On Gentzen's first consistency proof for arithmetic

Wilfried Buchholz
Ludwig-Maximilians Universität München

February 14, 2014

Introduction

If nowadays "Gentzen's consistency proof for arithmetic" is mentioned, one usually refers to [Ge38] while Gentzen's first (published) consistency proof, i.e. [Ge36], is widely unknown or ignored. The present paper is intended to change this unsatisfactory situation by presenting [Ge36, IV. Abschnitt] in a slightly modified and modernized form.

The method from [Ge36] can be roughly summarized as follows: By recursion on the build-up of d, for each derivation d in a suitably designed finitary proof system $\mathcal{Z}$ of first order arithmetic a family $(d[n])_{n\in|\mathsf{tp}(d)|}$ of *reduced* $\mathcal{Z}$-derivations is defined such that $\dfrac{\ldots\mathsf{End}(d[n])\ldots(n\in|\mathsf{tp}(d)|)}{\mathsf{End}(d)}$ (where $\mathsf{End}(d)$ denotes the endsequent of d) forms an inference $\mathsf{tp}(d)$ in cutfree ω-arithmetic with repetition rule Rep. Obviously, if d is a derivation of falsum $\perp$, i.e. if $\mathsf{End}(d)=\perp$, then $\mathsf{tp}(d)$ can only be an instance of Rep, so that $d[0]$ is again a derivation of $\perp$. In a second step, to each d an ordinal $o(d)<\varepsilon_0$ is assigned such that $o(d[n])<o(d)$ for all $n\in|\mathsf{tp}(d)|$. Then the consistency of $\mathcal{Z}$ follows by (quantifierfree) transfinite induction up to ε_0.

Actually Gentzen's terminology is somewhat different. First (in §13 of [Ge36]) Gentzen defines *reduction steps on sequents*. Such a reduction step $\mathcal{I}$ may involve a certain 'option' (Wahlfreiheit), so that the result of applying $\mathcal{I}$ to a sequent Π actually is a family of sequents $\big(\mathcal{I}(\Pi,n)\big)_{n\in|\mathcal{I}|}$. Then (in §14 of [Ge36]) for each $\mathcal{Z}$-derivation d (whose endsequent is not an axiom) a *reduction step on derivations*, $d\rightsquigarrow(d[n])_{n\in|\mathcal{I}|}$, is defined such that $\forall n\in|\mathcal{I}|\big(\mathsf{End}(d[n])=\mathcal{I}(\mathsf{End}(d),n)\big)$, where $\mathcal{I}$ is a reduction step on sequents, uniquely determined by d. Here, in contrast to Gentzen, we also regard Rep as a reduction step on sequents — with $|\mathsf{Rep}|=\{0\}$ and $\mathsf{Rep}(\Pi,0)=\Pi$.

The outline of the paper is as follows. In §1 and §2 we repeat relevant parts of [Ge36] using to a great extent Gentzen's own words (in the translation by M. E. Szabo [Sz69]). Thereby we do not hesitate to deviate from the original text (in content or form) whenever we think it is appropriate or facilitates understanding. The main point where we deviate from [Ge36] (besides omitting conjunction &) is the following: In the reduction steps on sequents concerning an antecedent formula $\forall xF$ or $\neg A$ (13.51, 13.53) we always require that this formula is retained in the reduced sequent while Gentzen allows to omit it. As a consequence we also have to modify the reduction steps on atomic $\mathcal{Z}$-derivations (which will be deferred till §5). In §3 we present the main definitions and proofs of §2 in a more condensed style (and with some further modifications). This facilitates the work in §4 where we assign to each $\mathcal{Z}$-derivation d an ordinal $o(d)<\varepsilon_0$ and prove that each reduction step on a derivation d lowers its ordinal, i.e. we prove that $o(d[n])<o(d)$ for all $n\in|\mathsf{tp}(d)|$. Our ordinal assignment is essentially that of [KB81] which on first sight looks very different from Gentzen's original assignment in [Ge36], where certain finite decimal fractions were used as notations for ordinals $<\varepsilon_0$. But in the appendix we will show that actually both ordinal assignments are rather closely related. In §6 we give an interpretation of $\mathcal{Z}$ in an infinitary system $\mathcal{Z}^\infty$. This way we obtain a semantic explanation for Gentzen's reduction steps on $\mathcal{Z}$-derivations and for the ordinal assignment of §4. Finally, in §7 we indicate how the approach of §§3,4 can easily be adapted to calculi with multisuccedent sequents.

§1 Formal language, reduction steps on sequents

The following symbols will serve for the formation of formulae: *Variables* (for natural numbers) which are divided into *free* and *bound* variables; the constant 0 and the unary function symbol S (successor); *predicate symbols* (each of a fixed arity); the logical connectives $\neg, \forall$. [2]

Terms are generated from the constant 0 and free variables by iterated application of S.
The terms $0, \mathsf{S}0, \mathsf{SS}0, \ldots$ are called *numerals*. In the following we identify numerals and natural numbers.

Formulas:

1. If P is an n-ary predicate symbol and $t_1, \ldots, t_n$ are terms, then $Pt_1 \ldots t_n$ is a (prime) formula.
   If $t_1, \ldots, t_n$ are numerals, then $Pt_1 \ldots t_n$ is called a *minimal formula*.
2. If A is a formula, then so is $\neg A$.
3. From a given formula we obtain another formula by replacing a free variable by a bound variable x not yet occurring in the formula and prefixing $\forall x$.

We assume that to each minimal formula a *truth value* "true" or "false" is assigned.

We use $\perp$ as abbreviation for some fixed *false minimal formula* (e.g. $0 = \mathsf{S}0$).

Abbreviation. $A \approx B \;:\Leftrightarrow$ either $A = B$ or A, B are both false minimal formulas.

Remark. $A \approx \perp \;\Leftrightarrow\; A$ is a false minimal formula.

A *sequent* is an expression of the form $\Gamma {\rightarrow} B$ where Γ is a finite sequence of formulae.

The formulae in Γ are called the *antecedent formulae* and B the *succedent formula* of the sequent.
We also call Γ the *antecedent* of $\Gamma {\rightarrow} B$.

A formula (sequent) is called *closed* if no free variable occurs in it.

Abbreviations.
$A \in \Gamma \;:\Leftrightarrow\; A$ occurs in the sequence Γ.
$\Gamma \subseteq \Gamma' \;:\Leftrightarrow$ for all formulas A, if $A \in \Gamma$ then $A \in \Gamma'$ (e.g. $A, B, A, A \subseteq B, B, A, C$).

Definition (Reduction steps on sequents)

On a closed sequent Π an individual *reduction step* can be carried out in the following way.

13.21. Suppose that the succedent formula of the sequent Π has the form $\forall x F(x)$. In that case we replace it by a formula $F(n)$, i.e., by a formula which results from $F(x)$ by the substitution of an *arbitrarily chosen numeral* n for the variable x.

13.23. Suppose that the succedent formula of the sequent Π has the form $\neg A$. In that case we replace it by the formula $\perp$ and, at the same time, adjoin the formula A to the antecedent of the sequent.

13.4. Suppose that the succedent formula of the sequent Π is a true minimal formula; or: that the succedent formula is a false minimal formula and that one of the antecedent formulae of Π is also a false minimal formula. Then we say that the sequent Π has (or, is in) *endform*, and no reduction step is defined.

13.5. Suppose that the succedent formula of Π is a false minimal formula, and that none of the antecedent formulae of Π is a false minimal formula. In that case the following two different kinds of reduction step are permissible (counterpart of 13.2):

13.51. Suppose that an antecedent formula has the form $\forall x F(x)$. We adjoin a formula $F(k)$ (k an arbitrary numeral) to the antecedent.

[2] We omit conjunction '&' in order to keep the focus on the essential things.

13.53. Suppose that an antecedent formula has the form $\neg A$. We replace the succedent formula by A.

In condensed form these reduction steps are described by the following schemata (reading them bottom-up):

$(\mathsf{R}_{\forall xF(x)})\ \dfrac{\ldots\Gamma\to F(n)\ldots(n\in\mathbb{N})}{\Gamma\to\forall xF(x)}$; $\qquad (\mathsf{R}_{\neg A})\ \dfrac{A,\Gamma\to\bot}{\Gamma\to\neg A}$;

$(\mathsf{L}^k_{\forall xF(x)})\ \dfrac{F(k),\Gamma\to C}{\Gamma\to C}$ with $C\approx\bot$ and $\forall xF(x)\in\Gamma$;

$(\mathsf{L}^0_{\neg A})\ \dfrac{\Gamma\to A}{\Gamma\to C}$ with $C\approx\bot$ and $\neg A\in\Gamma$.

In the sequel, each of the symbols $\mathsf{R}_{\forall xF}$, $\mathsf{R}_{\neg A}$, $\mathsf{L}^0_{\forall xF}$, $\mathsf{L}^0_{\neg A}$ is used as the name of the respective reduction step (as shown above). But the above schemata can also be read as inferences in ω-arithmetic; therefore the symbols $\mathsf{R}_{\forall xF}$, $\mathsf{R}_{\neg A}$, $\mathsf{L}^0_{\forall xF}$, $\mathsf{L}^0_{\neg A}$ will also be called *inference symbols*. Another reason is that this term has already been used in several previous publications (e.g., in [Bu97]) — and "reduction step symbol" would sound too clumsy.

§2 Reduction steps on derivations

Definition (The system $\mathcal{Z}$ of pure number theory).

Derivable objects of $\mathcal{Z}$ are sequents.

The *axioms* (or *initial sequents*) of $\mathcal{Z}$ will be specified in §5.

Inference Rules

$\forall$-introduction: $\dfrac{\Gamma\to F(a)}{\Gamma\to\forall xF(x)}$, if the free variable a does not occur in the conclusion.

$\neg$-introduction: $\dfrac{A,\Gamma\to\bot}{\Gamma\to\neg A}$

complete induction: $\dfrac{\Gamma\to F(0)\qquad F(a),\Gamma\to F(\mathsf{S}a)}{\Gamma\to F(t)}$, if the free variable a does not occur in the conclusion.

chain rule: $\dfrac{\Gamma_0\to A_0\ \ \ldots\ldots\ \ \Gamma_l\to A_l}{\Gamma\to C}$, if there exists $j\le l$ such that $C\approx A_j$ and $\forall i\le j(\Gamma_i\subseteq\Gamma,A_0,\ldots,A_{i-1})$.

In addition we require that no free variable is vanishing, i.e., that every free variable occurring in one of the premises $\Gamma_i\to A_i$, also occurs in the conclusion $\Gamma\to C$.

Abbreviation.
$d\vdash\Gamma\to C\ :\Leftrightarrow\ d$ is a $\mathcal{Z}$-derivation (i.e., a derivation in $\mathcal{Z}$) and the endsequent of d is $\Gamma\to C$.

A derivation is called *closed* if its endsequent is closed.

For each closed derivation d, whose endsequent is not in endform (13.4) we shall now define *the reduction step on d* and at the same time prove the following: by such a step the derivation is transformed into another closed derivation and its endsequent is thereby modified in the following way: At most one reduction step according to 13.2 or 13.5 is carried out on the sequent. (It may thus happen that an endsequent remains entirely unchanged.) The reduction step on a derivation is unambiguous, except in the case in which the endsequent undergoes a transformation according to a reduction step on sequents involving a *choice* (13.21); here the choice may be made arbitrarily; if this has been done, the reduction step is then also unambiguous. If the endsequent of d has endform according to 13.4, *no* reduction step is defined for this derivation.

Definitions.

1. The result of carrying out the reduction step on d is denoted by $d[n]$ where in case 13.21, n is the *'arbitrarily choosen numeral'*, and $n = 0$ otherwise.

2. If the reduction step on d causes a reduction step on the endsequent Π of d then $\mathsf{tp}(d)$ denotes the name of this latter reduction step[1] and $\mathsf{tp}(d)(\Pi, n)$ denotes the result of applying $\mathsf{tp}(d)$ to Π, where n plays the same role as in 1.

3. If the reduction step on d does not change the endsequent of d, we set $\mathsf{tp}(d) := \mathsf{Rep}$ and $\mathsf{Rep}(\Pi, n) := \Pi$.

4. The *arity* of d is defined by $\mathsf{arity}(d) := \begin{cases} \mathbb{N} & \text{if } \mathsf{tp}(d) = \mathsf{R}_{\forall x F} \\ \emptyset & \text{if the endsequent of } d \text{ has endform} \\ \{0\} & \text{otherwise} \end{cases}$

Summing up, by recursion on the build-up of d we will define $\mathsf{tp}(d)$ and $d[n]$ and simultaneously prove

Theorem 2.1.
If d is a closed $\mathcal{Z}$-derivation of Π and if $n \in \mathsf{arity}(d)$, then $d[n] \vdash \mathsf{tp}(d)(\Pi, n)$.

In the following we assume that d is a closed $\mathcal{Z}$-derivation whose endsequent is *not* in endform.

14.21. The axioms of $\mathcal{Z}$ are treated later (in §5).

14.22. We now consider the case where the endsequent is the result of the application of a *rule of inference* and we presuppose that for the derivations of the premises the reduction step has already been defined and the validity of the associated assertion (i.e. Theorem 2.1) demonstrated.

14.23. Suppose that the endsequent of d is the result of a $\forall$-*introduction* or a $\neg$-*introduction.* It (i.e. the endsequent) is then eliminated and its premise taken for the new endsequent, where, in the case of a $\forall$-introduction, every occurrence of the free variable a must be replaced throughout the derivation d_0 of this premise by an arbitrarily chosen numeral n.

The derivation has obviously remained correct, and the endsequent has become a reduced endsequent in the sense of 13.21 or 13.23.

In other words:

If $d = \left\{ \dfrac{\begin{array}{c} d_0(a) \\ | \\ \Gamma \to F(a) \end{array}}{\Gamma \to \forall x F(x)} \right.$, then $d[n] := d_0(n) = \left\{ \begin{array}{c} d_0(n) \\ | \\ \Gamma \to F(n) \end{array} \right.$ and $\mathsf{tp}(d) := \mathsf{R}_{\forall x F(x)}$.

If $d = \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ A, \Gamma \to \bot \end{array}}{\Gamma \to \neg A} \right.$, then $d[0] := d_0$ and $\mathsf{tp}(d) := \mathsf{R}_{\neg A}$.

14.24. Suppose that the endsequent of d is the result of a *'complete induction'.*

$d = \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ \Gamma \to F(0) \end{array} \qquad \begin{array}{c} d_1(a) \\ | \\ F(a), \Gamma \to F(\mathsf{S}a) \end{array}}{\Gamma \to F(n)} \right.$ (Since d is closed, the induction term is a numeral k.)

Then we set $d[0] := \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ \Gamma \to F(0) \end{array} \quad \begin{array}{c} d_1(0) \\ | \\ F(0), \Gamma \to F(1) \end{array} \quad \begin{array}{c} d_1(1) \\ | \\ F(1), \Gamma \to F(2) \end{array} \;\ldots\ldots\; \begin{array}{c} d_1(k-1) \\ | \\ F(k-1), \Gamma \to F(k) \end{array}}{\Gamma \to F(k)} \right.$

and $\mathsf{tp}(d) := \mathsf{Rep}$.

[1] cf. end of §1

14.25. The last case to be considered is that in which the endsequent is the conclusion of a *'chain rule'* inference: $d = \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ \Gamma_0 \to A_0 \end{array} \ldots\ldots \begin{array}{c} d_l \\ | \\ \Gamma_l \to A_l \end{array}}{\Theta \to D} \right.$

The premise whose succedent formula provides the succedent formula of the endsequent, I shall call the *'major premise'*. If the succedent of the endsequent is a false minimal formula, we choose as major premise the *first* premise (in the given order) whose succedent formula is *also* a false minimal formula. This does not change the correctness of the 'chain rule' inference.

So there is a $j \le l$ such that $A_j \approx D$, $\forall i \le j(\Gamma_i \subseteq \Theta, A_0, \ldots, A_{i-1})$ and, if A_j is a false minimal formula then none of $A_0, \ldots, A_{j-1}$ is a false minimal formula.

From these preliminaries it follows that the major premise $\Gamma_j \to A_j$ can in no case be in endform (13.4), for otherwise the endsequent $\Theta \to D$ would obviously also have to be in endform, and this was assumed not to be the case. Hence a *reduction step* can be carried out on the derivation of the major premise. In respect on this reduction step, i.e. in respect on $\mathsf{tp}(d_j)$, I distinguish four cases (14.251-14.254).

14.251. Suppose that the major premise undergoes a change according to 13.2 in the reduction step on its derivation d_j, i.e. $\mathsf{tp}(d_j) = \mathsf{R}_{A_j}$ and $A_j = D$. In that case the endsequent is subjected to the appropriate reduction step for sequents according to 13.2; any *choice* that arises is to be made arbitrarily. The reduction step for derivations is then carried out on the derivation d_j of the major premise and, whenever a choice exists, the *same* choice is to be made as before. The succedent formulae of both sequents are now the same once again and the 'chain rule' inference is once again correct. Thus, the reduction step for the entire derivation d is completed.

In other words, $\mathsf{tp}(d) := \mathsf{tp}(d_j)$ and

$$d[n] := \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ \Gamma_0 \to A_0 \end{array} \ldots \begin{array}{c} d_j[n] \\ | \\ \Gamma_j \to F(n) \end{array} \ldots \begin{array}{c} d_l \\ | \\ \Gamma_l \to A_l \end{array}}{\Theta \to F(n)} \right. \quad \text{if } A_j = D = \forall x F(x);$$

$$d[0] := \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ \Gamma_0 \to A_0 \end{array} \ldots \begin{array}{c} d_j[0] \\ | \\ A, \Gamma_j \to \bot \end{array} \ldots \begin{array}{c} d_l \\ | \\ \Gamma_l \to A_l \end{array}}{A, \Theta \to \bot} \right. \quad \text{if } A_j = D = \neg A.$$

14.252. Suppose that the major premise undergoes a change according to 13.5 in the reduction step on its derivation, and the affected antecedent formula B also occurs in the antecedent of the endsequent, i.e. $\mathsf{tp}(d_j) = \mathsf{L}^k_B$ with $B \in \Theta$. In that case the reduction step is carried out on the derivation of the major premise and the endsequent is modified according to the *corresponding* reduction step on sequents (13.5), so that the 'chain rule' inference becomes again correct.

In other words, $\mathsf{tp}(d) := \mathsf{tp}(d_j)$ and

$$d[0] := \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ \Gamma_0 \to A_0 \end{array} \ldots \begin{array}{c} d_j[0] \\ | \\ F(k), \Gamma_j \to A_j \end{array} \ldots \begin{array}{c} d_l \\ | \\ \Gamma_l \to A_l \end{array}}{F(k), \Theta \to D} \right. \quad \text{if } B = \forall x F(x);$$

$$d[0] := \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ \Gamma_0 \to A_0 \end{array} \ldots \begin{array}{c} d_j[0] \\ | \\ \Gamma_j \to A \end{array} \ldots \begin{array}{c} d_l \\ | \\ \Gamma_l \to A_l \end{array}}{\Theta \to A} \right. \quad \text{if } B = \neg A.$$

14.253. (Principal case) Suppose that the major premise, say $\Delta\to C$, undergoes a change according to 13.5 in the reduction step on its derivation and that the affected antecedent formula (V) is a formula that does not occur among the antecedent formulae of the endsequent, since it agrees with the succedent formula of an earlier premise; suppose further that *this* premise, call it $\Gamma\to V$, undergoes a *change* during the reduction step on its derivation which, in that case, must necessarily be a change according to 13.2. (Since V cannot be a minimal formula.) – Suppose that the endsequent of the whole derivation has the form $\Theta\to D$. I shall distinguish *two subcases* depending on whether V has the form $\forall x F(x)$ or $\neg A$.

Suppose first that V has the form $\forall x F(x)$. In that case an antecedent formula $F(k)$ is adjoined in the reduction step according to 13.51 on $\Delta\to C$; in the reduction step on $\Gamma\to\forall x F(x)$ which must be carried out according to 13.21, the *same* symbol k may be chosen for the numeral to be substituted, so that $\Gamma\to F(k)$ results. We now form three 'chain-rule' inferences: the premises of the first are those of the original 'chain-rule' inference, but with $\Gamma\to F(k)$ in place of $\Gamma\to\forall x F(x)$; its conclusion: $\Theta\to F(k)$. A correct result. The premises of the second are those of the original 'chain rule' inference, except that $\Delta\to C$ is replaced by the sequent that was reduced according to 13.51; its conclusion: $F(k), \Theta\to D$. This is also a correct 'chain rule' inference. The third 'chain-rule' inference again yields the endsequent $\Theta\to D$ from $\Theta\to F(k)$ and $F(k), \Theta\to D$. Together with each one of the sequents used we must of course write down the complete derivation of each sequent so that altogether we now have another correct derivation.

If V has the form $\neg A$, then $\Delta\to C$ is reduced to $\Delta\to A$, and $\Gamma\to\neg A$ to $A, \Gamma\to\bot$. We now form, as before, two 'chain-rule' inferences with the conclusions $A, \Theta\to\bot$ and $\Theta\to A$. With their order interchanged, these two yield by a third 'chain-rule' inference again $\Theta\to D$. (Note that D, like C and $\bot$, is a false minimal formula.)

In other words, if $d = \left\{ \begin{array}{c} \begin{array}{cc} d_i & d_j \\ | & | \\ \ldots\Gamma\to V & \ldots\Delta\to C\ \ldots \end{array} \\ \hline \Theta\to D \end{array} \right.$ with major premise $\Delta\to C$, $\mathsf{tp}(d_j) = \mathsf{L}^k_V$ and $V \notin \Theta$,

we set $\mathsf{tp}(d) := \mathsf{Rep}$, while the reduced derivation $d[0]$ depends on the form of V.

If $V = \forall x F(x)$, then $d[0] := \left\{ \begin{array}{c} \begin{array}{cc} d\{0\} & d\{1\} \\ | & | \\ \Theta\to F(k) & F(k), \Theta\to D \end{array} \\ \hline \Theta\to D \end{array} \right.$

where $\quad d\{0\} := \left\{ \begin{array}{c} \begin{array}{cc} d_i[k] & d_j \\ | & | \\ \ldots\ \Gamma\to F(k)\ \ldots & \Delta\to C\ \ldots \end{array} \\ \hline \Theta\to F(k) \end{array} \right.$ and $\quad d\{1\} := \left\{ \begin{array}{c} \begin{array}{cc} d_i & d_j[0] \\ | & | \\ \ldots\Gamma\to V\ \ldots & F(k), \Delta\to C\ \ldots \end{array} \\ \hline F(k), \Theta\to D \end{array} \right.$

If $V = \neg A$, then $\ d[0] := \left\{ \begin{array}{c} \begin{array}{cc} d\{0\} & d\{1\} \\ | & | \\ \Theta\to A & A, \Theta\to\bot \end{array} \\ \hline \Theta\to D \end{array} \right.$

where $\quad d\{0\} := \left\{ \begin{array}{c} \begin{array}{cc} d_i & d_j[0] \\ | & | \\ \ldots\Gamma\to V\ \ldots & \Delta\to A\ \ldots \end{array} \\ \hline \Theta\to A \end{array} \right.$ and $\quad d\{1\} := \left\{ \begin{array}{c} \begin{array}{cc} d_i[0] & d_j \\ | & | \\ \ldots\ A, \Gamma\to\bot\ \ldots & \Delta\to C\ \ldots \end{array} \\ \hline A, \Theta\to\bot \end{array} \right.$

14.254. We are still left with the following possibilities: the major premise remains *unchanged* in the reduction step on its derivation; or: its change is of the kind assumed at 14.253, and the premise $\Gamma\to V$ remains *unchanged* in the reduction step on its derivation. — In both cases we carry out the reduction step on the derivation of the unchanged remaining premise, and this completes the reduction.

However, if this reduction step on the derivation of the unchanged remaining premise is according to 14.253, we proceed somewhat *differently*, namely: we carry out this reduction step, but *without* completing the prescribed third ‘chain-rule’ inference; instead, we take the two premises of this ‘chain-rule’ inference and insert them in place of its conclusion in the sequence of premises of that ‘chain-rule’ inference which concludes the derivation as a whole. This obviously leaves that ‘chain-rule’ inference correct. The endsequent is not changed.

Let us have a closer look on one of these cases; namely the case where the premise $\Delta{\to}C$ $(=\Gamma_j{\to}A_j)$ remains unchanged in the reduction step on its derivation d_j, and where this reduction step is according to 14.253.

Then $d_j[0] = \left\{ \dfrac{\begin{matrix} d_j\{0\} \\ | \\ \Gamma_j{\to}B \end{matrix} \quad \begin{matrix} d_j\{1\} \\ | \\ B,\Gamma_j{\to}A_j' \end{matrix}}{\Gamma_j{\to}A_j} \right.$ for some B and $A_j' \approx A_j \approx D$.

We set $\mathsf{tp}(d) := \mathsf{Rep}$ and

$$d[0] := \left\{ \dfrac{\begin{matrix} d_0 \\ | \\ \Gamma_0{\to}A_0 \end{matrix} \ \dots \ \begin{matrix} d_{j-1} \\ | \\ \Gamma_{j-1}{\to}A_{j-1} \end{matrix} \quad \begin{matrix} d_j\{0\} \\ | \\ \Gamma_j{\to}B \end{matrix} \quad \begin{matrix} d_j\{1\} \\ | \\ B,\Gamma_j{\to}A_j' \end{matrix} \quad \begin{matrix} d_{j+1} \\ | \\ \Gamma_{j+1} \to A_{j+1} \end{matrix} \ \dots \ \begin{matrix} d_l \\ | \\ \Gamma_l{\to}A_l \end{matrix}}{\Theta{\to}D} \right.$$

The definition of the *reduction step on a derivation* and the proof of Theorem 2.1 are now complete.

As an immediate consequence from Theorem 2.1 one obtains

Corollary 2.1.
If $d \vdash {\to}\bot$ then $d[0] \vdash {\to}\bot$.

Proof:

By Theorem 2.1 we get $d[0] \vdash \mathsf{tp}(d)({\to}\bot, 0)$. Since no reduction step is applicable to ${\to}\bot$, it cannot be that $\mathsf{tp}(d)$ is R_A or L^k_A. Hence $\mathsf{tp}(d) = \mathsf{Rep}$ and thus $\mathsf{tp}(d)({\to}\bot, 0) = {\to}\bot$.

Remark (Consistency of $\mathcal{Z}$). In §4 we will assign to each $\mathcal{Z}$-derivation d an ordinal $o(d) < \varepsilon_0$ and prove that $o(d[n]) < o(d)$ whenever $d[n]$ is defined (Theorem 4.2). Together with Corollary 2.1 this implies the consistency of $\mathcal{Z}$ via (quantifierfree) induction up to ε_0.

§3 Reduction steps on derivations revisited

In this section we present the contents of §§1,2 in a more condensed style. In the course of this we also carry out some minor modifications on Gentzen’s original approach, namely

- In the reduction steps $\mathsf{L}^k_{\forall xF}$ and $\mathsf{L}^0_{\neg A}$ it is no longer required that the succedent C is a false minimal formula. Accordingly the notion “endform” will be modified, and the condition “$A_j \approx C$” in the chain rule will be replaced by “$A_j \in \{C, \bot\}$”.
- Each chain rule inference will now have an explicitly shown *rank* which is an upper bound on the ranks of all its cut formulas.

Some preliminary definitions and abbreviations

1. $A \approx \top \ :\Leftrightarrow\ A$ is a true minimal formula.
2. $\Gamma{\to}C$ has (or, is in) *endform* $:\Leftrightarrow\ C \approx \top$ or Γ contains a false minimal formula.
3. $\mathrm{rk}(C) := \begin{cases} 0 & \text{if } C \text{ is atomic} \\ \mathrm{rk}(A)+1 & \text{if } C = \forall x A \text{ or } C = \neg A \end{cases}$

4. If X is a formula or sequent, then $\mathrm{FV}(X)$ denotes the set of all free variables occurring in X.
5. Π ranges over sequents; for $\Pi = \Gamma \to C$ we set $A, \Pi := A, \Gamma \to C$ and $\Pi{\cdot}A := \Gamma \to A$.
6. An *inference symbol* is an expression of one of the following three kinds:
   R_A with $\mathrm{rk}(A) > 0$ or $A \approx \top$, $\quad \mathsf{L}^k_A$ with $\mathrm{rk}(A) > 0$ or $A \approx \bot$, $\quad \mathsf{Rep}$.
7. For each inference symbol $\mathcal{I}$ we define
   - its *arity* $|\mathcal{I}| := \begin{cases} \mathbb{N} & \text{if } \mathcal{I} = \mathsf{R}_{\forall x F} \\ \emptyset & \text{if } \mathcal{I} = \mathsf{R}_A \text{ or } \mathcal{I} = \mathsf{L}^k_A \text{ with } \mathrm{rk}(A) = 0 \\ \{0\} & \text{otherwise} \end{cases}$
   - the result of applying (the reduction step denoted by) $\mathcal{I}$ to Π under choice n:
   $$\mathcal{I}(\Pi, n) := \begin{cases} \Pi{\cdot}F(n) & \text{if } \mathcal{I} = \mathsf{R}_{\forall x F} \\ F(k), \Pi & \text{if } \mathcal{I} = \mathsf{L}^k_{\forall x F} \\ A, \Pi{\cdot}\bot & \text{if } \mathcal{I} = \mathsf{R}_{\neg A} \\ \Pi{\cdot}A & \text{if } \mathcal{I} = \mathsf{L}^k_{\neg A} \\ \Pi & \text{otherwise} \end{cases}$$
   - the relation $\mathcal{I} \lhd \Pi$ ($\mathcal{I}$ is permissible for Π):
   $\mathcal{I} \lhd \Gamma \to C \quad :\Leftrightarrow \quad \mathcal{I} = \mathsf{R}_C \;$ or $\; \mathcal{I} = \mathsf{L}^k_A$ with $A \in \Gamma \;$ or $\; \mathcal{I} = \mathsf{Rep}$

Definition.
The figure $\dfrac{\Gamma_0 \to A_0 \quad \ldots \quad \Gamma_l \to A_l}{\Gamma \to C}$ is called a *chain rule inference of rank* r if there exists a $j \le l$ such that $A_j \in \{C, \bot\}$ and $\forall i \le j(\Gamma_i \subseteq \Gamma, A_0, \ldots, A_{i-1})$ and $\forall i < j(\mathrm{rk}(A_i) \le r)$.

Inductive Definition of $d \vdash \Pi$ (d is a $\mathcal{Z}$-derivation with endsequent Π)

1. Atomic derivations (axioms): cf. §5.
2. If $d_0 \vdash \Gamma \to F(a)$ and $a \notin \mathrm{FV}(\Gamma \to \forall x F(x))$, then $\mathrm{I}^a_{\forall x F(x)} d_0 \vdash \Gamma \to \forall x F(x)$.
3. If $d_0 \vdash A, \Gamma \to \bot$, then $\mathrm{I}_{\neg A} d_0 \vdash \Gamma \to \neg A$.
4. If $d_0 \vdash \Gamma \to F(0)$ and $d_1 \vdash F(a), \Gamma \to F(\mathsf{S}a)$ and $a \notin \mathrm{FV}(\Gamma \to F(t))$, then $\mathrm{Ind}^{a,t}_F d_0 d_1 \vdash \Gamma \to F(t)$.
5. If $d_i \vdash \Pi_i$ with $\mathrm{FV}(\Pi_i) \subseteq \mathrm{FV}(\Pi)$ for $i = 0, \ldots, l$, and
   if $\dfrac{\Pi_0 \quad \ldots \quad \Pi_l}{\Pi}$ is a *chain rule inference of rank* r, then $\quad \mathsf{K}^r_\Pi d_0 \ldots d_l \vdash \Pi$.

A derivation is called *closed* iff its endsequent is closed.

Lemma 3.1.
Assume $\Pi_i = \Gamma_i \to A_i$ $(i = 0, \ldots, j_0)$ and $\Pi = \Gamma \to C$ with
$A_{j_0} \in \{C, \bot\}$ and $\forall i \le j_0(\Gamma_i \subseteq \Gamma, A_0, \ldots, A_{i-1})$.

Further, let $\mathcal{I}_0, \ldots, \mathcal{I}_{j_0}$ be inference symbols such that $\forall i \le j_0(\mathcal{I}_i \lhd \Pi_i \;\&\; \mathcal{I}_i \not\lhd \Pi)$.

Then $\exists i, j, k(i < j \le j_0 \;\&\; \mathcal{I}_i = \mathsf{R}_{A_i} \;\&\; \mathcal{I}_j = \mathsf{L}^k_{A_i} \;\&\; 0 < \mathrm{rk}(A_i))$.

Proof :
From $\mathcal{I}_{j_0} \lhd \Pi_{j_0} \;\&\; \mathcal{I}_{j_0} \not\lhd \Pi \;\&\; A_{j_0} \in \{C, \bot\}$ it follows that $\mathcal{I}_{j_0} \in \mathsf{L}$ (i.e. $\mathcal{I}_{j_0} = \mathsf{L}^k_B$ for some B and k). Hence there exists the least $j \le j_0$ such that $\mathcal{I}_j \in \mathsf{L}$. Assume $\mathcal{I}_j = \mathsf{L}^k_B$. Then $\mathsf{L}^k_B \lhd \Pi_j \;\&\; \mathsf{L}^k_B \not\lhd \Pi$ which implies $B \in \Gamma_j \setminus \Gamma \subseteq \{A_0, \ldots, A_{j-1}\}$. So we have $\mathcal{I}_j = \mathsf{L}^k_{A_i}$ for some $i < j$. By minimality of j and since $i < j \le j_0$, we have $\mathcal{I}_i \notin \mathsf{L}$ and $\mathcal{I}_i \lhd \Pi_i \;\&\; \mathcal{I}_i \not\lhd \Pi$, which implies $\mathcal{I}_i = \mathsf{R}_{A_i}$. Finally, from $\mathsf{R}_{A_i} \lhd \Pi_i \;\&\; \mathsf{L}^k_{A_i} \lhd \Pi_j$ we conclude $(\mathrm{rk}(A_i) = 0 \Rightarrow A_i \approx \top) \;\&\; (\mathrm{rk}(A_i) = 0 \Rightarrow A_i \approx \bot)$, hence $\mathrm{rk}(A_i) > 0$.

Definition 3.2. ($\mathsf{tp}(d)$ and $d[n]$)
For each closed $\mathcal{Z}$-derivation d we now define an inference symbol $\mathsf{tp}(d)$ and, for each $n \in |\mathsf{tp}(d)|$, a closed $\mathcal{Z}$-derivation $d[n]$. In the main case 5.1. where d is 'critical' we also define the auxiliary derivations $d\{0\}$, $d\{1\}$ and the formula $\mathsf{A}(d)$. The whole definition proceeds by recursion on the build-up of d. In parallel we observe that $\mathsf{tp}(d)$ is permissible for Π (i.e., $\mathsf{tp}(d) \triangleleft \Pi$) whenever $d \vdash \Pi$.

1. d atomic: cf. §5.

2. $d = \mathrm{I}^a_{\forall x F} d_0$: Then $\mathsf{tp}(d) := \mathsf{R}_{\forall x F}$ and $d[n] := d_0(a/n)$.

3. $d = \mathrm{I}_{\neg A} d_0$: Then $\mathsf{tp}(d) := \mathsf{R}_{\neg A}$ and $d[0] := d_0$.

4. $d = \mathrm{Ind}^{a,k}_F d_0 d_1$ with $d_0 \vdash \Gamma \to F(0)$ and $d_1 \vdash F(a), \Gamma \to F(\mathsf{S}a)$:
Then $\mathsf{tp}(d) := \mathsf{Rep}$ and $d[0] := \mathsf{K}^r_{\Gamma \to F(k)} d_0\, d_1(a/0) \dots d_1(a/k{-}1)$, where $r := \mathrm{rk}(F)$.

5. $d = \mathsf{K}^r_\Pi d_0 \dots d_l$ with $\Pi = \Gamma \to C$ and $d_i \vdash \Pi_i = \Gamma_i \to A_i$ $(i \le l)$:
Abbreviation: $\mathsf{K}^r_{\Pi'}(i/d'_1 \dots d'_m) := \mathsf{K}^r_{\Pi'} d_0 \dots d_{i-1} d'_1 \dots d'_m d_{i+1} \dots d_l$.
Let j_0 be minimal s.t. $A_{j_0} \in \{C, \bot\}$ & $\forall i \le j_0 (\, \Gamma_i \subseteq \Gamma, A_0, \dots, A_{i-1}\,)$.
We say that d is *critical* if $\forall i \le j_0 (\mathsf{tp}(d_i) \not\triangleleft \Pi)$.

5.1. d critical:
Then due to Lemma 3.1, and since $\forall i \le l (\mathsf{tp}(d_i) \triangleleft \Pi_i)$ there exists a pair (i,j) such that $i < j \le j_0$, $\mathsf{tp}(d_i) = \mathsf{R}_{A_i}$, $\mathsf{tp}(d_j) = \mathsf{L}^k_{A_i}$ (for some k) and $0 < \mathrm{rk}(A_i)$.
We take the least such pair and set $\mathsf{tp}(d) := \mathsf{Rep}$ and $d[0] := \mathsf{K}^{r-1}_\Pi d\{0\} d\{1\}$ where

$$d\{0\} := \mathsf{K}^r_{\Pi \centerdot \mathsf{A}(d)} \begin{cases} (i/d_i[k]) & \text{if } A_i = \forall x F \\ (j/d_j[0]) & \text{if } A_i = \neg A \end{cases}, \quad d\{1\} := \mathsf{K}^r_{\mathsf{A}(d), \Pi} \begin{cases} (j/d_j[0]) & \text{if } A_i = \forall x F \\ (i/d_i[0]) & \text{if } A_i = \neg A \end{cases},$$

and $\mathsf{A}(d) := \begin{cases} F(k) & \text{if } A_i = \forall x F \\ A & \text{if } A_i = \neg A \end{cases}$.

5.2. d not critical: Let $i \le j_0$ be minimal such that $\mathsf{tp}(d_i) \triangleleft \Pi$.

5.2.1. d_i critical:
Then $\mathsf{tp}(d) := \mathsf{Rep}$ and $d[0] := \mathsf{K}^{r'}_\Pi(i/d_i\{0\}, d_i\{1\})$ with $r' := \max\{\mathrm{rk}(\mathsf{A}(d_i)), r\}$.

5.2.2. d_i not critical: Then $\mathsf{tp}(d) := \mathsf{tp}(d_i)$ and $d[n] := \mathsf{K}^r_{\mathsf{tp}(d)(\Pi,n)}(i/d_i[n])$.

Lemma 3.3. If $d \vdash \Pi$, then $\mathsf{tp}(d) \triangleleft \Pi$.

Theorem 3.4. For $d \vdash \Pi$ the following holds:
(a) If $d = \mathsf{K}^r_\Pi d_0 \dots d_l$ is critical, then $d\{0\} \vdash \Pi \centerdot \mathsf{A}(d)$, $d\{1\} \vdash \mathsf{A}(d), \Pi$, and $\mathrm{rk}(\mathsf{A}(d)) < r$.
(b) $\forall n \in |\mathsf{tp}(d)| \big(d[n] \vdash \mathsf{tp}(d)(\Pi, n) \big)$.

Proof by simultaneous induction on the build-up of d:
(a) The premise "d critical" yields that we are in Case 5.1 of Definition 3.2.

Subcase $A_i = \forall x F$:
By assumption we have $d_\nu \vdash \Pi_\nu$ for all $\nu \le l$. From $d_i \vdash \Pi_i$ and $d_j \vdash \Pi_j$ together with $\mathsf{tp}(d_i) = \mathsf{R}_{A_i}$ and $\mathsf{tp}(d_j) = \mathsf{L}^k_{A_i}$ we get $d_i[k] \vdash \Pi_i \centerdot F(k)$ and $d_j[0] \vdash F(k), \Pi_j$ by IH(b).

Since $\dfrac{\Pi_0 \;\dots\; \Pi_{i-1} \;\; \Pi_i \centerdot F(k) \;\dots}{\Pi \centerdot F(k)}$ and $\dfrac{\Pi_0 \;\dots\; \Pi_{j-1} \;\; F(k), \Pi_j \;\dots\; \Pi_{j_0} \;\dots}{F(k), \Pi}$ are chain inferences of degree r, we conclude $d\{0\} \vdash \Pi \centerdot F(k)$ and $d\{1\} \vdash F(k), \Pi$. Further $\mathrm{rk}(\mathsf{A}(d)) = \mathrm{rk}(F(k)) < \mathrm{rk}(A_i) \le r$.

Subcase $A_i = \neg A$:

Similar to the previous case, only that now $d_j[0] \vdash \Pi_j{\cdot}A$ and $d_i[0] \vdash A, \Pi_i{\cdot}\bot$, and we apply the chain inferences $\dfrac{\Pi_0 \ \ldots \ \Pi_{j-1} \ \ \Pi_j{\cdot}A \ \ldots}{\Pi{\cdot}A}$ and $\dfrac{\Pi_0 \ \ldots \ \Pi_{i-1} \ \ A, \Pi_i{\cdot}\bot \ \ldots}{A, \Pi}$ to obtain $d\{0\} \vdash \Pi{\cdot}A$ and $d\{1\} \vdash A, \Pi$.

(b) We follow the case distinction of Definition 3.2.:

1. d atomic: cf. §5.

2.-4.: Left to the reader.

5. $d = \mathsf{K}^r_\Pi d_0 \ldots d_l$:

5.1. d critical: Then $\mathsf{tp}(d) = \mathsf{Rep}$ and $d[0] = \mathsf{K}^{r-1}_\Pi d\{0\}d\{1\}$. By (a) we have $d\{0\} \vdash \Pi{\cdot}\mathsf{A}(d)$, $d\{1\} \vdash \mathsf{A}(d), \Pi$, and $\mathrm{rk}(\mathsf{A}(d)) < r$. Hence $d[0] \vdash \Pi$, i.e. $\forall n \in |\mathsf{tp}(d)|(d[n] \vdash \mathsf{tp}(d)(\Pi, n))$.

5.2. d not critical, and i is minimal s.t. $\mathsf{tp}(d_i) \lhd \Pi$:

5.2.1. d_i critical: By IH(a) we have, $d_i\{0\} \vdash \Pi_i{\cdot}\mathsf{A}(d_i)$ and $d_i\{1\} \vdash \mathsf{A}(d_i), \Pi_i$. Further,
$\dfrac{\Pi_0 \ \ldots \ \Pi_{i-1} \ \ \Pi{\cdot}\mathsf{A}(d_i) \quad \mathsf{A}(d_i), \Pi_i \quad \Pi_{i+1} \ \ldots \ \Pi_l}{\Pi}$ is a chain inference of degree $r' := \max\{\mathrm{rk}(\mathsf{A}(d_i)), r\}$.

Hence $d[0] = \mathsf{K}^{r'}_\Pi(i/d_i\{0\}d_i\{1\}) \vdash \Pi$, which yields the claim, since $\mathsf{tp}(d) = \mathsf{Rep}$.

5.2.2. d_i not critical: Then $\mathsf{tp}(d) = \mathsf{tp}(d_i)$, and by IH(b) we have $d_i[n] \vdash \mathsf{tp}(d_i)(\Pi_i, n)$ for all $n \in |\mathsf{tp}(d_i)|$.
Further, $\dfrac{\Pi_0 \ \ldots \ \Pi_{i-1} \ \ \mathsf{tp}(d_i)(\Pi_i, n) \ \ \Pi_{i+1} \ \ldots \ \Pi_l}{\mathsf{tp}(d_i)(\Pi, n)}$ is a chain inference of rank r.

Since $\mathsf{tp}(d) = \mathsf{tp}(d_i)$, we conclude $d[n] = \mathsf{K}^r_{\mathsf{tp}(d)(\Pi,n)}(i/d_i[n]) \vdash \mathsf{tp}(d)(\Pi, n)$ for all $n \in |\mathsf{tp}(d)|$.

Corollary. If $d \vdash \to\bot$, then $d[0] \vdash \to\bot$.

Proof:

From $d \vdash \to\bot$ by Lemma 3.3 we get $\mathsf{tp}(d) \lhd \to\bot$, which implies $\mathsf{tp}(d) = \mathsf{Rep}$. Now by Theorem 3.4b we conclude $d[0] \vdash \mathsf{Rep}(\to\bot, 0)$, i.e. $d[0] \vdash \to\bot$.

§4 Ordinal assignment and termination proof

In this section we will assign to each $\mathcal{Z}$-derivation d an ordinal $o(d) < \varepsilon_0$ and prove that if d is a closed $\mathcal{Z}$-derivation then $o(d[n]) < o(d)$ for all $n \in |\mathsf{tp}(d)|$. The ordinal $o(d)$ will be defined via the auxiliary notions $\mathrm{dg}(d)$ (degree of d) and $\tilde{\mathsf{o}}(d)$ (pre-ordinal of d). [1]

Definition of $\mathrm{dg}(d) < \omega$ **and** $\tilde{\mathsf{o}}(d), o(d) < \varepsilon_0$

For atomic d cf. §5.

Otherwise

$$\mathrm{dg}(d) := \begin{cases} \mathrm{dg}(d_0) & \text{if } d = \mathrm{I}^a_{\forall xF} d_0 \text{ or } d = \mathrm{I}_{\neg A} d_0 \\ \max\{\mathrm{dg}(d_0){-}1, \mathrm{dg}(d_1){-}1, r\} & \text{if } d = \mathrm{Ind}^{a,t}_F d_0 d_1 \text{ with } r := \mathrm{rk}(F) \\ \max\{\mathrm{dg}(d_0){-}1, \ldots, \mathrm{dg}(d_l){-}1, r\} & \text{if } d = \mathsf{K}^r_\Pi d_0 \ldots d_l \end{cases}$$

$$\tilde{\mathsf{o}}(d) := \begin{cases} \tilde{\mathsf{o}}(d_0) + 1 & \text{if } d = \mathrm{I}^a_{\forall xF} d_0 \text{ or } d = \mathrm{I}_{\neg A} d_0 \\ \omega^{\tilde{\mathsf{o}}(d_0)} \# \omega^{\tilde{\mathsf{o}}(d_1)+1} & \text{if } d = \mathrm{Ind}^{a,t}_F d_0 d_1 \\ \omega^{\tilde{\mathsf{o}}(d_0)} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_l)} & \text{if } d = \mathsf{K}^r_\Pi d_0 \ldots d_l \end{cases}$$

$o(d) := \omega_{\mathrm{dg}(d)}(\tilde{\mathsf{o}}(d))$, where $\omega_0(\alpha) := \alpha$, $\omega_{n+1}(\alpha) := \omega^{\omega_n(\alpha)}$.

[1] This ordinal assignment is essentially that of [KB81].

Remark. $\tilde{\mathsf{o}}(d(a/t)) = \tilde{\mathsf{o}}(d)$ and $\mathrm{dg}(d(a/t)) = \mathrm{dg}(d)$.

Lemma 4.1. For each closed $\mathcal{Z}$-derivation d the following holds:

(a) If d is not critical then $\mathrm{dg}(d[n]) \le \mathrm{dg}(d)$ & $\tilde{\mathsf{o}}(d[n]) < \tilde{\mathsf{o}}(d)$, for all $n \in |\mathsf{tp}(d)|$.

(b) If d is critical then:

 (i) $\mathrm{dg}(d\{\nu\}) \le \mathrm{dg}(d)$ & $\tilde{\mathsf{o}}(d\{\nu\}) < \tilde{\mathsf{o}}(d)$, for $\nu = 0, 1$.

 (ii) $\mathrm{dg}(d[0]) < \mathrm{dg}(d)$ & $\tilde{\mathsf{o}}(d[0]) < \omega^{\tilde{\mathsf{o}}(d)}$ & $\mathrm{rk}(\mathsf{A}(d)) < \mathrm{dg}(d)$

Proof by induction on the build-up of d:
Notation: In the following we omit the subscript of K^r_Π.

Assume $d \vdash \Pi$. As before we follow the case distinction of Definition 3.2.

1. d atomic: cf. §5.

2. $d = \mathrm{I}^a_{\forall x F} d_0$: Then $\mathsf{tp}(d) = \mathsf{R}_{\forall x F}$ and $d[n] = d_0(a/n)$.

So we have $\mathrm{dg}(d[n]) = \mathrm{dg}(d_0(a/n)) = \mathrm{dg}(d_0) = \mathrm{dg}(d)$ and $\tilde{\mathsf{o}}(d[n]) = \tilde{\mathsf{o}}(d_0(a/n)) = \tilde{\mathsf{o}}(d_0) < \tilde{\mathsf{o}}(d)$.

3. $d = \mathrm{I}_{\neg A} d_0$: similar to 2.

4. $d = \mathrm{Ind}^{a,k}_F d_0 d_1$: Then $\mathsf{tp}(d) = \mathsf{Rep}$ and $d[0] = \mathsf{K}^r d_0\, d_1(a/0) \ldots d_1(a/k{-}1)$, where $r = \mathrm{rk}(F)$.

So we have $\mathrm{dg}(d[0]) \le \max\{\mathrm{dg}(d_0){-}1, \mathrm{dg}(d_1){-}1, r\} = \mathrm{dg}(d)$ and
$\tilde{\mathsf{o}}(d[0]) = \omega^{\tilde{\mathsf{o}}(d_0)} \# \omega^{\tilde{\mathsf{o}}(d_1)}\cdot k < \omega^{\tilde{\mathsf{o}}(d_0)} \# \omega^{\tilde{\mathsf{o}}(d_1)+1} = \tilde{\mathsf{o}}(d)$.

5. $d = \mathsf{K}^r d_0 \ldots d_l$:

5.1. d critical: Then $\mathsf{tp}(d) = \mathsf{Rep}$ and $d[0] = \mathsf{K}^{r-1} d\{0\} d\{1\}$ where
either $d\{0\} = \mathsf{K}^r(i/d_i[k])$ & $d\{1\} = \mathsf{K}^r(j/d_j[0])$ or $d\{0\} = \mathsf{K}^r(j/d_j[0])$ & $d\{1\} = \mathsf{K}^r(i/d_i[0])$.

By IH(a), $\mathrm{dg}(d_i[k]) \le \mathrm{dg}(d_i)$ & $\tilde{\mathsf{o}}(d_i[k]) < \tilde{\mathsf{o}}(d_i)$ and $\mathrm{dg}(d_j[0]) \le \mathrm{dg}(d_j)$ & $\tilde{\mathsf{o}}(d_j[0]) < \tilde{\mathsf{o}}(d_j)$.

This yields $\mathrm{dg}(d\{\nu\}) \le \mathrm{dg}(d)$ & $\tilde{\mathsf{o}}(d\{\nu\}) < \tilde{\mathsf{o}}(d)$ for $\nu = 0, 1$.

Hence $\mathrm{dg}(d[0]) = \max\{\mathrm{dg}(d\{0\}){-}1, \mathrm{dg}(d\{1\}){-}1, r{-}1\} < \mathrm{dg}(d)$ and $\tilde{\mathsf{o}}(d[0]) = \omega^{\tilde{\mathsf{o}}(d\{0\})} \# \omega^{\tilde{\mathsf{o}}(d\{1\})} < \omega^{\tilde{\mathsf{o}}(d)}$.

By Theorem 3.4a we have $\mathrm{rk}(\mathsf{A}(d)) < r$, thence $\mathrm{rk}(\mathsf{A}(d)) < \mathrm{dg}(d)$.

5.2. d not critical, and i is minimal s.t. $\mathsf{tp}(d_i) \triangleleft \Pi$:

5.2.1. d_i critical: Then $\mathsf{tp}(d) = \mathsf{Rep}$ and $d[0] = \mathsf{K}^{r'}(i/d_i\{0\}d_i\{1\})$ with $r' = \max\{\mathrm{rk}(\mathsf{A}(d_i)), r\}$.

By IH(b) we have $\mathrm{dg}(d_i\{\nu\}) \le \mathrm{dg}(d_i)$ & $\tilde{\mathsf{o}}(d_i\{\nu\}) < \tilde{\mathsf{o}}(d_i)$ for $\nu = 0, 1$, and also $\mathrm{rk}(\mathsf{A}(d_i)) < \mathrm{dg}(d_i)$.
The latter yields $r' \le \max\{\mathrm{dg}(d_i){-}1, r\} \le \mathrm{dg}(d)$. Hence

$\mathrm{dg}(d[0]) = \max\{\mathrm{dg}(d_0){-}1, \ldots, \mathrm{dg}(d_i\{0\}){-}1, \mathrm{dg}(d_i\{1\}){-}1, \ldots, \mathrm{dg}(d_l){-}1, r'\} \le$
$\le \max\{\mathrm{dg}(d_0){-}1, \ldots, \mathrm{dg}(d_i){-}1, \ldots, \mathrm{dg}(d_l){-}1, r'\} \le \mathrm{dg}(d)$ and
$\tilde{\mathsf{o}}(d[0]) = \omega^{\tilde{\mathsf{o}}(d_0)} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_i\{0\})} \# \omega^{\tilde{\mathsf{o}}(d_i\{1\})} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_l)} < \omega^{\tilde{\mathsf{o}}(d_0)} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_i)} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_l)} = \tilde{\mathsf{o}}(d)$.

5.2.2. d_i not critical: Then $\mathsf{tp}(d) = \mathsf{tp}(d_i)$ and $d[n] = \mathsf{K}^r(i/d_i[n])$.

By IH(a), $\mathrm{dg}(d_i[n]) \le \mathrm{dg}(d_i)$ and $\tilde{\mathsf{o}}(d_i[n]) < \tilde{\mathsf{o}}(d_i)$.
Hence $\mathrm{dg}(d[n]) = \max\{\mathrm{dg}(d_0){-}1, \ldots, \mathrm{dg}(d_i[n]){-}1, \ldots, \mathrm{dg}(d_l){-}1, r\} \le \mathrm{dg}(d)$ and
$\tilde{\mathsf{o}}(d[n]) = \omega^{\tilde{\mathsf{o}}(d_0)} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_i[n])} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_l)} < \omega^{\tilde{\mathsf{o}}(d_0)} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_i)} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_l)} = \tilde{\mathsf{o}}(d)$.

Theorem 4.2.
If d is a closed $\mathcal{Z}$-derivation, then $o(d[n]) < o(d)$ for all $n \in |\mathsf{tp}(d)|$.

Proof :
By Lemma 4.1 we have $\tilde{\mathsf{o}}(d[n]) < \omega_{\mathrm{dg}(d)-\mathrm{dg}(d[n])}(\tilde{\mathsf{o}}(d))$ and thus
$o(d[n]) = \omega_{\mathrm{dg}(d[n])}(\tilde{\mathsf{o}}(d[n])) < \omega_{\mathrm{dg}(d)}(\tilde{\mathsf{o}}(d)) = o(d)$.

§5 Treatment of atomic derivations

At several places in the preceding sections we had postponed the treatment of atomic derivations. This will now be caught up.

The *logical axioms of* $\mathcal{Z}$ are all sequents of the following kinds:

- $\Gamma \to A$ with $A \in \Gamma$
- $\Gamma \to F(t)$ with $\forall x F(x) \in \Gamma$
- $\Gamma \to \bot$ with $A, \neg A \in \Gamma$
- $\Gamma \to A$ with A atomic and $\neg\neg A \in \Gamma$

The *mathematical axioms of* $\mathcal{Z}$ are given by a set of sequents $\mathrm{Ax}(\mathcal{Z})$ satisfying the following conditions:

- $\Pi \in \mathrm{Ax}(\mathcal{Z}) \Rightarrow \Pi(a/t) \in \mathrm{Ax}(\mathcal{Z})$ and $A, \Pi \in \mathrm{Ax}(\mathcal{Z})$.
- $\mathrm{FV}(\Pi) = \emptyset \Rightarrow (\, \Pi \in \mathrm{Ax}(\mathcal{Z}) \Leftrightarrow \Pi$ has endform$\,)$.

Definition of the atomic $\mathcal{Z}$-derivations

0. If $\Pi \in \mathrm{Ax}(\mathcal{Z})$, then $\mathsf{Ax}^0_\Pi \vdash \Pi$.

1. If $\Pi = \Gamma \to C$ with $C \in \Gamma$ then $\mathsf{Ax}^1_\Pi \vdash \Pi$.

2.1. If $\Pi = \Gamma \to F(t)$ with $\forall x F \in \Gamma$ then $\mathsf{Ax}^{\forall x F, t}_\Pi \vdash \Pi$.

2.2. If $\Pi = \Gamma \to \bot$ with $\neg A, A \in \Gamma$ then $\mathsf{Ax}^{\neg A, 0}_\Pi \vdash \Pi$.

3. If $\Pi = \Gamma \to A$ with $\mathrm{rk}(A) = 0$ & $\neg\neg A \in \Gamma$ then $\mathsf{Ax}^{\neg\neg}_\Pi \vdash \Pi$.

Definition of $\mathsf{tp}(d)$ and $d[n]$ for closed atomic $\mathcal{Z}$-derivations d

0. $d = \mathsf{Ax}^0_{\Gamma \to C}$: Then $\Gamma \to C$ has endform, and we set
$$\mathsf{tp}(d) := \begin{cases} \mathsf{R}_C & \text{if } C \approx \top \\ \mathsf{L}^0_A & \text{if } C \not\approx \top \text{ and } A \text{ is the first formula in } \Gamma \text{ s.t. } A \approx \bot \end{cases}$$

1. $d = \mathsf{Ax}^1_{\Gamma \to C}$ with $C \in \Gamma$:

1.1. $\mathrm{rk}(C) = 0$: Then $\mathsf{tp}(d) := \begin{cases} \mathsf{R}_C & \text{if } C \approx \top \\ \mathsf{L}^0_C & \text{if } C \approx \bot \end{cases}$

1.2. $\mathrm{rk}(C) > 0$: Then $\mathsf{tp}(d) := \mathsf{R}_C$ and $d[n] := \mathsf{Ax}^{C,n}_{\mathsf{tp}(d)(\Pi,n)}$.

2. $d = \mathsf{Ax}^{C,k}_\Pi$: Then $\mathsf{tp}(d) := \mathsf{L}^k_C$ and $d[0] := \mathsf{Ax}^1_{\mathsf{tp}(d)(\Pi,0)}$

3. $d = \mathsf{Ax}^{\neg\neg}_{\Gamma \to A}$:

3.1. $A \approx \top$: Then $\mathsf{tp}(d) := \mathsf{R}_A$.

3.2. $A \approx \bot$: Then $\mathsf{tp}(d) := \mathsf{L}^0_{\neg\neg A}$ and $d[0] := \mathrm{I}_{\neg A}\mathsf{Ax}^0_{A,\Gamma \to \bot}$.

Lemma 5.1.
If $d \vdash \Pi$ with $\mathrm{FV}(\Pi) = \emptyset$ and d atomic, then:
(a) $\mathsf{tp}(d) \lhd \Pi$.
(b) $d[n] \vdash \mathsf{tp}(d)(\Pi, n)$ for all $n \in |\mathsf{tp}(d)|$.

Proof :
(a) Left to the reader.

(b) Abbreviation: $\Pi' := \mathsf{tp}(d)(\Pi, n)$.

1.2. $d = \mathsf{Ax}^1_\Pi$ with $\Pi = \Gamma \to C$ and $C \in \Gamma$ & $\mathrm{rk}(C) > 0$:
Then $\mathsf{tp}(d) = \mathsf{R}_C$ and $\Pi' = \begin{cases} \Gamma \to F(n) & \text{if } C = \forall x F(x) \\ A, \Gamma \to \bot & \text{if } C = \neg A \end{cases}$. Hence $d[n] = \mathsf{Ax}^{C,n}_{\Pi'} \vdash \Pi'$.

2.1. $d = \mathsf{Ax}_{\Pi}^{\forall xF,k}$ with $\Pi = \Gamma{\to}F(k)$: Then $\mathsf{tp}(d) = \mathsf{L}^k_{\forall xF}$ and $\Pi' = F(k), \Gamma{\to}F(k)$. Hence $d[0] = \mathsf{Ax}^1_{\Pi'} \vdash \Pi'$.

2.2. $d = \mathsf{Ax}_{\Pi}^{\neg A,0}$ with $\Pi = \Gamma{\to}\bot$ and $A, \neg A \in \Gamma$: Then $\mathsf{tp}(d) = \mathsf{L}^0_{\neg A}$ and $\Pi' = \Gamma{\to}A$. Hence $d[0] = \mathsf{Ax}^1_{\Pi'} \vdash \Pi'$.

3.2. $d = \mathsf{Ax}_{\Pi}^{\neg\neg}$ with $\Pi = \Gamma{\to}A$, $A \approx \bot$, and $\neg\neg A \in \Gamma$:
Then $d' := \mathsf{Ax}^0_{A,\Gamma\to\bot} \vdash A, \Gamma{\to}\bot$ and thus $d[0] = \mathrm{I}_{\neg A} d' \vdash \Gamma{\to}\neg A$.
Further, $\Pi' = \mathsf{L}^0_{\neg\neg A}(\Pi, n) = \Gamma{\to}\neg A$.

Definition of $\mathrm{dg}(d), \tilde{\mathsf{o}}(d), o(d)$ **for atomic** $\mathcal{Z}$**-derivations** d

$\mathrm{dg}(d) := 0$ and $o(d) := \omega_{\mathrm{dg}(d)}(\tilde{\mathsf{o}}(d)) = \tilde{\mathsf{o}}(d)$, where
$\tilde{\mathsf{o}}(\mathsf{Ax}^0_{\Pi}) := 0$, $\tilde{\mathsf{o}}(\mathsf{Ax}^1_{\Gamma\to C}) := 2\mathrm{rk}(C)$, $\tilde{\mathsf{o}}(\mathsf{Ax}^{C,t}_{\Pi}) := 2\mathrm{rk}(C) - 1$, $\tilde{\mathsf{o}}(\mathsf{Ax}^{\neg\neg}_{\Pi}) := 2$.

Lemma 5.2. If d is a closed atomic $\mathcal{Z}$-derivation, then $o(d[n]) < o(d)$ for all $n \in |\mathsf{tp}(d)|$.

Proof: Left to the reader.

§6 Embedding of $\mathcal{Z}$ into an infinitary system $\mathcal{Z}^\infty$

In this section we give an interpretation of the finitary system $\mathcal{Z}$ in an infinitary system $\mathcal{Z}^\infty$ of ω-arithmetic. This way we obtain an explanation of the reduction steps on $\mathcal{Z}$-derivations and the assignment of ordinals to $\mathcal{Z}$-derivations introduced in §§3-5.

Derivable objects of $\mathcal{Z}^\infty$ are closed sequents $\Pi = \Gamma{\to}C$.

The inference symbols of $\mathcal{Z}^\infty$ are:

R_A with $\mathrm{rk}(A) > 0$ or $A \approx \top$, L^k_A with $\mathrm{rk}(A) > 0$ or $A \approx \bot$, and Cut_D for arbitrary sentences D.

We set $\mathsf{Cut}_D \lhd \Pi$ for each Π, $|\mathsf{Cut}_D| := \{0, 1\}$, $\mathsf{Cut}_D(\Pi, 0) := \Pi{\cdot}D$ and $\mathsf{Cut}_D(\Pi, 1) := D, \Pi$.

$$\mathrm{rk}(\mathcal{I}) := \begin{cases} \mathrm{rk}(D) & \text{if } \mathcal{I} = \mathsf{Cut}_D \\ -1 & \text{otherwise} \end{cases}$$

The following definition introduces the relation $\mathfrak{d} \vdash^\alpha_m \Pi$ which is short for
"*$\mathfrak{d}$ is a $\mathcal{Z}^\infty$-derivation of* Π *with ordinal height* $\le \alpha$ *and cutrank* $\le m$".

Inductive Definition of $\mathfrak{d} \vdash^\alpha_m \Pi$

If $\mathcal{I}$ is an inference symbol of $\mathcal{Z}^\infty$ with $\mathrm{rk}(\mathcal{I}) < m$, and if
$\mathcal{I} \lhd \Pi \;\&\; \forall n \in |\mathcal{I}| \exists \alpha_n < \alpha(\, \mathfrak{d}_n \vdash^{\alpha_n}_m \mathcal{I}(\Pi, n)\,)$, then $\mathcal{I}\big(\mathfrak{d}_n\big)_{n\in|\mathcal{I}|} \vdash^\alpha_m \Pi$.

Definition of $\mathsf{last}(\mathfrak{d})$**:** If $\mathfrak{d} = \mathcal{I}\big(\mathfrak{d}_n\big)_{n\in|\mathcal{I}|}$, then $\mathsf{last}(\mathfrak{d}) := \mathcal{I}$.

Remark. If $\mathfrak{d} \vdash^\alpha_m \Pi$ then $\mathsf{last}(\mathfrak{d}) \lhd \Pi$.

Theorem and Definition 6.1.

If $\dfrac{\Pi_0 \;\ldots\; \Pi_l}{\Pi}$ is a chain inference of rank $r \le m$, and if $\mathfrak{d}_i \vdash^{\alpha_i}_{m+1} \Pi_i$ for $i = 0, \ldots, l$,

then there exists a $\mathcal{Z}^\infty$-derivation $\mathfrak{d} = \mathcal{K}^r_\Pi(\mathfrak{d}_0, \ldots, \mathfrak{d}_l) \vdash^\alpha_m \Pi$ with $\alpha := \omega^{\alpha_0} \# \ldots \# \omega^{\alpha_l}$.

Proof by induction on α:
Assume $\Pi = \Gamma{\to}C$ and $\Pi_i = \Gamma_i{\to}A_i$, and let j_0 be minimal such that
$A_{j_0} \in \{C, \bot\}$ & $\forall i \le j_0(\, \Gamma_i \subseteq \Gamma, A_0, \ldots, A_{i-1}\,)$.

1. $\forall i \le j_0(\mathsf{last}(\mathfrak{d}_i) \not\lhd \Pi)$. By Lemma 3.1 there is the least pair (i, j) such that $i < j \le j_0$, $\mathsf{last}(\mathfrak{d}_j) = \mathsf{L}^k_{A_i}$ (for some k), $\mathsf{last}(\mathfrak{d}_i) = \mathsf{R}_{A_i}$, and $0 < \mathrm{rk}(A_i) \le r$. Then $\mathfrak{d}_i = \mathsf{R}_{A_i}(\mathfrak{d}_{in})_n$ and $\mathfrak{d}_j = \mathsf{L}^k_{A_i}\mathfrak{d}_{j0}$.

Let $\mathfrak{d} := \mathsf{Cut}_D(\mathfrak{e}_0, \mathfrak{e}_1)$ with $D := \begin{cases} F(k) & \text{if } A_i = \forall x F \\ A & \text{if } A_i = \neg A \end{cases}$, and

$\mathfrak{e}_0 := \mathcal{K}^r_{\Pi\text{•}D} \begin{cases} (i/\mathfrak{d}_{ik}) & \text{if } A_i = \forall x F \\ (j/\mathfrak{d}_{j0}) & \text{if } A_i = \neg A \end{cases}$ and $\mathfrak{e}_1 := \mathcal{K}^r_{D,\Pi} \begin{cases} (j/\mathfrak{d}_{j0}) & \text{if } A_i = \forall x F \\ (i/\mathfrak{d}_{i0}) & \text{if } A_i = \neg A \end{cases}$.

The IH yields $\mathfrak{e}_0 \vdash^{\alpha'}_m \Pi\text{•}D$ and $\mathfrak{e}_1 \vdash^{\alpha''}_m D, \Pi$ with $\alpha', \alpha'' < \alpha$.
Since $\mathrm{rk}(D) < \mathrm{rk}(A_i) \le r \le m$, it follows that $\mathfrak{d} \vdash^{\alpha}_m \Pi$

2. Otherwise: Let $i \le j_0$ be minimal such that $\mathsf{last}(\mathfrak{d}_i) \lhd \Pi$, and let $\mathcal{I} := \mathsf{last}(d_i)$.

2.1. $\mathcal{I} = \mathsf{Cut}_D$: Then $\mathfrak{d}_i = \mathsf{Cut}_D(\mathfrak{d}_{i0}, \mathfrak{d}_{i1})$ with
$\mathfrak{d}_{i0} \vdash^{\alpha_{i0}}_{m+1} \Pi_i\text{•}D$ & $\mathfrak{d}_{i1} \vdash^{\alpha_{i1}}_{m+1} D, \Pi_i$ & $\alpha_{i0}, \alpha_{i1} < \alpha$ & $\mathrm{rk}(D) \le m$.
We set $\mathfrak{d} := \mathcal{K}^{r'}_{\Pi}(\mathfrak{d}_0, \ldots, \mathfrak{d}_{i-1}, \mathfrak{d}_{i0}, \mathfrak{d}_{i1}, \mathfrak{d}_{i+1}, \ldots, \mathfrak{d}_l)$ with $r' := \max\{\mathrm{rk}(D), r\} \le m$.
From $\mathfrak{d}_{i0} \vdash^{\alpha_{i0}}_{m+1} \Pi_i\text{•}D$ & $\mathfrak{d}_{i1} \vdash^{\alpha_{i1}}_{m+1} D, \Pi_i$ & $\alpha_{i0}, \alpha_{i1} < \alpha_i$ and $\mathfrak{d}_\nu \vdash^{\alpha_\nu}_{m+1} \Pi_\nu$ for $\nu \in \{0, \ldots, l\} \setminus \{i\}$ by IH we obtain $\mathfrak{d} \vdash^{\beta}_m \Pi$ with $\beta := \omega^{\alpha_0} \# \ldots \# \omega^{\alpha_{i-1}} \# \omega^{\alpha_{i0}} \# \omega^{\alpha_{i1}} \# \omega^{\alpha_{i+1}} \# \ldots \# \omega^{\alpha_l} < \alpha$.

2.2. $\mathcal{I} \notin \mathsf{Cut}$: Then $\mathfrak{d} := \mathcal{I}\big(\mathcal{K}^r_{\mathcal{I}(\Pi,n)}(i/\mathfrak{d}_{in})\big)_{n \in |\mathcal{I}|}$, where $\mathfrak{d}_i = \mathcal{I}(\mathfrak{d}_{in})_{n \in |\mathcal{I}|}$.

Abbreviation. $\mathcal{Z}^\infty \vdash^\alpha_m \Pi \;:\Leftrightarrow\; \exists \mathfrak{d}$ such that $\mathfrak{d} \vdash^\alpha_m \Pi$.

Corollary 6.2. $\mathcal{Z}^\infty \vdash^\alpha_{m+1} \Pi \;\Rightarrow\; \mathcal{Z}^\infty \vdash^{\omega^\alpha}_m \Pi$.

Proof: This follows from Theorem 6.1 for $l = 0$.

Having the operations $\mathcal{K}^r_\Pi$ at hand it is now easy to embed $\mathcal{Z}$ into the infinitary system $\mathcal{Z}^\infty$.

Definition of a $\mathcal{Z}^\infty$-derivation d^∞ for each closed $\mathcal{Z}$-derivation d

1. For atomic d we define $d^\infty := \mathsf{tp}(d)\big(d[n]^\infty\big)_{n \in |\mathsf{tp}(d)|}$ by recursion on $o(d) < \omega$.
   Especially, in case $d = \mathsf{Ax}^{\neg\neg}_{\Gamma \to A}$ with $A \approx \bot$ we have
   $d^\infty = \mathsf{L}^0_{\neg\neg A} d[0]^\infty = \mathsf{L}^0_{\neg\neg A}(\mathrm{I}_{\neg A} \mathsf{Ax}^0_{A,\Gamma\to\bot})^\infty = \mathsf{L}^0_{\neg\neg A} \mathsf{R}_{\neg A} \mathsf{tp}(\mathsf{Ax}^0_{A,\Gamma\to\bot}) = \mathsf{L}^0_{\neg\neg A} \mathsf{R}_{\neg A} \mathsf{L}^0_A$.
2. $(\mathrm{I}^a_{\forall x F} d_0)^\infty := \mathsf{R}_{\forall x F}\big(d_0(a/n)^\infty\big)_{n \in \mathbb{N}}$
3. $(\mathrm{I}_{\neg A} d_0)^\infty := \mathsf{R}_{\neg A} d_0^\infty$
4. $(\mathrm{Ind}^{a,k}_F d_0 d_1)^\infty := \mathcal{K}^r_{\Gamma \to F(k)}(d_0^\infty, d_1(a/0)^\infty, \ldots, d_1(a/k{-}1)^\infty)$
5. $(\mathsf{K}^r_\Pi d_0 \ldots d_l)^\infty := \mathcal{K}^r_\Pi(d_0^\infty, \ldots, d_l^\infty)$

Theorem 6.3. If $d \vdash \Pi$ and $\mathrm{FV}(\Pi) = \emptyset$, then $d^\infty \vdash^{\tilde{\mathsf{o}}(d)}_{\mathrm{dg}(d)} \Pi$.

Proof by induction on the build-up of d using Theorem 6.1:
Assume $\Pi = \Gamma \to C$.

1. d atomic: Left to the reader.

2. $d = \mathrm{I}^a_{\forall x F} d_0$: Then $C = \forall x F$ and $d_0(n) \vdash \Gamma \to F(n)$. By IH, $d_0(n)^\infty \vdash^{\tilde{\mathsf{o}}(d_0)}_{\mathrm{dg}(d_0)} \Gamma \to F(n)$ $(\forall n)$.
Hence $d^\infty = \mathsf{R}_{\forall x F}(d_0(n)^\infty)_{n \in \mathbb{N}} \vdash^{\tilde{\mathsf{o}}(d)}_{\mathrm{dg}(d)} \Pi$.

3. $d = \mathrm{I}_{\neg A} d_0$: Similar to 2.

4. $d = \mathrm{Ind}^{a,k}_F d_0 d_1$ with $d_0 \vdash \Gamma \to F(0)$, $d_1 \vdash F(a), \Gamma \to F(\mathsf{S}a)$, and $\Pi = \Gamma \to F(k)$:
By IH, $d_0^\infty \vdash^{\tilde{\mathsf{o}}(d_0)}_{\mathrm{dg}(d_0)} \Gamma \to F(0)$ and $d_1(a/n)^\infty \vdash^{\tilde{\mathsf{o}}(d_1)}_{\mathrm{dg}(d_1)} F(n), \Gamma \to F(\mathsf{S}n)$ $(\forall n)$.
From this by Theorem 6.1 we obtain $d^\infty = \mathcal{K}^r_{\Gamma \to F(k)}(d_0^\infty, d_1(a/0)^\infty, \ldots, d_1(a/k{-}1)^\infty) \vdash^{\tilde{\mathsf{o}}(d)}_{\mathrm{dg}(d)} \Gamma \to F(k)$, since $r \le \mathrm{dg}(d)$ and $\mathrm{dg}(d_0), \mathrm{dg}(d_1) \le \mathrm{dg}(d){+}1$ and $\omega^{\tilde{\mathsf{o}}(d_0)} \# \omega^{\tilde{\mathsf{o}}(d_1)} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_1)} < \omega^{\tilde{\mathsf{o}}(d)}$.

5. $d = \mathsf{K}^r_\Pi d_0 \ldots d_l$ with $d_i \vdash \Pi_i$ $(i = 0, \ldots, l)$:

Note that $\mathrm{dg}(d) = \max\{\mathrm{dg}(d_0)-1, \ldots, \mathrm{dg}(d_l)-1, r\}$ and therefore (1) $\mathrm{dg}(d_i) \le \mathrm{dg}(d)+1$, (2) $r \le \mathrm{dg}(d)$. By IH we have $d_i^\infty \vdash^{\tilde{\mathsf{o}}(d_i)}_{\mathrm{dg}(d_i)} \Pi_i$ and therefore, by (1), $d_i^\infty \vdash^{\tilde{\mathsf{o}}(d_i)}_{\mathrm{dg}(d)+1} \Pi_i$ $(i = 0, \ldots, l)$. From this by (2) and Theorem 6.1 we conclude $d^\infty = \mathcal{K}^r_\Pi(d_0^\infty, \ldots, d_l^\infty) \vdash^{\alpha}_{\mathrm{dg}(d)} \Pi$ with $\alpha = \omega^{\tilde{\mathsf{o}}(d_0)} \# \ldots \# \omega^{\tilde{\mathsf{o}}(d_l)} = \tilde{\mathsf{o}}(d)$.

Corollary 6.4. If $d \vdash \Pi$ and $\mathrm{FV}(\Pi) = \emptyset$, then $\mathcal{Z}^\infty \vdash^{o(d)}_0 \Pi$.

Theorem 6.5.

(i) If $\mathsf{tp}(d) = \mathsf{Rep}$, then $d^\infty = \begin{cases} \mathsf{Cut}_{\mathsf{A}(d)}(d\{0\}^\infty, d\{1\}^\infty) & \text{if } d \text{ critical} \\ d[0]^\infty & \text{otherwise} \end{cases}$

(ii) If $\mathcal{I} := \mathsf{tp}(d) \ne \mathsf{Rep}$, then $d^\infty = \mathcal{I}\big(d[n]^\infty\big)_{n \in |\mathcal{I}|}$

Proof by induction over the build-up of d, comparing definitions 3.2 and 6.1.

§7 Multisuccedent sequents

The approach of §§3,4 can easily be adapted to calculi with multisuccedent sequents by generalizing the chain rule as follows:[1]

(GCR) The figure $\dfrac{\Pi_0 \;\ldots\; \Pi_l}{\Pi}$ is called a *(generalized) chain rule inference of rank* r if Π can be derived from (weakenings of) the sequents $\Pi_0, \ldots, \Pi_l$ by a finite number of cuts of rank $\le r$.

By adding this rule to the proof system of [Ge38] and taking the ordinal assignment from §4 of the present paper a certain simplification of [Ge38] can be achieved, especially the somewhat unpleasent concept of "Höhenlinie" can be avoided.

In the following we review the essential concepts of §§3,4 in a kind of axiomatic presentation, thereby adjusting everthing to the multisuccedent context. The main ingredient here is Lemma 7.1 which replaces Lemma 3.1. The above rule (GCR) will be captured by the inductively defined relation "$(\Pi_0, \ldots, \Pi_l) \Vdash_r \Pi$".

Definitions.

A *sequent* is an expression $\Gamma{\to}\Delta$ where Γ and Δ are finite (possibly empty) sequences of formulas.

For $\Pi = \Gamma{\to}\Delta$ we set
$\mathsf{L}(\Pi) := \Gamma$ and $\mathsf{R}(\Pi) := \Delta$;
$A, \Pi := A, \Gamma \to \Delta$ and $\Pi, A := \Gamma \to \Delta, A$.

Inference symbols R_A, L^k_A, Rep and their *arities* are the same as in §3.

For each inference symbol $\mathcal{I}$, sequent Π, and $n \in |\mathcal{I}|$ the sequent $\mathcal{I}(\Pi, n)$ is defined by

$$\mathcal{I}(\Pi, n) := \begin{cases} \Pi, F(n) & \text{if } \mathcal{I} = \mathsf{R}_{\forall x F} \\ F(k), \Pi & \text{if } \mathcal{I} = \mathsf{L}^k_{\forall x F} \\ A, \Pi & \text{if } \mathcal{I} = \mathsf{R}_{\neg A} \\ \Pi, A & \text{if } \mathcal{I} = \mathsf{L}^0_{\neg A} \\ \Pi & \text{otherwise} \end{cases}$$

The relation $\mathcal{I} \lhd \Pi$ is defined by:
$\mathsf{R}_A \lhd \Pi \;:\Leftrightarrow\; A \in \mathsf{R}(\Pi)$,
$\mathsf{L}^k_A \lhd \Pi \;:\Leftrightarrow\; A \in \mathsf{L}(\Pi)$,
$\mathsf{Rep} \lhd \Pi \;:\Leftrightarrow\; 0 = 0$.

[1] A similar rule is used in [KB81].

Abbreviation. $\Pi \subseteq \Pi' \;:\Leftrightarrow\; \mathsf{L}(\Pi) \subseteq \mathsf{L}(\Pi') \;\&\; \mathsf{R}(\Pi) \subseteq \mathsf{R}(\Pi')$.

Inductive Definition of $(\Pi_0, \dots, \Pi_l) \Vdash_r \Pi$
Let $\vec{\Pi} := (\Pi_0, \dots, \Pi_l)$.
1. If $\Pi_i \subseteq \Pi$ for some $i \le l$, then $\vec{\Pi} \Vdash_r \Pi$.
2. If $\vec{\Pi} \Vdash_r \Pi, C$ and $\vec{\Pi} \Vdash_r C, \Pi$ with $\mathrm{rk}(C) \le r$, then $\vec{\Pi} \Vdash_r \Pi$.

Lemma 7.1. ("Existence of a suitable cut")
If $\vec{\Pi} = (\Pi_0, \dots, \Pi_l) \Vdash_r \Pi$ and $\forall i \le l(\mathcal{I}_i \triangleleft \Pi_i \;\&\; \mathcal{I}_i \ntriangleleft \Pi)$, then there are $i, j \le l$ such that $\mathcal{I}_i = \mathsf{R}_B \;\&\; \mathcal{I}_j = \mathsf{L}_B^k \;\&\; \mathrm{rk}(B) \le r$ for some B, k.

Proof by induction over the definition of $\vec{\Pi} \Vdash_r \Pi$:
From the second premise we conclude $\forall i \le l(\Pi_i \not\subseteq \Pi)$. Together with $\vec{\Pi} \Vdash_r \Pi$ this implies that there exists a C of rank $\le r$ such that $\vec{\Pi} \Vdash_r \Pi, C$ and $\vec{\Pi} \Vdash_r C, \Pi$.

Case 1: $\forall i \le l(\mathcal{I}_i \ntriangleleft \Pi, C)$ or $\forall i \le l(\mathcal{I}_i \ntriangleleft C, \Pi)$. Then the claim follows immediately from the IH.

Case 2: Otherwise. Then there exist $i, j \le l$ such that $\mathcal{I}_i \triangleleft \Pi, C$ and $\mathcal{I}_j \triangleleft C, \Pi$.

From $\mathcal{I}_i \triangleleft \Pi, C \;\&\; \mathcal{I}_i \ntriangleleft \Pi$ it follows that $\mathcal{I}_i = \mathsf{R}_C$.

From $\mathcal{I}_j \triangleleft C, \Pi \;\&\; \mathcal{I}_j \ntriangleleft \Pi$ it follows that $\mathcal{I}_j = \mathsf{L}_C^k$ for some k.

Assumption 0.

$\mathcal{D}$ is a set of (derivation) terms, and to each $d \in \mathcal{D}$ there is assigned a sequent $\mathsf{End}(d)$, an inference symbol $\mathsf{tp}(d)$, and, for each $n \in |\mathsf{tp}(d)|$, a term $d[n] \in \mathcal{D}$.

Abbreviation. $d \vdash \Pi \;:\Leftrightarrow\; d \in \mathcal{D} \;\&\; \mathsf{End}(d) = \Pi$

Assumption 1.

If $(\Pi_0, \dots, \Pi_l) \Vdash_r \Pi$ and $d_0 \vdash \Pi_0, \dots, d_l \vdash \Pi_l$ then $\mathsf{K}_\Pi^r d_0 \dots d_l \vdash \Pi$.

Definitions. Assume $d = \mathsf{K}_\Pi^r d_0 \dots d_l \vdash \Pi$ with $d_i \vdash \Pi_i$ and $\mathsf{tp}(d_i) \triangleleft \Pi_i$ for all $i \le l$.

- d is critical $:\Leftrightarrow\; \forall i \le l(\mathsf{tp}(d_i) \ntriangleleft \Pi)$
- If d is critical we take the least pair (i, j) such that
  $i, j \le l \;\&\; \mathsf{tp}(d_i) = \mathsf{R}_B \;\&\; \mathsf{tp}(d_j) = \mathsf{L}_B^k \;\&\; \mathrm{rk}(B) \le r$ for some B, k
  (which exists according to Lemma 7.1), and define
  $\mathsf{A}(d) := \begin{cases} F(k) & \text{if } B = \forall x F(x) \\ A & \text{if } B = \neg A \end{cases}$
  $d\{0\} := \mathsf{K}_{\Pi, \mathsf{A}(d)}^r \begin{cases} (i/d_i[k]) & \text{if } B = \forall x F \\ (j/d_j[0]) & \text{if } B = \neg A \end{cases} \qquad d\{1\} := \mathsf{K}_{\mathsf{A}(d), \Pi}^r \begin{cases} (j/d_j[0]) & \text{if } B = \forall x F \\ (i/d_i[0]) & \text{if } B = \neg A \end{cases}$

Assumption 2:
If $d = \mathsf{K}_\Pi^r d_0 \dots d_l \vdash \Pi$ with $d_i \vdash \Pi_i$ and $\mathsf{tp}(d_i) \triangleleft \Pi_i$ for all $i \le l$, then the following holds

(a) If d is critical, then $\mathsf{tp}(d) = \mathsf{Rep}$ and $d[0] = \mathcal{K}_\Pi^{r-1} d\{0\} d\{1\}$

(b) If d is not critical and $i \le l$ is the least number s.t. $\mathsf{tp}(d_i) \triangleleft \Pi$, then

$$\mathsf{tp}(d) = \begin{cases} \mathsf{Rep} & \text{if } d_i \text{ critical} \\ \mathsf{tp}(d_i) & \text{otherwise} \end{cases}$$

$$d[n] = \begin{cases} \mathsf{K}_\Pi^{r'}(i/d_i\{0\}d_i\{1\}) \text{ with } r' := \max\{\mathrm{rk}(\mathsf{A}(d_i)), r\} & \text{if } d_i \text{ critical} \\ \mathsf{K}_{\mathsf{tp}(d)(\Pi, n)}^r(i/d_i[n]) & \text{otherwise} \end{cases}$$

Assumption 3:

There are mappings $\mathrm{dg} : \mathcal{D} \to \omega$ and $\tilde{\mathsf{o}} : \mathcal{D} \to On$ such that such that for each $d = \mathsf{K}^r_\Pi d_0 \dots d_l$ we have $\mathrm{dg}(d) = \max\{\mathrm{dg}(d_0){-}1, \dots, \mathrm{dg}(d_l){-}1, r\}$, and $\tilde{\mathsf{o}}(d) = \omega^{\tilde{\mathsf{o}}(d_0)} \# \dots \# \omega^{\tilde{\mathsf{o}}(d_l)}$.

Abbreviations.

For $d \in \mathcal{D}$ and $\Pi := \mathsf{End}(d)$ we set:

$d \in \mathcal{D}_1 \;:\Leftrightarrow\; \mathsf{tp}(d) \vartriangleleft \Pi \;\&\; \forall n \in |\mathsf{tp}(d)| \big(d[n] \vdash \mathsf{tp}(d)(\Pi, n)\big)$

$$d \in \mathcal{D}_2 \;:\Leftrightarrow \begin{cases} \mathrm{dg}(d[0]) < \mathrm{dg}(d) & \text{if } d \text{ critical} \\ \forall n \in |\mathsf{tp}(d)| (\mathrm{dg}(d[n]) \leq \mathrm{dg}(d)) & \text{otherwise} \end{cases}$$

$$d \in \mathcal{D}_3 \;:\Leftrightarrow \begin{cases} \tilde{\mathsf{o}}(d[0]) < \omega^{\tilde{\mathsf{o}}(d)} & \text{if } d \text{ critical} \\ \forall n \in |\mathsf{tp}(d)| (\tilde{\mathsf{o}}(d[n]) < \tilde{\mathsf{o}}(d)) & \text{otherwise} \end{cases}$$

Theorem 7.2. For $\nu = 1, 2, 3$ the following holds:
If $d = \mathsf{K}^r_\Pi d_0 \dots d_l \in \mathcal{D}$ with $d_0, \dots, d_l \in \mathcal{D}_\nu$, then $d \in \mathcal{D}_\nu$.

Proof:
Cf. the proofs of Theorem 3.4 and Lemma 4.1.

APPENDIX

In this appendix we will show how Gentzen's original ordinal assignment [Ge36, §15] can be transformed into the assignment which we have used in §4. This transformation consists in essentially four steps.

Step 1: We do not use exactly the same set of decimal fractions as Gentzen did. Gentzen defined his set of *Ordnungszahlen* (let's call it $\mathcal{O}_G$) by: $\mathcal{O}_G := \{n.u : n \in \mathbb{N} \;\&\; u \in \mathcal{M}_n\}$ where $\mathcal{M}_0 := \{1, 11, 111, \dots, 2\}$, $\mathcal{M}_{n+1} := \{u_0 0^{n+1} u_1 0^{n+1} \dots 0^{n+1} u_l : l \geq 0 \;\&\; u_0, \dots, u_l \in \mathcal{M}_n \;\&\; 0.u_l <_{\mathbb{R}} \dots <_{\mathbb{R}} 0.u_0\}$. This corresponds to representing ordinals in base 2 Cantor normal form, while here we shall use base ω. Instead of $\mathcal{O}_G$ we define the set $\mathcal{O} := \{n.u : n \in \mathbb{N} \;\&\; u \in M_n\}$, where $M_0 := \{1\}$, $M_{n+1} := \{u_0 0^{n+1} u_1 0^{n+1} \dots 0^{n+1} u_l : l \geq 0 \;\&\; u_0, \dots, u_l \in M_n \;\&\; 0.u_l \leq_{\mathbb{R}} \dots \leq_{\mathbb{R}} 0.u_0\}$.

Step 2: We define an embedding of $(\mathcal{O}, <_{\mathbb{R}})$ into the set theoretic ordinals, namely for each 'Ordnungszahl' $n.u \in \mathcal{O}$ we define an ordinal $|n.u| \in On$ such that $\forall n.u, m.v \in \mathcal{O}(n.u <_{\mathbb{R}} m.v \Rightarrow |n.u| < |m.v|)$ (Lemma 3).

Step 3: We modify Gentzen's assignment of 'Ordnungszahlen' to derivations ([Ge36, §15.2]) according to the alterations made in step 1. For each derivation d we define its *numerus* $\rho(d) \in \mathbb{N}$, *mantissa* $\mu(d) \in \bigcup_{n \in \mathbb{N}} M_n$, and *'Ordnungszahl'* $\mathrm{Ord}(d) := \rho(d).\mu(d) \in \mathcal{O}$. Actually we only consider the crucial case where d ends with a chain rule inference.

Step 4: We show how the ordinal $|\mathrm{Ord}(d)|$ can be defined directly by recursion on the build-up of d, without referring to the decimal fraction $\mathrm{Ord}(d)$. Then we compare the involved recursion equations with the corresponding equations in the definition of $\tilde{\mathsf{o}}(d)$, $o(d)$ in §4.

Step 1.

Let $\{0,1\}^+$ denote the set of all finite nonempty words u over the alphabet $\{0,1\}$, and let
$\{0,1\}^{(+)} := \{u \in \{0,1\}^+ :$ the first and the last letter of u is $1\}$.

Further, let 0^n denote the word consisting of n zeros. Each expression $n.u$ (with $n \in \mathbb{N}$ and $u \in \{0,1\}^{(+)}$) will be identified with the real number denoted by it in the usual way.

Definition of $M_n \subseteq \{0,1\}^{(+)}$
1. $M_0 := \{1\}$;
2. $M_{n+1} := \{u_0 0^{n+1} u_1 0^{n+1} \dots 0^{n+1} u_l : l \geq 0 \ \&\ u_0,\dots,u_l \in M_n \ \&\ 0.u_l \leq_{\mathbb{R}} \dots \leq_{\mathbb{R}} 0.u_0\}$.

Further we set $M := \bigcup_{n \in \mathbb{N}} M_n$. The elements of M are called *mantissas*.

Definition. $\mathrm{h} : M \to \mathbb{N}$, $\mathrm{h}(u) := \min\{n : u \in M_n\}$.

Remark. $M_n \subseteq M_{n+1}$, and $\mathrm{h}(u)$ is the maximal number of consecutive zeros in u.

Lemma 1.
If $u = u_0 0^{n+1} \dots 0^{n+1} u_l \in M_{n+1}$ and $v = v_0 0^{n+1} \dots 0^{n+1} v_k \in M_{n+1}$ with $u_0,\dots,u_l,v_0,\dots,v_k \in M_n$, then $0.u <_{\mathbb{R}} 0.v$ if, and only if, $l < k \ \&\ \forall i \leq l (u_i = v_i)$ or $\exists j \leq \min\{l,k\}\big(\forall i < j (u_i = v_i) \ \&\ 0.u_j <_{\mathbb{R}} 0.v_j\big)$.

Proof: Straightforeward.

Definition. $\mathcal{O} := \{n.u : n < \omega \ \&\ u \in M_n\}$ (*Ordnungszahlen*)

Step 2.

Definition of $|u|_n \in On$ **for** $u \in M_n$
1. $|1|_0 := 0$. 2. If $u = u_0 0^{n+1} \dots 0^{n+1} u_l \in M_{n+1}$ then $|u|_{n+1} := \omega^{|u_0|_n} + \dots + \omega^{|u_l|_n}$.

As usual we set $\omega_0(\alpha) := \alpha$, $\omega_{n+1}(\alpha) := \omega^{\omega_n(\alpha)}$.

Lemma 2. For $u \in M_n$ the following holds:
(a) $|u|_{n+k} = \omega_k(|u|_n)$,
(b) $\omega_n(0) \leq |u|_n < \omega_{n+1}(0)$.

Definition. For $n.u \in \mathcal{O}$ let $|n.u| := |u|_n \in On$.

Lemma 3. $n.u \in \mathcal{O} \ \&\ m.v \in \mathcal{O} \ \&\ n.u <_{\mathbb{R}} m.v \ \Rightarrow\ |n.u| < |m.v|$.

Proof by induction on the length of u:
Case $n < m$: Then $|n.u| = |u|_n < \omega_{n+1}(0) \leq \omega_m(0) \leq |v|_m = |m.v|$.
Case $n = m$: Then $0.u <_{\mathbb{R}} 0.v$ and $u, v \in M_n$ with $n > 0$. Hence $u = u_0 0^n \dots 0^n u_l \in M_n$ and $v = v_0 0^n \dots 0^n v_k \in M_n$ with $u_0,\dots,u_l,v_0,\dots,v_k \in M_{n-1}$. By Lemma 1 it follows that one of the following two cases applies:
(i) $l < k \ \&\ \forall i \leq l (u_i = v_i)$: Then trivially $|u|_n < |v|_n$.
(ii) $\forall i < j (u_i = v_i) \ \&\ 0.u_j <_{\mathbb{R}} 0.v_j$ for some $j \leq \min\{l,k\}$:
Then $\forall i \in \{j,\dots,l\}(0.u_i <_{\mathbb{R}} 0.v_j)$ and therefore, by IH, $\forall i \in \{j,\dots,l\}(|u_i|_{n-1} < |v_j|_{n-1})$.
Hence $|u|_n = \omega^{|v_0|_{n-1}} + \dots + \omega^{|v_{j-1}|_{n-1}} + \omega^{|u_j|_{n-1}} + \dots + \omega^{|u_l|_{n-1}} < \omega^{|v_0|_{n-1}} + \dots + \omega^{|v_j|_{n-1}} \leq |v|_n$.

Step 3.

The following are more or less Gentzen's own words (in [Ge36, 15.2]) — of course with some alterations enforced by the modifications made in step 1.

To each given derivation d we assign an 'Ordnungszahl' $\mathrm{Ord}(d) := \rho(d).\mu(d) \in \mathcal{O}$ *according to the following recursive rule: (...) If the endsequent of d is the conclusion of a 'chain rule' inference (i.e., if $d = \mathsf{K}^r_\Pi d_0 \dots d_l$) we consider the mantissas $u_i = \mu(d_i)$ of the 'Ordnungszahlen' of the derivations d_i; suppose that ν is the maximum number of consecutive zeros in all of these mantissas (i.e., $\nu = \max_{i \leq l} \mathrm{h}(u_i)$). The mantissas are written down from left to right according to their size (the largest one first) and any two successive mantissas are seperated by $\nu{+}1$ zeros. (It may well be that several successive mantissas are equal.) The result is the mantissa $\mu(d)$ of the ordinal number for the whole derivation; i.e., $\mu(d) := u_{\sigma(0)} 0^{\nu+1} u_{\sigma(1)} 0^{\nu+1} \dots 0^{\nu+1} u_{\sigma(l)}$*

where σ is an appropriate permutation of $\{0,\dots,l\}$, and $u_i = \mu(d_i)$. As the numerus $\rho(d)$ we take the least natural number ρ whose excess over the maximum number of consecutive zeros in the mantissa is ≥ 0 and, firstly, is not more than 1 less than the corresponding excess in any of the ordinal numbers for the derivations of the premises and, secondly, is not less than the rank of the succedent formula of any one of the premises preceding the major premise (14.25). W.l.o.g. we may assume here that $l \geq 1$ and therefore $\mathrm{h}(\mu(d)) = \nu{+}1$. So $\rho(d)$ is the least number ρ such that (i) $\rho - (\nu{+}1) \geq \rho(d_i) - \mathrm{h}(u_i) - 1$ for $i = 0,\dots,l$, and (ii) $\rho - (\nu{+}1) \geq r$, which amounts to: $\rho(d) - \mathrm{h}(\mu(d)) = \max(\{\rho(d_i) - \mathrm{h}(\mu(d_i)) - 1 : i \leq l\} \cup \{r\})$.

Step 4:

Let $\mathrm{h}(d) := \mathrm{h}(\mu(d))$, $\mathrm{exc}(d) := \rho(d) - \mathrm{h}(d)$, and $\widehat{\mathsf{o}}(d) := |\mu(d)|_{\mathrm{h}(d)}$

Then

(1) $|\mathrm{Ord}(d)| = \omega_{\mathrm{exc}(d)}(\widehat{\mathsf{o}}(d))$,

and for $d = \mathsf{K}^r_\Pi d_0 \dots d_l$ we have the recursion equations

(2) $\mathrm{h}(d) = \max_{i\leq l} \mathrm{h}(d_i) + 1$, and

(3) $\mathrm{exc}(d) = \max(\{\mathrm{exc}(d_i) - 1 : i \leq l\} \cup \{r\})$.

(4) $\widehat{\mathsf{o}}(d) = \omega^{\alpha_0} \# \dots \# \omega^{\alpha_l}$ with $\alpha_i := \omega_{\nu-\mathrm{h}(d_i)}(\widehat{\mathsf{o}}(d_i))$ and $\nu := \max_{i\leq l} \mathrm{h}(d_i)$.

Proof of (1) and (4):

(1) $|\mathrm{Ord}(d)| = |\rho(d).\mu(d)| = |\mu(d)|_{\rho(d)} = \omega_{\rho(d)-\mathrm{h}(d)}(\widehat{\mathsf{o}}(d)) = \omega_{\mathrm{exc}(d)}(\widehat{\mathsf{o}}(d))$.

(4) By definition, $\mu(d) = u_{\sigma(0)}0^{\nu+1}\dots 0^{\nu+1}u_{\sigma(l)}$ with $u_i = \mu(d_i)$ and $\nu = \max_{i\leq l} \mathrm{h}(d_i)$. Hence $\nu{+}1 = \mathrm{h}(\mu(d)) = \mathrm{h}(d)$, $\widehat{\mathsf{o}}(d) = |\mu(d)|_{\nu+1} = \omega^{|u_0|_\nu} \# \dots \# \omega^{|u_l|_\nu}$, and $|u_i|_\nu = |\mu(d_i)|_\nu \overset{\text{L.2a}}{=} \omega_{\nu-\mathrm{h}(d_i)}(|\mu(d_i)|_{\mathrm{h}(d_i)})$.

Observation: In case that $\mathrm{h}(\mu(d_0)) = \dots = \mathrm{h}(\mu(d_l))$ we have

(5) $\widehat{\mathsf{o}}(d) = \omega^{\widehat{\mathsf{o}}(d_0)} \# \dots \# \omega^{\widehat{\mathsf{o}}(d_l)}$.

Now compare (1), (3), (5) with the corresponding clauses in the definitions of $o(d)$, $\mathrm{dg}(d)$, $\tilde{\mathsf{o}}(d)$ in §4:

(1)' $o(d) = \omega_{\mathrm{dg}(d)}(\tilde{\mathsf{o}}(d))$

(3)' $\mathrm{dg}(d) = \max(\{\mathrm{dg}(d_i) - 1 : i \leq l\} \cup \{r\})$

(5)' $\tilde{\mathsf{o}}(d) = \omega^{\tilde{\mathsf{o}}(d_0)} \# \dots \# \omega^{\tilde{\mathsf{o}}(d_l)}$

Acknowledgement: I would like to thank the anonymous referee for careful reading and valuable comments which helped to improve the paper.

References

[Bu97] Buchholz, W.: *Explaining Gentzen's Consistency Proof within infinitary Proof Theory.* In: Gottlob, G., Leitsch, A., Mundici, D. (eds.) Computational Logic and Proof Theory. 5th Kurt Gödel Colloquium, KGC'97, Lecture Notes in Computer Science, Vol.1298. Springer 1997

[Ge36] Gentzen, G.: *Die Widerspruchsfreiheit der reinen Zahlentheorie.* Math. Ann. 112 (1936), pp. 493-565.

[Ge38] Gentzen, G.: *Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie.* Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften. Neue Folge 4, pp. 19-44 (1938)

[KB81] Kogan-Bernstein, L.M.: *Simplification of Gentzen's reductions in the classical arithmetic.* Zap. Nauch. Sem. LOMI 105 (1981) English translation: J.MATH.SCI. Vol.22(3) (1983)

[Sz69] Szabo, M.E.: *The collected papers of Gerhard Gentzen.* Amsterdam: North-Holland 1969.