.$$

Proof. By induction on n, considering all cases in Definition 12. We use $(\sim A)^{\mathsf{c}} = \sim(A^{\mathsf{c}})$ and $\Delta \cap \Gamma^{\mathsf{c}} = \emptyset$ for the base case, and Lemma 15 ($\vee, \wedge$-inversion) in case of the quantifier rules. $\square$

Theorem 134. (Adequacy for $\mathcal{L}^1$)

If $\mathcal{T} \subseteq \mathcal{L}^1$ is a set of *sentences* then $\mathcal{T} \models A$ iff $\mathcal{T} \vdash A$.

Proof. For $\Gamma \subseteq \mathcal{L}^1$ we show $\mathcal{T} \vdash^{n} \Gamma \Rightarrow \mathcal{T} \models \Gamma^{\vee}$ by induction on n. For the other direction w.l.o.g. A is a *closed* formula. We assume $\mathcal{T} \models A$, hence $\mathcal{T}^* \models A^*$. By Theorem 132 we have $\mathcal{T}^{*\mathsf{c}} \cup \{\exists x \mathsf{S}(x), \exists x \mathsf{C}(x)\} \models A^{*\mathsf{c}}$, and by completeness of first order logic we get $\mathcal{T}^{*\mathsf{c}} \cup \{\exists x \mathsf{S}(x), \exists x \mathsf{C}(x)\} \vdash A^{*\mathsf{c}}$. By Lemma 14, and Lemma 15, and Corollary 17, there is some $\Gamma \subseteq \mathcal{T}^*$ such that $\vdash_0 \sim\mathsf{S}(u), \sim\mathsf{C}(v), (\neg\Gamma)^{\mathsf{c}}, A^{*\mathsf{c}}$, hence $\vdash \neg\Gamma, A^*$ by Lemma 133, i.e. $\mathcal{T}^* \vdash A^*$. By induction on n we get $\mathcal{T}^* \vdash^{n} \Delta^* \Rightarrow \mathcal{T} \vdash^{n} \Delta$ for $\Delta \subseteq \mathcal{L}^1$, hence $\mathcal{T} \vdash A$. $\square$

For the rest of this section we are heading towards an analogue of Theorem 132 for the translation $A \mapsto A^{\mathsf{s}}$, i.e. we show that there is a set of axioms $\mathcal{A} \subseteq \mathcal{L}^0_{\mathsf{C}}$, such that any set theory $\mathcal{T} \cup \{\forall x \exists Y(x = Y)\} \subseteq \mathcal{L}^1$ corresponds to the set theory $\mathcal{T}^{\mathsf{s}} \cup \mathcal{A}$ in $\mathcal{L}^0_{\mathsf{C}}$. We need the following four technical lemmas for the proof of this claim.

Lemma 135.

If $\mathcal{M}$ is a $\mathcal{L}^1$-structure with $\mathcal{M} \models \forall x \exists Y(x = Y)$ and $|\mathcal{M}| \cap \|\mathcal{M}\| = \emptyset$ then we have that

(1) There is a mapping $f : |\mathcal{M}| \to \|\mathcal{M}\|$ such that for all $a \in |\mathcal{M}|$

$$\mathcal{M} \models (x = Y)[a/x][f(a)/Y].$$

(2) For any f in (1) and $A \in \mathcal{L}^{1^*}$ with all *free* variables in $\underline{x}, \underline{Y}, Z$, and $\underline{a} \in |\mathcal{M}|$, $\underline{b} \in |\mathcal{M}| \cup \|\mathcal{M}\|$, $c \in |\mathcal{M}|$ we have

$$\mathcal{M}^{\mathsf{s}} \models A^{\mathsf{c}}[\underline{a}/\underline{x}][\underline{b}, c/\underline{Y^+}, Z^+] \;\Leftrightarrow\; \mathcal{M}^{\mathsf{s}} \models A^{\mathsf{c}}[\underline{a}/\underline{x}][\underline{b}, f(c)/\underline{Y^+}, Z^+].$$

(3) If $A \in \mathcal{L}^{1^*}$ with all *free* variables in $\underline{x}, \underline{Y}$, and $\underline{a} \in |\mathcal{M}|$, $\underline{b} \in \|\mathcal{M}\|$ then

$$\mathcal{M} \models A[\underline{a}/\underline{x}][\underline{b}/\underline{Y}] \quad \Leftrightarrow \quad \mathcal{M}^{\mathsf{s}} \models A^{\mathsf{c}}[\underline{a}/\underline{x}][\underline{b}/\underline{Y^+}].$$

Proof.

(1) By the definition of $\models$.

(2) By induction on A, using Part 1.

(3) By induction on A. $|\mathcal{M}| \cap \|\mathcal{M}\| = \emptyset$ is used in case $A = x \in^1 Y$, i.e. for $b \in \|\mathcal{M}\|$ we have $a \in^1_{\mathcal{M}} b \Leftrightarrow (a \in^0_{\mathcal{M}} b$ or $a \in^1_{\mathcal{M}} b)$. Part 2 is used in case $A = \exists X B[X]$ or $A = \forall X B[X]$, e.g. if $\mathcal{M}^{\mathsf{s}} \models A^{\mathsf{c}}[\underline{a}/\underline{x}][\underline{b}, c/\underline{Y^+}, Z^+]$ for all $c \in \|\mathcal{M}\|$ then this also holds for all $c \in |\mathcal{M}| \cup \|\mathcal{M}\|$ by Part 2. □

Lemma 136.

(1) If $\mathcal{M}$ is a $\mathcal{L}^1$-structure with $|\mathcal{M}| \cap \|\mathcal{M}\| = \emptyset$ then $\mathcal{M}^{\mathsf{s}} \models \mathcal{A}^+_{\mathsf{C}}$.

(2) If $\mathcal{N}$ is a $\mathcal{L}^0_{\mathsf{C}}$-structure with $\mathcal{N} \models \mathcal{A}^+_{\mathsf{C}}$, and $A \in \mathcal{L}^{1^*}$ with all *free* variables in $\overline{x}, \overline{Y}$, and $\underline{a} \in \mathsf{S}_{\mathcal{N}}$, $\underline{b} \in |\mathcal{N}|$ then

$$\mathcal{N} \models A^{\mathsf{c}}[\underline{a}/\underline{x}][\underline{b}/\underline{Y^+}] \quad \Leftrightarrow \quad \mathcal{N} \models A^{\mathsf{s}}[\underline{a}/\underline{x}][\underline{b}/\underline{Y^+}].$$

(3) For *sentences* $A \in \mathcal{L}^{1^*}$ we have $\mathcal{A}^+_{\mathsf{C}} \models A^{\mathsf{c}} \leftrightarrow A^{\mathsf{s}}$.

Proof.

(1) By the definition of $\mathcal{M}^{\mathsf{s}}$.

(2) By induction on A.

(3) By Part 2. $\square$

Lemma 137.

If $\mathcal{N}$ is a $\mathcal{L}^0_{\mathsf{C}}$-structure with $\mathcal{N} \models \mathcal{A}^+_{\mathsf{C}}$ then we have that

(1) $|\mathcal{N}^\flat| \cap \|\mathcal{N}^\flat\| = \emptyset$ and $\mathcal{N}^\flat \models \forall x \exists Y (x = Y)$.

(2) If $A \in \mathcal{L}^{1^*}$ with all *free* variables in $\underline{x}, \underline{Y}$, and $\underline{a} \in \mathsf{S}_{\mathcal{N}}$, $\underline{b} \in \mathsf{C}_{\mathcal{N}}$ then

$$\mathcal{N}^\flat \models A[\underline{(a,0)}/\underline{x}][\underline{(b,1)}/\underline{Y}] \quad \Leftrightarrow \quad \mathcal{N} \models A^{\mathsf{c}}[\underline{a}/\underline{x}][\underline{b}/\underline{Y^+}]$$

(3) If $A \in \mathcal{L}^{1^*}$ is a *sentence* then

$$(\mathcal{N}^\flat)^{\mathsf{s}} \models A^{\mathsf{c}} \quad \Leftrightarrow \quad \mathcal{N} \models A^{\mathsf{c}}.$$

Proof.

(1) By the definition of $\mathcal{N}^\flat$.

(2) By induction on A.

(3) By using Parts 1+2 and Lemma 135.(3). $\square$

Lemma 138.

If $\mathcal{N}$ is a $\mathcal{L}^0_{\mathsf{C}}$-structure with $\mathcal{N} \models \exists x \mathsf{S}(x) \wedge \forall x \forall y (x \in^1 y \to \mathsf{S}(x))$ then we have that

(1) $\mathcal{N}^{\mathsf{s}} \models \mathcal{A}^+_{\mathsf{C}}$.

(2) If $A \in \mathcal{L}^{1^*}$ with all *free* variables in $\underline{x}, \underline{Y}$, and $\underline{a} \in \mathsf{S}_{\mathcal{N}}$, $\underline{b} \in |\mathcal{N}|$ then

$$\mathcal{N} \models A^{\mathsf{s}}[\underline{a}/\underline{x}][\underline{b}/\underline{Y^+}] \quad \Leftrightarrow \quad \mathcal{N}^{\mathsf{s}} \models A^{\mathsf{s}}[\underline{a}/\underline{x}][\underline{b}/\underline{Y^+}].$$

Proof.

(1) By the definition of $\mathcal{N}^{\mathsf{s}}$ and because $\mathcal{N} \models \exists x \mathsf{S}(x) \wedge \forall x \forall y (x \in^1 y \to \mathsf{S}(x))$.

(2) By induction on A, using that S and $\in^0$ do not occur in A^{s}. $\square$

Theorem 139.

If $\mathcal{T} \subseteq \mathcal{L}^{1^*}$ is a set of *sentences* and $A \in \mathcal{L}^{1^*}$ is a *sentence* then the following are equivalent:

(1) $\mathcal{T} \cup \{\forall x \exists Y(x = Y)\} \models A$,

(2) $\mathcal{T}^{\mathsf{c}} \cup \mathcal{A}^+_{\mathsf{C}} \models A^{\mathsf{c}}$,

(3) $\mathcal{T}^{\mathsf{s}} \cup \{\exists x \mathsf{S}(x), \forall x \forall y(x \in^1 y \to \mathsf{S}(x))\} \models A^{\mathsf{s}}$.

Proof. (1)→(2) If $\mathcal{N} \models \mathcal{T}^{\mathsf{c}} \cup \mathcal{A}^+_{\mathsf{C}}$ then $\mathcal{N}^\flat \models \mathcal{T} \cup \{\forall x \exists Y(x = Y)\}$ by Lemma 137.(1+2), hence $\mathcal{N}^\flat \models A$ by (1), and $(\mathcal{N}^\flat)^{\mathsf{s}} \models A^{\mathsf{c}}$ by Lemma 137.(1) and 135.(3), and $\mathcal{N} \models A^{\mathsf{c}}$ by Lemma 137.(3). (2)→(1) To show $\mathcal{C} \models \mathcal{D}$ for sets of sentences $\mathcal{C}, \mathcal{D} \subseteq \mathcal{L}^1$, we may consider $\mathcal{L}^1$-structures $\mathcal{M}$ with $|\mathcal{M}| \cap \|\mathcal{M}\| = \emptyset$ only. If $\mathcal{M} \models \mathcal{T} \cup \{\forall x \exists Y(x = Y)\}$ and $|\mathcal{M}| \cap \|\mathcal{M}\| = \emptyset$ then $\mathcal{M}^{\mathsf{s}} \models \mathcal{T}^{\mathsf{c}} \cup \mathcal{A}^+_{\mathsf{C}}$ by Lemma 135.(3) and 136.(1), and $\mathcal{M}^{\mathsf{s}} \models A^{\mathsf{c}}$ by (2), hence $\mathcal{M} \models A$ by Lemma 135.(3). (3)→(2) If $\mathcal{N} \models \mathcal{T}^{\mathsf{c}} \cup \mathcal{A}^+_{\mathsf{C}}$ then $\mathcal{N} \models \mathcal{T}^{\mathsf{s}}$ by Lemma 136.(2), and $\mathcal{N} \models A^{\mathsf{s}}$ by (3), hence $\mathcal{N} \models A^{\mathsf{c}}$ by Lemma 136.(2). (2)→(3) If $\mathcal{N} \models \mathcal{T}^{\mathsf{s}} \cup \{\exists x \mathsf{S}(x), \forall x \forall y(x \in^1 y \to \mathsf{S}(x))\}$ then $\mathcal{N}^{\mathsf{s}} \models \mathcal{A}^+_{\mathsf{C}}$ and $\mathcal{N}^{\mathsf{s}} \models \mathcal{T}^{\mathsf{s}}$ by Lemma 138, hence $\mathcal{N}^{\mathsf{s}} \models \mathcal{T}^{\mathsf{c}}$ by Lemma 136.(2), and $\mathcal{N}^{\mathsf{s}} \models A^{\mathsf{c}}$ by (2), that is $\mathcal{N}^{\mathsf{s}} \models A^{\mathsf{s}}$ by Lemma 136.(2), and finally $\mathcal{N} \models A^{\mathsf{s}}$ by Lemma 138. $\square$

C. Another Notation System for Ordinals

This section is in complete analogy to Section 5. Instead of building the notation system on the Cantor normal form of ordinals, we now use the binary Veblen function and the Veblen normal form, i.e. we use that each ordinal $\alpha \neq \emptyset$ can be uniquely represented in the form $\alpha = \varphi_{\beta_1}(\gamma_1) + ... + \varphi_{\beta_p}(\gamma_p)$ with $\varphi_{\beta_q}(\gamma_q) \leq \varphi_{\beta_r}(\gamma_r)$ for $r < q$, and $\gamma_q < \varphi_{\beta_q}(\gamma_q)$ for $q \leq p$. The resulting notation system is similar to the standard notation system for the ordinal Γ_0, see e.g. Pohlers [15]. The notation $\widetilde{\alpha}$ for the ordinal α is defined recursively

$$\widetilde{\alpha} \quad := \quad \begin{cases} \alpha & \alpha = \varphi_\alpha(\emptyset) \text{ or } \alpha = \emptyset, \\ \langle\!\langle \langle\!\langle \widetilde{\beta_1}, \widetilde{\gamma_1} \rangle\!\rangle, ..., \langle\!\langle \widetilde{\beta_p}, \widetilde{\gamma_p} \rangle\!\rangle \rangle\!\rangle & \alpha =_{\mathsf{VNF}} \varphi_{\beta_1}(\gamma_1) + ... + \varphi_{\beta_p}(\gamma_p) \neq \varphi_\alpha(\emptyset). \end{cases}$$

Once again, we are going to define this notation system in a generic way without refering to ordinals, hence we can easily get notation systems going beyond the ordinals analogously to Section 6.

Definition 140. (Binary Veblen Function)

We define the expression $\varphi_\alpha(\beta)$ (analogous Definition 18) such that

$$\begin{aligned} z \in \varphi_\alpha(\beta) \quad &:= \quad \exists f(\mathit{Veb}[f] \wedge \alpha \in \mathit{dom}(f) \wedge \\ &\qquad\qquad \beta \in \mathit{dom}(f(\alpha)) \wedge z \in f(\alpha)(\beta)), \end{aligned}$$

$$\begin{aligned} \mathit{Veb}[f] \quad := \quad & \mathit{Fun}[f] \wedge \exists\alpha\exists\beta(\mathit{dom}(f) = \alpha \wedge (\forall\alpha_1 \in \alpha)(\forall\beta_1 \in \beta)(\\ & \mathit{Fun}[f(\alpha_1)] \wedge \mathit{dom}(f(\alpha_1)) = \beta \wedge \mathit{ran}(f(\alpha_1)) \subseteq \beta \wedge \\ & (\forall\beta_0 \in \beta_1) f(\alpha_1)(\beta_0) \in f(\alpha_1)(\beta_1) \wedge \\ & (((\alpha_1 = \emptyset \wedge \beta_1 \neq \emptyset \wedge (\forall\gamma_0 \in \beta_1)(\forall\gamma_1 \in \beta_1)\gamma_0 + \gamma_1 \in \beta_1) \vee \\ & (\alpha_1 \neq \emptyset \wedge (\forall\alpha_0 \in \alpha_1) f(\alpha_0)(\beta_1) = \beta_1)) \leftrightarrow \\ & (\exists\beta_0 \in \beta) f(\alpha_1)(\beta_0) = \beta_1))). \end{aligned}$$

Definition 141. (Veblen Normal Form)

$$\begin{aligned} \mathit{VNF}[f, p, \alpha] \quad := \quad & \mathit{Fun}[f] \wedge \mathit{dom}(f) = p \wedge \exists h(\forall p_1 \in p)(\\ & \exists\beta\exists\gamma(f(p_1) = \langle\!\langle \beta, \gamma \rangle\!\rangle \wedge h(p_1) = \varphi_\beta(\gamma) \wedge \gamma \in h(p_1)) \wedge \\ & (\forall p_0 \in p_1) h(p_0) \in h(p_1)' \wedge \alpha = \Sigma_p h). \end{aligned}$$

Theorem 142. (Veblen Normal Form)

(1) $\mathsf{NBG} \vdash \forall\alpha(\alpha = \emptyset \vee \exists! f \exists p\, VNF[f, p, \alpha])$,

(2) $\mathsf{NBG} \vdash \forall\alpha\forall f\forall p(VNF[f, p, \alpha] \rightarrow \alpha = \emptyset \vee \alpha = \varphi_\alpha(\emptyset) \vee f(\emptyset)(\emptyset) \in \alpha)$.

Proof. See e.g. Pohlers [15]. □

Definition 143. (Ordinal Notation System)

We define the expression $\widetilde{\alpha}$ (analogous Definition 18) such that

$$z \in \widetilde{\alpha} \quad := \quad \exists f(OT_\Gamma[f] \wedge \alpha \in dom(f) \wedge z \in f(\alpha)),$$

$$\begin{aligned} OT_\Gamma[f] \quad := \quad & Fun[f] \wedge \exists\alpha(dom(f) = \alpha \wedge (\forall\alpha_0 \in \alpha)(\\ & (\alpha_0 = \emptyset \wedge f(\alpha_0) = \alpha_0) \vee (\alpha_0 = \varphi_{\alpha_0}(\emptyset) \wedge f(\alpha_0) = \alpha_0) \vee \\ & (\alpha_0 \neq \emptyset \wedge \alpha_0 \neq \varphi_{\alpha_0}(\emptyset) \wedge Fun[f(\alpha_0)] \wedge \\ & \exists g\exists p(VNF[g, p, \alpha_0] \wedge dom(f(\alpha_0)) = p \wedge \\ & (\forall p_0 \in p) f(\alpha_0)(p_0) = \langle\!\langle f(g(p_0)(\emptyset)), f(g(p_0)(\overline{1}))\rangle\!\rangle))). \end{aligned}$$

Lemma 144. (Ordinal Notation System)

$$\begin{aligned} & \mathsf{NBG} \vdash \forall\alpha\forall f\forall p(\alpha \neq \emptyset \wedge \alpha \neq \varphi_\alpha(\emptyset) \wedge VNF[f, p, \alpha] \rightarrow \\ & \qquad Fun[\widetilde{\alpha}] \wedge dom(\widetilde{\alpha}) = p \wedge (\forall p_1 \in p)\widetilde{\alpha}(p_1) = \langle\!\langle \widetilde{f(p_1)(\emptyset)}, \widetilde{f(p_1)(\overline{1})}\rangle\!\rangle). \end{aligned}$$

Proof. By Definition (i.e. by induction on the ordinals). □

Definition 145. (Ordering Relation)

We define the expressions y^φ and $\widetilde{Y}_X$ (analogous Definition 18) such that

$$\begin{aligned} z \in y^\varphi \quad := \quad & \exists f\exists g(z = \langle f, g\rangle \wedge (\\ & (\langle f(\emptyset), g(\emptyset)\rangle \in y \wedge \langle f(\overline{1}), \langle\!\langle g\rangle\!\rangle\rangle \in y) \vee \\ & (f(\emptyset) = g(\emptyset) \wedge \langle f(\overline{1}), g(\overline{1})\rangle \in y) \vee \\ & (\langle g(\emptyset), f(\emptyset)\rangle \in y \wedge \langle \langle\!\langle f\rangle\!\rangle, g(\overline{1})\rangle \in y))), \\ z \in \widetilde{Y}_X \quad := \quad & \exists y(Ex^\varphi[X, Y, y] \wedge z \in y), \end{aligned}$$

$$\begin{aligned} Ex^\varphi[X, Y, y] \quad := \quad & (\forall x \in y)\exists f\exists g(x = \langle f, g\rangle \wedge \{f, g\} \subseteq \mathcal{H}_X \cup X \wedge \\ & (f \in X \wedge g \in X \wedge \langle f, g\rangle \in Y) \vee \\ & (f \in X \wedge g \notin X \wedge g \neq \emptyset \wedge \\ & \qquad (\langle \langle\!\langle f, \emptyset\rangle\!\rangle, g(\emptyset)\rangle \in y^\varphi \vee \langle\!\langle f, \emptyset\rangle\!\rangle = g(\emptyset))) \vee \\ & (f \notin X \wedge g \in X \wedge (f = \emptyset \vee \langle f(\emptyset), \langle\!\langle g, \emptyset\rangle\!\rangle\rangle \in y^\varphi) \vee \\ & (f \notin X \wedge g \notin X \wedge Lex[y^\varphi, f, g])). \end{aligned}$$

See Definition 60 for $Lex[y^\varphi, f, g]$.

Definition 146. (Generic Notation System)

We define the expressions $\widetilde{\mathcal{O}t}^0_{X,Y}$, $\widetilde{\mathcal{O}t}^1_{X,Y}$ $\widetilde{\mathcal{O}t}_{X,Y}$, η_X (analogous Definition 18) such that

$$
\begin{array}{rcl}
z \in \widetilde{\mathcal{O}t}^0_{X,Y} & := & \exists y(\tilde{OT}[X,Y,y] \wedge z \in y),\\
z \in \widetilde{\mathcal{O}t}^1_{X,Y} & := & z \in \widetilde{\mathcal{O}t}^0_{X,Y} \vee (\exists w \in X) z = \langle\!\langle \langle\!\langle w, \emptyset \rangle\!\rangle \rangle\!\rangle,\\
z \in \widetilde{\mathcal{O}t}_{X,Y} & := & z \in \widetilde{\mathcal{O}t}^0_{X,Y} \vee z \in X,\\
z \in \eta_X & := & \exists y((y \in X \wedge z = \langle y, \langle\!\langle \langle\!\langle y, \emptyset \rangle\!\rangle \rangle\!\rangle \rangle) \vee (y \notin X \wedge z = \langle y, y \rangle)),\\
z \in \eta^{inv}_X & := & z \in \eta^{-1}_X \wedge (\forall y \in X) z \neq \langle \langle\!\langle \langle\!\langle y, \emptyset \rangle\!\rangle \rangle\!\rangle, \langle\!\langle \langle\!\langle y, \emptyset \rangle\!\rangle \rangle\!\rangle \rangle,
\end{array}
$$

$$
\begin{array}{rcl}
\tilde{OT}[X,Y,y] & := & (\forall f \in y) f \in \mathcal{H}_X \wedge (\forall p \in dom(f)) \exists g \exists h(\\
& & \{g,h\} \subseteq y \cup X \wedge f(p) = \langle\!\langle g,h \rangle\!\rangle \wedge \langle h, \langle\!\langle f(p) \rangle\!\rangle \rangle \in Y \wedge\\
& & (\forall q \in p)(f(q') = f(q) \vee \langle f(q'), f(q) \rangle \in Y^{\varphi})) \wedge\\
& & ((dom(f) \neq \overline{1} \vee f(\emptyset)(\emptyset) \notin X \vee f(\emptyset)(\overline{1}) \neq \emptyset)).
\end{array}
$$

Lemma 147. (Strict Total Order)

$\mathsf{NBG} \vdash Lin[X,Y] \wedge \mathcal{H}_X \cap X = \emptyset \to Lin[\widetilde{\mathcal{O}t}_{X,\widetilde{Y}_X}, \widetilde{Y}_X]$.

Definition 148. (Addition)

We define the expression $f \;\tilde{+}_{X,Y}\; g$ (analogous Definition 18) such that

$$
\begin{array}{rcl}
z \in x \;\tilde{+}^0_Y\; g & := & (\langle x, g(\emptyset) \rangle \in Y^{\varphi} \wedge g \neq \emptyset \wedge z \in g) \vee\\
& & (\langle x, g(\emptyset) \rangle \notin Y^{\varphi} \vee g = \emptyset) \wedge z \in \langle\!\langle x \rangle\!\rangle ^\frown g,\\
z \in f \;\tilde{+}^1_Y\; g & := & \exists p \exists h(p = dom(f) \wedge dom(h) = p' \wedge h(p) = g \wedge\\
& & z \in h(\emptyset) \wedge (\forall q \in p) h(q) = f(q) \;\tilde{+}^0_Y\; h(q')),\\
z \in f \;\tilde{+}_{X,Y}\; g & := & z \in \eta^{inv}_X(\eta_X(f) \;\tilde{+}^1_Y\; \eta_X(g)).
\end{array}
$$

Definition 149. (Multiplication)

We define the expression $f \;\tilde{\cdot}_{X,Y}\; g$ (analogous Definition 18) such that

$$
\begin{array}{rcl}
z \in \downarrow(f) & := & (f = \langle\!\langle \emptyset, f(\overline{1}) \rangle\!\rangle \wedge z \in f(\overline{1})) \vee\\
& & (f \neq \langle\!\langle \emptyset, f(\overline{1}) \rangle\!\rangle \wedge z \in \langle\!\langle f \rangle\!\rangle),\\
z \in \uparrow(f) & := & ((dom(f) \neq \overline{1} \vee f(\emptyset)(\emptyset) = \emptyset) \wedge z \in \langle\!\langle \emptyset, f \rangle\!\rangle) \vee\\
& & (dom(f) = \overline{1} \wedge f(\emptyset)(\emptyset) \neq \emptyset \wedge z \in f(\emptyset)),\\
z \in f \;\tilde{\cdot}^0_Y\; x & := & (f \neq \emptyset \wedge x = \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \wedge z \in f) \vee\\
& & (f \neq \emptyset \wedge x \neq \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \wedge z \in \langle\!\langle \uparrow(\downarrow(f(\emptyset)) \;\tilde{+}^1_Y\; \downarrow(x)) \rangle\!\rangle),\\
z \in f \;\tilde{\cdot}^1_Y\; g & := & \exists p \exists h(p = dom(g) \wedge dom(h) = p' \wedge h(\emptyset) = \emptyset \wedge\\
& & z \in h(p) \wedge (\forall q \in p) h(q') = h(q) \;\hat{+}^1_Y\; (f \;\tilde{\cdot}^0_Y\; g(q))),\\
z \in f \;\tilde{\cdot}_{X,Y}\; g & := & z \in \eta^{inv}_X(\eta_X(f) \;\tilde{\cdot}^1_Y\; \eta_X(g)).
\end{array}
$$

Definition 150. (Exponentiation)

We define the expression $f \mathbin{\tilde{\wedge}_{X,Y}} x$ (analogous Definition 18) such that

$$
\begin{array}{rcl}
z \in f^{\dashv} & := & (f(\emptyset) = \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \wedge \exists p(p' \in dom(f) \wedge z = \langle p, f(p') \rangle)) \vee \\
 & & (f(\emptyset) \neq \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \wedge z \in f), \\
z \in f^{\dashv\!\vdash} & := & (f = \langle\!\langle \emptyset, f(\overline{1}) \rangle\!\rangle \wedge z \in \langle\!\langle \emptyset, f(\overline{1})^{\dashv} \rangle\!\rangle) \vee \\
 & & (f \neq \langle\!\langle \emptyset, f(\overline{1}) \rangle\!\rangle \wedge z \in f), \\
z \in f \mathbin{\tilde{\wedge}^{0}_{Y}} x & := & ((x = \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \vee f = \emptyset) \wedge z \in f) \vee (x \neq \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \wedge \\
 & & ((dom(f) = \overline{1} \wedge f(\emptyset) = \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \wedge z \in f) \vee \\
 & & (dom(f) \notin \overline{2} \wedge f(\emptyset) = \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \wedge z \in \langle\!\langle \uparrow(\langle\!\langle x^{\dashv\!\vdash} \rangle\!\rangle) \rangle\!\rangle) \vee \\
 & & (f(\emptyset) \neq \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \wedge z \in \langle\!\langle \uparrow(\downarrow(f(\emptyset)) \mathbin{\tilde{\cdot}^{1}_{Y}} \langle\!\langle x \rangle\!\rangle) \rangle\!\rangle))), \\
z \in f \mathbin{\tilde{\wedge}^{1}_{Y}} g & := & \exists p \exists h(p = dom(g) \wedge dom(h) = p' \wedge h(\emptyset) = \langle\!\langle \langle\!\langle \emptyset, \emptyset \rangle\!\rangle \rangle\!\rangle \wedge \\
 & & z \in h(p) \wedge (\forall q \in p) h(q') = h(q) \mathbin{\tilde{\cdot}^{1}_{Y}} (f \mathbin{\tilde{\wedge}^{0}_{Y}} g(q))), \\
z \in f \mathbin{\tilde{\wedge}_{X,Y}} g & := & z \in \eta^{inv}_{X}(\eta_X(f) \mathbin{\tilde{\wedge}^{1}_{Y}} \eta_X(g)),
\end{array}
$$

Theorem 151. (Ordinal Notation System)

Let $\widetilde{\mathcal{O}n}$, $\mathcal{O}n_\Gamma$ and $<_\Gamma$ be elementarily definable classes in NBG, such that

$$
\begin{array}{rcl}
\widetilde{\mathcal{O}n} & := & \{x \mid \exists \alpha(x = \widetilde{\alpha})\}, \\
\mathcal{O}n_\Gamma & := & \{x \mid \exists \alpha(x = \alpha \wedge \alpha = \varphi_\alpha(\emptyset))\}, \\
<_\Gamma & := & \widetilde{\in}_{\mathcal{O}n_\Gamma},
\end{array}
$$

where $\widetilde{\in}_{\mathcal{O}n_\Gamma}$ as in Definition 145 with $\in$ as in Lemma 26. If we write $+_\Gamma$, $\cdot_\Gamma$, $\wedge_\Gamma$ for $\hat{+}_{\mathcal{O}n_\Gamma,<_\Gamma}$, $\hat{\cdot}_{\mathcal{O}n_\Gamma,<_\Gamma}$, $\hat{\wedge}_{\mathcal{O}n_\Gamma,<_\Gamma}$, respectively, then we have that

(1) $\mathsf{NBG} \vdash \forall\alpha\forall\beta(\alpha < \beta \leftrightarrow \widetilde{\alpha} <_\Gamma \widetilde{\beta})$,

(2) $\mathsf{NBG} \vdash \forall\alpha(\widetilde{\alpha} \in \widetilde{\mathcal{O}t}_{\mathcal{O}n_\Gamma \cap \alpha', <_\Gamma})$,

(3) $\mathsf{NBG} \vdash \widetilde{\mathcal{O}n} = \widetilde{\mathcal{O}t}_{\mathcal{O}n_\Gamma, <_\Gamma}$,

(4) $\mathsf{NBG} \vdash \forall\alpha\forall\beta(\widetilde{\alpha + \beta} = \widetilde{\alpha} +_\Gamma \widetilde{\beta} \wedge \widetilde{\alpha \cdot \beta} = \widetilde{\alpha} \cdot_\Gamma \widetilde{\beta} \wedge \widetilde{\alpha^\beta} = \widetilde{\alpha} \wedge_\Gamma \widetilde{\beta})$.

Bibliography

[1] A. Cantini, *On the relation between choice and comprehension principles in second order arithmetic*, ***Journal of Symbolic Logic***, vol. 51 (1986), no. 2, pp. 360–373.

[2] S. Feferman and G. Jäger, *Choice principles, the bar rule and autonomously iterated comprehension schemes in analysis*, ***Journal of Symbolic Logic***, vol. 48 (1983), no. 1, pp. 63–70.

[3] S. Feferman and W. Sieg, *Inductive definitions and subsystems of analysis*, ***Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies*** (W. Buchholz, S. Feferman, W. Pohlers, and W. Sieg, eds.), Lecture Notes in Mathematics, vol. 897, Springer, 1981, pp. 16–77.

[4] S. Feferman and W. Sieg, *Proof theoretic equivalences between classical and constructive theories for analysis*, ***Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies*** (W. Buchholz, S. Feferman, W. Pohlers, and W. Sieg, eds.), Lecture Notes in Mathematics, vol. 897, Springer, 1981, pp. 78–142.

[5] Th. Forster, ***Logic, Induction and Sets***, Cambridge University Press, 2003.

[6] H. Friedman, *Iterated inductive definitions and* Σ^1_2-AC, ***Intuitionism and proof theory***, North-Holland Amsterdam, 1970, pp. 435–442.

[7] G. Jäger, *Beweistheorie von* KPN, ***Archiv für mathematische Logik und Grundlagenforschung***, vol. 20 (1980), pp. 53–64.

[8] G. Jäger, *On Feferman's operational set theory* OST, ***Annals of Pure and Applied Logic***, vol. 150 (2007), pp. 19–39.

[9] G. Jäger, *Full operational set theory with unbounded existential quantification and power set*, ***Annals of Pure and Applied Logic***, vol. 200 (2009), pp. 33–52.

[10] G. Jäger and J. Krähenbühl, Σ^1_1*-choice in a theory of sets and classes*, ***Ways of Proof Theory*** (Festschrift on the occasion of Wolfram Pohlers' retirement, Münster, 2008), (R. Schindler, ed.), Ontos-Verlag, 2010, pp. 283–314.

[11] G. Jäger and T. Strahm, *Upper bounds for metapredicative Mahlo in explicit mathematics and admissible set theory*, ***Journal of Symbolic Logic***, vol. 66 (2001), pp. 935–958.

[12] Th. Jech, ***Set theory, the third millennium edition***, Springer, 2002.

[13] A. Levy, *The role of classes in set theory*, ***Sets and Classes. On the Work of Paul Bernays*** (G.-H. Müller, ed.), Studies in Logic and the Foundations of Mathematics, North-Holland, 1976, pp. 277–323.

[14] E. Mendelson, ***Introduction to mathematical logic, fifth edition***, Chapman & Hall, 2010.

[15] W. Pohlers, ***Proof theory, the first step into impredicativity***, Springer, 2009.

[16] K. Schütte, ***Proof Theory***, Springer, 1977.

[17] S. G. Simpson, ***Subsystems of second order arithmetic***, Springer, 1999.

[18] G. Takeuti and W. M. Zaring, ***Introduction to axiomatic set theory***, Springer, 1971.

Index