Proof Theory

Gaisi Takeuti

June 23, 2021

Contents

1 First Order Predicate Calculus **1**
1.1 Formalization of statements 1
1.2 Formal proofs and related concepts 3
1.3 A formulation of intuitionistic predicate calculus 9
1.4 Axiom systems 9
1.5 The cut-elimination theroem 11
1.6 Some consequences of the cut-elimination theorem 17
1.7 The predicate calculus with equality PROBLEM 25
1.8 The completeness theorem 28

2 Peano Arithmetic **36**
2.1 A formulation of Peano arithmetic 36
2.2 The Incompleteness Theorem 40
2.3 A Discussion of Ordinals from a Finitist Standpoint 46
2.4 A Consistency Proof of **PA** 55

3 TODO ALL the problems **62**

1 First Order Predicate Calculus

In this chapter we shall present Gentzen's formulation of the first order predicate calculus **LK** (logistischer klassischer Kalkül). Intuitionisitic logic is known as **LJ** (logistischer intuitionistischer Kalkül)

1.1 Formalization of statements

Definition 1.1. **Terms** are defined inductively as follows:

1. Every individual constant is a term
2. Every free variable is a term
3. If f^i is a function constant with i argument-places and $t_1, \dots, t_i$ are terms, then $f^i(t_1, \dots, t_i)$ is a term
4. Terms are only those expressions obtained by 1-3.

Definition 1.2. **Formulas** are defined inductively as:

3. If A is a formula, a is a free variable and x is a bound variable not occurring in A, then $\forall x A'$ and $\exists x A'$ are formulas, where A' is the expression obtained from A by writing x in place of a at each occurrence of a in A

Definition 1.3. Let A be an expression, let $\tau_1, \dots, \tau_n$ be distinct primitive symbols, and let $\sigma_1, \dots, \sigma_n$ be any symbols. By

$$\left(A \frac{\tau_1, \dots, \tau_n}{\sigma_1, \dots, \sigma_n}\right)$$

we mean the expression obtained from A by writing $\sigma_1, \dots, \sigma_n$ in place of $\tau_1, \dots, \tau_n$ respectively at each occurrence of $\tau_1, \dots, \tau_n$. Such an operation is called the **(simultaneous) replacement of** $(\tau_1, \dots, \tau_n)$ **by** $(\sigma_1, \dots, \sigma_n)$ **in** A.

Proposition 1.4. 1. *If A contains none of $\tau_1, \dots, \tau_n$, then*

$$\left(A \frac{\tau_1, \dots, \tau_n}{\sigma_1, \dots, \sigma_n}\right)$$

is A itself

2. *If $\sigma_1, \dots, \sigma_n$ are distinct primitive symbols, then*

$$\left(\left(A \frac{\tau_1, \dots, \tau_n}{\sigma_1, \dots, \sigma_n}\right) \frac{\sigma_1, \dots, \sigma_n}{\theta_1, \dots, \theta_n}\right)$$

is identical with

$$\left(A \frac{\tau_1, \dots, \tau_n}{\theta_1, \dots, \theta_n}\right)$$

Definition 1.5. 1. Let A be a formula and $t_1, \dots, t_n$ be terms. If there is a formula B and n distinct free variables $b_1, \dots, b_n$ s.t. A is

$$\left(B \frac{b_1, \dots, b_n}{t_1, \dots, t_n} \right)$$

then for each $i (1 \leq i \leq n)$ the occurrences of t_1 resulting from the above replacement are said to be **indicated** in A, and this fact is also expressed by writing B as $B(b_1, \dots, b_n)$ and A as $B(t_1, \dots, t_n)$

2. A term t is **fully indicated** in A, or every occurrence of t in A is indicated, if every occurrence of t is obtained by such a replacement

Proposition 1.6. *If A is a formula (where a is not necessarily fully indicated) and x is a bound variable not occurring in $A(a)$, then $\forall x A(x)$ and $\exists x A(x)$ are formulas*

1.2 Formal proofs and related concepts

Definition 1.7. For arbitrary Γ and Δ in the above notation, $\Gamma \to \Delta$ is called a **sequent**. Γ and Δ are called the **antecedent** and **succedent**, respectively, of the sequent and each formula in Γ and Δ is called a **sequent-formula**

Definition 1.8. An **inference** is an expression of the form

$$\frac{S_1}{S} \text{ or } \frac{S_1 \quad S_2}{S}$$

where S_1, S_2 and S are sequents. S_1 and S_2 are called the **upper sequents** and S is called the **lower sequent** of the inference

Structural rules

1. **Weakening**:

$$\text{left: } \frac{\Gamma \to \Delta}{D, \Gamma \to \Delta}; \quad \text{right: } \frac{\Gamma \to \Delta}{\Gamma \to \Delta, D}$$

D is called the **weakening formula**

2. **Contraction**:

$$\text{left: } \frac{D, D, \Gamma \to \Delta}{D, \Gamma \to \Delta} \quad \text{right: } \frac{\Gamma \to \Delta, D, D}{\Gamma \to \Delta, D}$$

3. **Exchange**

$$\text{left: } \frac{\Gamma, C, D, \Pi \to \Delta}{\Gamma, D, C, \Pi \to \Delta} \quad \text{right: } \frac{\Gamma \to \Delta, C, D, \Lambda}{\Gamma \to \Delta, D, C, \Lambda}$$

We will refer to these three kinds of inferences as "weak inferences", while all others will be called "strong inferences"

4. **Cut**

$$\frac{\Gamma \to \Delta, D \qquad D, \Pi \to \Lambda}{\Gamma, \Pi \to \Delta, \Lambda}$$

D is called the **cut formula** of this instance

Logical rules

1.

$$\neg : \text{left: } \frac{\Gamma \to \Delta, D}{\neg D, \Gamma \to \Delta}; \quad \neg : \text{right: } \frac{D, \Gamma \to \Delta}{\Gamma \to \Delta, \neg D}$$

D and $\neg D$ are called the **auxiliary formula** and the **principal formula** respectively, of this inference

2.

$$\frac{C, \Gamma \to \Delta}{C \wedge D, \Gamma \to \Delta} \wedge\text{left} \quad \text{and} \quad \frac{D, \Gamma \to \Delta}{C \wedge D, \Gamma \to \Delta} \wedge\text{left}$$

$$\frac{\Gamma \to \Delta, C \qquad \Gamma \to \Delta, D}{\Gamma \to \Delta, C \wedge D} \wedge\text{right}$$

C and D are called the auxiliary formulas and $C \wedge D$ is called the principal formula of this inference

3.

$$\frac{C, \Gamma \to \Delta \qquad D, \Gamma \to \Delta}{C \vee D, \Gamma \to \Delta} \vee\text{left}$$

$$\frac{\Gamma \to \Delta, C}{\Gamma \to \Delta, C \vee D} \vee\text{right} \quad \text{and} \quad \frac{\Gamma \to \Delta, D}{\Gamma \to \Delta, C \vee D} \vee\text{right}$$

C and D are called the auxiliary formulas and $C \vee D$ the principal formula of this inference

4.

$$\frac{\Gamma \to \Delta, C \qquad D, \Pi \to \Lambda}{C \supset D, \Gamma, \Pi \to \Delta, \Lambda} \supset\text{left} \qquad \frac{C, \Gamma \to \Delta, D}{\Gamma \to \Delta, C \supset D} \supset\text{right}$$

C and D are called the auxiliary formulas and $C \supset D$ the principal formula

1-4 are called **propositional inferences**

5.
$$\frac{F(t), \Gamma \to \Delta}{\forall x F(x), \Gamma \to \Delta}\ \forall\text{left} \qquad \frac{\Gamma \to \Delta, F(a)}{\Gamma \to \Delta, \forall x F(x)}\ \forall\text{right}$$

where t is an arbitrary term, and a does not occur in the lower sequent. $F(t)$ and $F(a)$ are called the auxiliary formulas and $\forall x F(x)$ the principal formula. The a in $\forall$right is called the **eigenvariable** of this inference

In $\forall$right all occurrences of a in $F(a)$ are indicated. In $\forall$left, $F(t)$ and $F(x)$ are

$$\left(F(a)\frac{a}{t}\right) \quad \text{and} \quad \left(F(a)\frac{a}{t}\right)$$

respectively, so not every t in $F(t)$ is necessarily indicated

6.
$$\frac{F(a), \Gamma \to \Delta}{\exists x F(x), \Gamma \to \Delta}\ \exists\text{left} \qquad \frac{\Gamma \to \Delta, F(t)}{\Gamma \to \Delta, \exists x F(x)}\ \exists\text{right}$$

where a does not occur in the lower sequent, and t is an arbitrary term

$F(a)$ and Ft are called the auxiliary formulas and $\exists x F(x)$ the principal formula. The a in $\exists$left is called the eigenvariable of this inference

In $\exists$left a is fully indicated

5 and 6 are called the **quantifier inferences**. The condition, that the eigenvariable must not occur in the lower sequent in $\forall$right and $\exists$left is called the **eigenvariable condition**

A sequent of the form $A \to A$ is called an **initial sequent** or axiom

Definition 1.9. A **proof** P (in **LK**), or **LK-proof**, is a tree of sequents satisfying the following conditions

1. The topmost sequents of P are initial sequents
2. Every sequent in P except the lowest one is an upper sequent of an inference whose lower sequent is also in P

Definition 1.10. 1. A sequence of sequents in a proof P is called a **thread** (of P) if the following conditions are satisfied

   (a) The sequence begins with an initial sequent and ends with the end-sequent

(b) Every sequent in the sequence except the last is an upper sequent of an inference, and is immediately followed by the lower sequent of this inference

2. Let $S_1, S_2$and S_3 be sequents in a proof P. We say S_1 is **above** S_2or S_2 is **below** S_1 if there is a thread containing both $S_1$and S_2 where S_1 appears before S_2. If $S_1$is above $S_2$and S_2 is above S_3, we say S_2 is **between** S_1 and S_3

3. An inference in P is said to be **below a sequent** S if its lower sequent is below S

4. Let P be a proof. A part of P which itself is a proof is called a **subproof** of P. For any sequent S in P, that part of P which consists of all sequents which are either Sitself or which occur above $S$is called a subproof of P (with end-sequent S)

5. Let P_0 be a proof of the form

$$\begin{array}{c} \vdots \\ \Gamma \to \Theta \\ \vdots\,(*) \end{array}$$

where (*) denotes the part of P_0 under $\Gamma \to \Theta$, and let Q be a proof ending with $\Gamma, D \to \Theta$. By a copy of P_0 from Q we mean a proof P of the form

$$\begin{array}{c} \vdots\,Q \\ \Gamma, D \to \Theta \\ \vdots\,(**) \end{array}$$

where $(**)$ differs from $(*)$ only in that for each sequent in $(*)$, say $\Gamma \to \Lambda$, the corresponding sequent in $(**)$ has the form $\Pi, D \to \Lambda$.

6. Let $S(a)$ or $\Gamma(a) \to \Delta(a)$, denote a sequent of the form $A_1(a), \dots, A_m(a) \to B_1(a), \dots, B_n(a)$. Then $S(t)$, or $\Gamma(t) \to \Delta(t)$, denotes the sequent $A_1(t), \dots, A_m(t) \to B_1(t), \dots, B_n(t)$

Definition 1.11. A proof in **LK** is called **regular** if it satisfies the condition that all eigenvariables are distinct from one another and if a free variable a occurs as an eigenvariable in a sequent S of the proof, then a occurs only in sequents above S

Lemma 1.12. 1. *Let $\Gamma(a) \to \Delta(a)$ be an ($\boldsymbol{LK}$-)provable sequent in which a is fully indicated, and let $P(a)$ be a proof of $\Gamma(a) \to \Delta(a)$. Let b be a free variable not occurring in $P(a)$. Then the tree $P(b)$, obtained from $P(a)$ by replacing a by b at each occurrence of a in $P(a)$, is also a proof and its end-sequent is $\Gamma(b) \to \Delta(b)$*

2. *For an arbitrary $\boldsymbol{LK}$-proof there exists a regular proof of the same end-sequent. Moreover, the required proof is obtained from the original proof simply by replacing free variables*

Proof. 1. By induction on the number of inference in $P(a)$. If $P(a)$ consists of simply an initial sequent $A(a) \to A(a)$, then $P(b)$ consists of the sequent $A(b) \to A(b)$.

Suppose that our proposition holds for proofs containing at most n inferences and suppose that $P(a)$ contains $n+1$ inferences. We treat the possible cases according to the last inferences in $P(a)$. Since other cases can be treated similarly, we consider only the case where the last inference, say J, is a $\forall$right. Suppose the eigenvariable of J is a, and $P(a)$ is of the form

$$\dfrac{\begin{array}{c}\vdots\, Q(a)\\ \Gamma \to \Lambda, A(a)\end{array}}{\Gamma \to \Lambda, \forall x A(x)}\ J$$

where $Q(a)$ is the subproof of $P(a)$ ending with $\Gamma \to \Lambda, A(a)$. a doesnt occur in Γ, Λ or $A(x)$. By the induction hypotheses the result of replacing all a's in $Q(a)$ by b is a proof whose end-sequent is $\Gamma \to \Lambda, A(b)$. Γ and Λ contain no b's. Thus we can apply a $\forall$right to this sequent using b as its eigenvariable

$$\dfrac{\begin{array}{c}\vdots\, Q(b)\\ \Gamma \to \Lambda, A(b)\end{array}}{\Gamma \to \Lambda, \forall x A(x)}$$

and so $P(b)$ is a proof ending with $\Gamma \to \Lambda, \forall x A(x)$. If a is not the eigenvariable of J, $P(a)$ is of the form

$$\dfrac{\begin{array}{c}\vdots\, Q(a)\\ \Gamma(a) \to \Lambda(a), A(a,c)\end{array}}{\Gamma(a) \to \Lambda(a), \forall x A(a,x)}$$

By the induction hypothesis the result of replacing all a's in $Q(a)$ by $b$is a proof and its end-sequent is $\Gamma(b) \rightarrow \Lambda(b), A(b,c)$

Since by assumption b doesn't occur in $P(a)$, b is not c and so we can apply a $\forall$right to this sequent, with c as its eigenvariable

2. By mathematical induction on the number l of applications of $\forall$right and $\exists$left in a given proof P. If $l = 0$ then take P itself. Otherwise, P can be represented in the form

$$\begin{array}{c} P_1 \quad P_2 \dots P_k \\ \vdots \, (*) \\ S \end{array}$$

where P_i is a subproof of P of the form

$$\cfrac{\begin{array}{c}\vdots \\ \Gamma_i \rightarrow \Delta_i, F_i(b_i)\end{array}}{\Gamma_i \rightarrow \Delta_i, \forall y_i F_i(y_i)} \, I_i \qquad \text{or} \qquad \cfrac{\begin{array}{c}\vdots \\ F_i(b_i), \Gamma_i \rightarrow \Delta_i\end{array}}{\exists y_i F_i(y_i), \Gamma_i \rightarrow \Delta_i} \, I_i$$

and I_i is a lowermost $\forall$right or $\exists$left in P

Let us deal with the case where I_i is $\forall$right. P_i has fewer applications of $\forall$right or $\exists$left than P, so by the induction hypothesis there is a regular proof P'_i of $\Gamma_i \rightarrow \Delta_i, F_i(b_i)$. Note that no free variable in $\Gamma_i \rightarrow \Delta_i, F_i(b_i)$ (including b_i) is used as an eigenvariable in P'_i. Suppose $c_1, \dots, c_m$ are all the eigenvariables in all the P_i's which occur in P above $\Gamma_i \rightarrow \Delta_i, \forall y_i F_i(y_i), i = 1, \dots, k$. Then change $c_1, \dots, c_m$ to $d_1, \dots, d_m$ respectively, where $d_1, \dots, d_m$ are the first m variables which occur neither in P nor in $P_i{}'$. If b_i occurs in P below $\Gamma_i \rightarrow \Delta_i, \forall y_i F_i(y_i)$ then change it to d_{m+i}

Let P''_i be the proof which is obtained from P'_i by the above replacement of varaibles. Then $P''_1, \dots, P''_k$ are each regular

$$\begin{array}{c} P''_1 \dots \cfrac{P''_i}{\Gamma_i \rightarrow \Delta_i, \forall y_i F_i(y_i)} \dots P''_n \\ \vdots \, (*) \\ S \end{array}$$

□

From now on we will assume that we are dealing with regular proofs whenever convenient

Lemma 1.13. *Let t be an arbitrary term. Let $\Gamma(a) \to \Delta(a)$ be a provable (in **LK**) sequent in which a is fully indicated, and let $P(a)$ be a proof ending with $\Gamma(a) \to \Delta(a)$ in which **every eigenvariable is different from** a **and not contained in** t. Then $P(t)$ is a proof whose end-sequent is $\Gamma(t) \to \Delta(t)$*

Lemma 1.14. *Let t be an arbitrary term. Let $\Gamma(a) \to \Delta(a)$ be a provable (in **LK**) sequent in which a is fully indicated, and let $P(a)$ be a proof of $\Gamma(a) \to \Delta(a)$. Let $P'(a)$ be a proof obtained from $P(a)$ by changing eigenvariables in such a way that in $P'(a)$ every eigenvariable is different from a and not contained in t. Then $P'(t)$ is a proof of $\Gamma(t) \to \Delta(t)$*

Proposition 1.15. *Let t be an arbitrary term and $S(a)$ a provable sequent in which a is fully indicated. Then $S(t)$ is also provable*

Proposition 1.16. *If a sequent is provable, then it is provable with a proof in which all the initial sequents consist of atmoic formulas. Furthermore, if a sequent is provable without cut, then it is provable without cut with a proof of the above sort*

Proof. It suffices to show that for an arbitrary formula A, $A \to A$ is provable without cut, starting with initial sequents consisting of atomic formulas. □

Definition 1.17. Two formulas A and B are **alphabetical variants** if for some $x_1, \dots, x_n, y_1, \dots, y_n$

$$\left(A \frac{x_1, \dots, x_n}{z_1, \dots, z_n}\right)$$

is

$$\left(B \frac{y_1, \dots, y_n}{z_1, \dots, z_n}\right)$$

where $z_1, \dots, z_n$ are bound variables occurring neither in A nor in B. The fact that A and B are alphabetical variants will be expressed by $A \sim B$

1.3 A formulation of intuitionistic predicate calculus

Definition 1.18. We can formalize the intuitionistic predicate calculus as a subsystem of **LK** which we call **LJ** following Gentzen (**J** stands for "intuitionistic"). **LJ**is obtained from **LK** by modifying it as follows

1. A sequent in **LJ** is of the form $\Gamma \to \Delta$ where Δ consists of at most one formula

2. Inferences in **LJ** are those obtained from those in **LK** by imposing the restriction that the succedent of each upper and lower sequent consists of at most one formula; thus there are no inferences in **LJ**corresponding to contraction right or exchange right

Proposition 1.19. *If a sequent S of **LJ** is provable in **LJ**, then it is also provable in **LK***

1.4 Axiom systems

Definition 1.20. The basic system is **LK**

1. A finite or infinite set $\mathcal{A}$ of sentences is called an **axiom system**, and each of these sentences is called an **axiom** of $\mathcal{A}$. Sometimes an axiom system is called a **theory**

2. A finite (possibly empty) sequence of formulas consisting only of axioms of $\mathcal{A}$ is called an **axiom sequence** of $\mathcal{A}$

3. If there exists an axiom sequence Γ_0 of $\mathcal{A}$ s.t. $\Gamma_0, \Gamma \to \Delta$ is **LK**-provable, then $\Gamma \to \Delta$ is said to be **provable from** $\mathcal{A}$ (in **LK**). We express this by $\mathcal{A}, \Gamma \to \Delta$

4. $\mathcal{A}$ is **inconsistent** (with **LK**) if the empty sequent $\to$ is provable from $\mathcal{A}$ (in **LK**)

5. If all function constants and predicate constants in a formula A occur in $\mathcal{A}$, then A is said to be **dependent on** $\mathcal{A}$

6. A sentence A is **consistent** if the axiom system $\{A\}$ is consistent

7. $\mathbf{LK}_{\mathcal{A}}$ is the system obtained from **LK** by adding $\to A$ as initial sequents for all A in $\mathcal{A}$

Proposition 1.21. *Let $\mathcal{A}$ be an axiom system. Then the following are equivalent*

1. *$\mathcal{A}$ is inconsistent (with **LK**)*

2. *for every formula A, A is provable from $\mathcal{A}$*

3. *for some formula A, $A$and $\neg A$ are both provable from $\mathcal{A}$*

Proof. $3 \to 1$. we have $\mathbf{LK} \vdash A \leftrightarrow \neg\neg A$. So from $\to \neg A$ we have $A \to$. THen we apply cut. $\square$

Proposition 1.22. *Let $\mathcal{A}$ be an axiom system. Then a sequent $\Gamma \to \Delta$ is* $\boldsymbol{LK}_{\mathcal{A}}$*-provable iff $\Gamma \to \Delta$ is provable from $\mathcal{A}$ (in **LK**)*

Corollary 1.23. *An axiom system $\mathcal{A}$ is consistent (with **LK**) iff* $\boldsymbol{LK}_{\mathcal{A}}$ *is consistent*

These definitions and the propositions hold also for **LJ**

1.5 The cut-elimination theroem

Theorem 1.24 (the cut-elimination theroem: Gentzen)**.** *If a sequent is (**LK**)-provable, then it is (**LK**)-provable without a cut*

Let Abe a formula. An inference of the following form is called a **mix** (w.r.t. A):

$$\frac{\Gamma \to \Delta \qquad \Pi \to \Lambda}{\Gamma, \Pi^* \to \Delta^*, \Lambda}\ A$$

where both Δ and Π contain the formula A, and Δ^* and Π^* are obtained from Δ and Π respectively by deleting all the occurrences of A in them. We call A the mix formula of this inference.

Let's call the system which is obtained from **LK** by replacing the cut rule by the mix rule, **LK***.

Lemma 1.25. ***LK** and **LK*** are equivalent, that is, a sequent S is **LK**-provable iff S is **LK***-provable*

mix is a strengthened version of cut

Theorem 1.26. *If a sequent is provable in **LK***, then it's provable in **LK*** without a mix*

Lemma 1.27. *If P is a proof of S (in **LK***) which contains (only) one mix, occurring as the last inference, then S is provable without a mix*

The **grade** of a formula A (denoted by $g(A)$) is the number of logical symbols contained in A. The grade of a mix is the grade of the mix formula. When a proof P has a mix as the last inference, we define the grade of P (denoted by $g(P)$) to be the grade of this mix.

Let P be a proof which contains a mix only as the last inference

$$J\,\frac{\Gamma \to \Delta \qquad \Pi \to \Lambda}{\Gamma, \Pi^* \to \Delta^*, \Lambda}\ (A)$$

We refer to the left and right upper sequents as S_1 and S_2 and the lower sequent as S. We call a thread in P a **left (right) thread** if it contains the left (right) upper sequent of the mix J. The **rank** of a thread $\mathcal{F}$ in P is defined as follows: if $\mathcal{F}$ is a left (right) thread, then the rank of $\mathcal{F}$ is the number consecutive sequents, counting upward from the left (right) upper sequent of J, that contains the mix formula in its succedent (antecedent). The rank of a thread $\mathcal{F}$ in P is denoted by $\text{rank}(\mathcal{F}; P)$. We define

$$\text{rank}_l(P) = \max_{\mathcal{F}}(\text{rank}(\mathcal{F}; P))$$

where $\mathcal{F}$ ranges over all the left threads in P, and

$$\text{rank}_r(P) = \max_{\mathcal{F}}(\text{rank}(\mathcal{F}; P))$$

where $\mathcal{F}$ ranges over all the right threads in P. The rank of P, $\text{rank}(P)$, is defined as

$$\text{rank}(P) = \text{rank}_l(P) + \text{rank}_r(P)$$

Note that $\text{rank}(P) \geq 2$

Proof. We prove the Lemma by double induction on the grade g and rank r of the proof P (i.e. transfinite induction on $\omega \cdot g + r$). We divide the proof into two main cases, namely $r = 2$ and $r > 2$

1. $r = 2$, $\text{rank}_l(P) = \text{rank}_r(P) = 1$

   (a) The left upper sequent S_1 is an initial sequent. In this case we may assume P is of the form

   $$J \, \frac{A \to A \qquad \Pi \to \Lambda}{A, \Pi^* \to \Lambda}$$

   We can obtain the lower sequent without a mix

   $$\cfrac{\cfrac{\cfrac{\Pi \to \Lambda}{\text{some exchanges}}}{A, \ldots, A, \Pi^* \to \Lambda}}{\cfrac{\text{some contractions}}{A, \Pi^* \to \Lambda}}$$

   (b) The right upper sequent S_2 is an initial sequent.

(c) Neither S_1 nor S_2 is an initial sequent, and S_1 is the lower sequent of a structural inference J_1. Since $\mathrm{rank}_l(P) = 1$, the formula A cannot appear in the succedent of the upper sequent of J_1. Hence

$$\dfrac{\dfrac{\Gamma \to \Delta_1}{\Gamma \to \Delta_1, A}\,J_1 \qquad \Pi \to \Lambda}{\Gamma, \Pi^* \to \Delta_1, \Lambda}\,J$$

where Δ_1 doesn't contain A. We can eliminate the mix as follows

$$\dfrac{\dfrac{\dfrac{\dfrac{\Gamma \to \Delta_1}{\text{some weakenings}}}{\Pi^*, \Gamma \to \Delta_1, \Lambda}}{\text{some exchanges}}}{\Gamma, \Pi^* \to \Delta_1, \Lambda}$$

(d) None of 1.1-1.3 holds but S_2 is the lower sequent of a structural inference. Similarly

(e) Both $S_1$and $S_2$are the lower sequents of logical inferences. In this case, since $\mathrm{rank}_l(P) = \mathrm{rank}_r(P) = 1$, the mix formula on each side must be the principal formula of the logical inference. We use induction on the grade, distinguishing several cases according to the outermost logical symbol of A

i. The outermost logical symbol of A is $\wedge$

$$\dfrac{\dfrac{\Gamma \to \Delta_1, B \qquad \Gamma \to \Delta_1, C}{\Gamma \to \Delta_1, B \wedge C} \qquad \dfrac{B, \Pi_1 \to \Lambda}{B \wedge C, \Pi_1 \to \Lambda}}{\Gamma, \Pi_1 \to \Delta_1, \Lambda}\,(B \wedge C)$$

where by assumption none of the proofs ending with $\Gamma \to \Delta_1, B; \Gamma \to \Delta_1, C$ or $B, \Pi_1 \to \Lambda$ contain a mix. Consider the following

$$\dfrac{\Gamma \to \Delta_1, B \qquad B, \Pi_1 \to \Lambda}{\Gamma, \Pi_1^{\#} \to \Delta_1^{\#}, \Lambda}\,(B)$$

This proof contains only one mix, a mix that occurs as its last inference. Furthermore the grade of the mix formula B is less than $g(A)$. So by induction hypothesis we can obtain a proof which contains no mixes and whose end-sequent is $\Gamma, \Pi_1^{\#} \to \Delta_1^{\#}, \Lambda$. From this we can obtain a proof without a mix with end-sequent $\Gamma, \Pi_1 \to \Delta_1, \Lambda$

ii. The outermost logical symbol of A is $\vee$. Similar.

iii. The outermost logical symbol of A is $\forall$ company

$$\dfrac{\dfrac{\Gamma \to \Delta_1, F(a)}{\Gamma \to \Delta_1, \forall x F(x)} \qquad \dfrac{F(t), \Pi_1 \to \Lambda}{\forall x F(x), \Pi_1 \to \Lambda}}{\Gamma, \Pi_1 \to \Delta_1, \Lambda}$$

(a being fully indicated in $F(a)$). By the eigenvariable condition, a does not occur in Γ, Δ_1 or $F(x)$. Since by assumption the proof ending with $\Gamma \to \Delta_1, F(a)$ contains no mix, we can obtain a proof without a mix, ending with $\Gamma \to \Delta_1, F(t)$. Consider

$$\dfrac{\Gamma \to \Delta_1, F(t) \qquad F(t), \Pi_1 \to \Lambda}{\Gamma, \Pi_1^{\#} \to \Delta_1^{\#}, \Lambda}\ (F(t))$$

iv. The outermost logical symbol of A is $\exists$. Similar.

v. The outermost logical symbol of A is $\neg$. Then the end of the derivation runs

$$\dfrac{\dfrac{A, \Gamma \to \Delta_1}{\Gamma \to \Delta_1, \neg A} \qquad \dfrac{\Pi_1 \to \Lambda, A}{\neg A, \Pi_1 \to \Lambda}}{\Gamma, \Pi_1 \to \Delta_1, \Lambda}$$

This is transformed into

$$\dfrac{\dfrac{\Pi_1 \to \Lambda, A \qquad A, \Gamma \to \Delta_1}{\Pi_2 \to \Gamma^{\#} \to \Lambda^{\#}, \Delta_1}}{\overline{\Gamma, \Pi_1 \to \Delta_1, \Lambda}}$$

vi. The outermost logical symbol of A is $\supset$.

$$\dfrac{\dfrac{C, \Gamma_1 \to \Delta_1, D}{\Gamma_1 \to \Delta_1, C \supset D} \qquad \dfrac{\Gamma \to \Delta, C \qquad D, \Pi \to \Lambda}{C \supset D, \Gamma, \Pi \to \Delta, \Lambda}}{\Gamma_1, \Gamma, \Pi \to \Delta_1, \Delta, \Lambda}$$

This is transformed into

$$\dfrac{\dfrac{\Gamma \to \Delta, C \qquad \dfrac{C, \Gamma_1 \to \Delta_1, D \qquad D, \Pi \to \Lambda}{C, \Gamma_1, \Pi^{\#} \to \Delta_1^{\#}, \Lambda}}{\Gamma, \Gamma_1^{\#}, \Pi^{\#\#} \to \Delta^{\#}, \Delta_1^{\#}, \Lambda}}{\overline{\Gamma_1, \Gamma, \Pi \to \Delta_1, \Delta, \Lambda}}$$

2. $r > 2$, i.e., $\text{rank}_l(P) > 1$ and/or $\text{rank}_r(P) > 1$

$$\dfrac{\dfrac{A, \Gamma \to \Delta_1}{\Gamma \to \Delta_1, \neg A} \qquad \dfrac{\Pi_1 \to \Lambda, A}{\neg A, \Pi_1 \to \Lambda}}{\Gamma, \Pi_1 \to \Delta_1, \Lambda}$$

This is transformed into

$$\dfrac{\dfrac{\Pi_1 \to \Lambda, A \qquad A, \Gamma \to \Delta_1}{\Pi_2 \to \Gamma^{\#} \to \Lambda^{\#}, \Delta_1}}{\Gamma, \Pi_1 \to \Delta_1, \Lambda}$$

We distinguish two main cases: The right rank is greater than 1 and the right rank is equal to 1

The induction hypothesis is that every proof Q which contains a mix only as the last inference, and which satisfies either $g(Q) < g(P)$, or $g(Q) = g(P)$ and $\text{rank}(Q) < \text{rank}(P)$, we can eliminate the mix

(a) $\text{rank}_r(P) > 1$

i. Γ or Δ (in S_1) contains A. Construct a proof as follows

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\Pi \to \Lambda}}{\text{exchanges/contractions}}}{A, \Pi^* \to \Lambda}}{\dfrac{\text{weakenings/exchanges}}{\Gamma, \Pi^* \to \Delta^*, \Lambda}} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\Gamma \to \Delta}}{\text{exchanges/contractions}}}{\Gamma \to \Delta^*, A}}{\dfrac{\text{weakenings/exchanges}}{\Gamma, \Pi^* \to \Delta^*, \Lambda}}$$

ii. S_2 is the lower sequent of an inference J_2, where J_2 is not a logical inference whose principal formula is A. The last part of P looks like this

$$\dfrac{\Gamma \to \Delta \qquad \dfrac{\Phi \to \Psi}{\Pi \to \Lambda} J_2}{\Gamma, \Pi^* \to \Delta^*, \Lambda}$$

where the proofs $\Gamma \to \Delta$ and $\Phi \to \Psi$ contain no mixes and Φ contains at least one A. Consider the following proof P':

$$\dfrac{\Gamma \to \Delta \qquad \Phi \to \Psi}{\Gamma, \Phi^* \to \Delta^*, \Psi} \ (A)$$

In P', the grade of the mix is equal to $g(P)$, $\text{rank}_l(P') = \text{rank}_l(P)$ and $\text{rank}_r(P') = \text{rank}_r(P) - 1$. Thus by induction hypothesis, $\Gamma, \Phi^* \to \Delta^*, \Psi$ is provable without a mix. Then we construct the proof

$$\begin{array}{c} \Gamma, \Phi^* \to \Delta^*, \Psi \\ \hline \text{some exchanges} \\ \hline \dfrac{\Phi^*, \Gamma \to \Delta^*, \Psi}{\Pi^*, \Gamma \to \Delta^*, \Lambda} \, J_2 \end{array}$$

iii. Γ contains no A's and S_2 is the lower sequent of a logical inference whose principal formula is A.

A. A is $B \supset C$. The last part of P is of the form

$$\dfrac{\Gamma \to \Delta \qquad \dfrac{\Pi_1 \to \Lambda_1, B \qquad C, \Pi_2 \to \Lambda_2}{B \supset C, \Pi_1, \Pi_2 \to \Lambda_1, \Lambda_2}}{\Gamma, \Pi_1^*, \Pi_2^* \to \Delta^*, \Lambda_1, \Lambda_2}$$

Consider the following proofs P_1 and P_2

$$\dfrac{\Gamma \to \Delta \qquad \Pi_1 \to \Lambda_1, B}{\Gamma_1^* \to \Delta^*, \Lambda_1, B} \, B \supset C \qquad\qquad \dfrac{\Gamma \to \Delta \qquad C, \Pi_2 \to \Lambda_2 \to \Lambda_2}{\Gamma, C, \Pi_2^* \to \Delta^*, \Lambda_2} \, B \supset C$$

assuming that $B \supset C$ is in Π_1 and Π_2. If $B \supset C$ is not in Π_i ($i = 1$ or 2), then Π_i^* is Π_i and P_i is defined as

$$\begin{array}{c} \Pi_1 \to \Lambda_1, B \\ \hline\hline \Gamma, \Pi_1^* \to \Delta^*, \Lambda_1, B \end{array} \qquad\qquad \begin{array}{c} C, \Pi_2 \to \Lambda_2 \\ \hline\hline \Gamma, C, \Pi_2^* \to \Delta^*, \Lambda_2 \end{array}$$

Note that $g(P_1) = g(P_2) = g(P)$, $\text{rank}_l(P_1) = \text{rank}_l(P_2) = \text{rank}_l(P)$ and $\text{rank}_r(P_1) = \text{rank}_r(P_2) = \text{rank}_r(P) - 1$. Hence by the induction hypothesis, the end-sequents of P_1 and P_2 are provable without a mix (say by P_1' and P_2'). Consider the following proof P'

$$\dfrac{\Gamma \to \Delta \qquad \dfrac{\begin{array}{c} \vdots \, P_1' \\ \Gamma, \Pi_1^* \to \Delta^*, \Lambda_1, B \end{array} \qquad \begin{array}{c} \vdots \, P_2' \\ \Gamma, C, \Pi_2^* \to \Delta^*, \Lambda_2 \\ \hline\hline C, \Gamma, \Pi_2^* \to \Delta^*, \Lambda_2 \end{array}}{B \supset C, \Gamma, \Pi_1^*, \Gamma, \Pi_2^* \to \Delta^*, \Lambda_1, \Delta^*, \Lambda_2}}{\Gamma, \Gamma, \Pi_1^*, \Gamma, \Pi_2^* \to \Delta^*, \Delta^*, \Lambda_1, \Delta^*, \Lambda_2} \, B \supset C$$

Then $g(P') = g(P)$, $\text{rank}_l(P') = \text{rank}_l(P)$, $\text{rank}_r(P') = 1$. Thus the end-sequent of P' is provable without a mix by the induction hypothesis. wefwaefwefwefaweewojweoifaewjfoi

B. A is $\exists x F(x)$. The last part of P looks like this

$$\dfrac{\Gamma \to \Delta \qquad \dfrac{F(a), \Pi_1 \to \Lambda}{\exists x F(x), \Pi_1 \to \Lambda}}{\Gamma, \Pi_1^* \to \Delta^*, \Lambda} \; \exists x F(x)$$

Let b be a free variable not occurring in P. Then the result of replacing a by b throughout the proof ending with $F(a), \Pi_1 \to \Lambda$ without a mix, ending with $F(b), \Pi_1 \to \Lambda$, since by the eigenvariable condition, a does not occur in Π_1 or Λ (Lemma 1.13)

Consider the following proof:

$$\dfrac{\Gamma \to \Delta \qquad F(b), \Pi_1 \to \Lambda}{\Gamma, F(b), \Pi_1^* \to \Delta^*, \Lambda} \; \exists x F(x))$$

By the induction hypothesis, the end-sequent of this proof can be proved without a mix (say by P')). Now consider the proof

$$\dfrac{\Gamma \to \Delta \qquad \dfrac{\dfrac{\vdots\, P'}{\Gamma, F(b), \Pi_1^* \to \Delta^*, \Lambda}}{\dfrac{F(b), \Gamma, \Pi_1^* \to \Delta^*, \Lambda}{\exists x F(x), \Gamma, \Pi_1^* \to \Delta^*, \Lambda}}}{\Gamma, \Gamma, \Pi_1^* \to \Delta^*, \Delta^*, \Lambda}$$

1. $\text{rank}_r(P) = 1$).

□

Theorem 1.28. *The cut-elimination theorem holds for* ***LJ***

1.6 Some consequences of the cut-elimination theorem

Definition 1.29. By a **subformula** of a formula A we mean a formula used in building up A.

Two formulas A and B are said to be **equivalent** in **LK**if $A \equiv B$ is provable in **LK**

In a formula A an occurrence of a logical symbol, say $\sharp$ is **in the scope** of an occurrences of a logical symbol, say $\natural$, if in the construction of A (from atomic formulas) the stage where $\sharp$ is the outermost logical symbol precedes the stage where $\natural$ is the outermost logical symbol. Further, a symbol $\sharp$ is said to be in the left scope of a $\supset$ if $\supset$ occurs in the form $B \supset C$ and $\sharp$ occurs in B

A formula is called **prenex** (in prenex form) if no quantifier in it is in the scope of a propositional connective.

A proof without a cut contains only subformulas of the formulas occurring in the end-sequent. A formula is provable iff it is provable by use of its subformulas only

Theorem 1.30 (consistency)**.** ***LK** and **LJ** are consistent*

Proof. Suppose $\rightarrow$ were provable in **LK**. Then by the cut-elimination theorem, it would be provable in **LK** without a cut. But this is impossible, by the subformula property of cut-free proofs $\square$

Theorem 1.31. *In a cut-free proof in **LK** (or **LJ**) all the formulas which occur in it are subformulas of the formulas in the end-sequent*

Theorem 1.32 (Gentzen's midsequent theorem for **LK**)**.** *Let S be a sequent which consists of prenex formulas only and is provable in **LK**. Then there is a cut-free proof of S which contains a sequent (called a **midsequent**), say S', which satisfies the following*

1. *S' is quantifier-free*

2. *Every inference above S' is either structural or propositional*

3. *Every inference below S' is either structural or a quantifier inference*

Thus a midsequent splits the proof into an upper part, which contains the propositional inferences, and a lower part, which contains the quantifier inferences.

*The above holds reading "**LJ** without $\vee$left" in place of **LK***

outline. Combining Proposition 1.16 and the cut-elimination theorem we may assume that there is a cut-free proof of S, say P, in which all the initial sequents consist of atmoic formulas only ($_{\text{why}}$ do we need atomic formula_). Let I be a quantifier inference in P. The number of propositional inference under I is called the order of I. The sum of orders for all the quantifier inferences in $P$is called the order of P. The proof is carried out by induction on the order of P.

Case 1: The order of a proof P is 0. If there is a propositional inference, take the lowermost such, and call its lower sequent S_0. Above this sequent there is no quantifier inference. Therefore if there is a quantifier in or above S_0, then it is introduced by weakening. Since the proof is cut-free, the weakening formula is a subformula of one of the formulas in the end-sequent. Hence no propositional inferences apply to it. (Since its in prenex form!) We can thus eliminate these weakenings and obtain a sequent S_0' corresponding to S_0. By adding some weakenings under S_0' we derive S and S_0' serves as the mid-sequent

If there is no propositional inference in P, then take the uppermost quantifier inferences. Its upper sequent serves as a midsequent

Case 2: The order of P is not 0. Then there is at least one propositional inference which is below a quantifier property. Moreover, there is a quantifier inference I with the following property: the uppermost logical inference under I is a propositional inference. Call it I'. We can lower the order by interchanging the positions of I and I'. Say I is $\forall$right, then proof P is

$$\begin{array}{c}\vdots\\ \dfrac{\Gamma \to \Theta, F(a)}{\Gamma \to \Theta, \forall x F(x)}\ I\\ \vdots\ (*)\\ \dfrac{\qquad}{\Delta \to \Lambda}\ I'\end{array}$$

where the (*)-part of P contains only structural inferences and Λ contains $\forall x F(x)$ as a sequent-formula. Transform P into the following proof P':

$$\begin{array}{c}\Gamma \to \Theta, F(a)\\ \vdots\ \text{structural inferences}\\ \Gamma \to F(a), \Theta, \forall x F(x)\\ \vdots\\ \dfrac{\qquad}{\Delta \to F(a), \Lambda}\ I'\\ \hline\hline \Delta, \Lambda, \forall x F(x)\ \ I\\ \hline\hline \Delta \to \Lambda\\ \vdots\end{array}$$

It is obvious that the order of P' is less than that of P $\square$

For technical reasons we introduce the predicate symbol $\top$ with 0 argument places, and admit $\rightarrow \top$ as an additional initial sequent. The system which is obtained from **LK** thus extended is denoted by **LK#**

Lemma 1.33. *Let $\Gamma \rightarrow \Delta$ be **LK**-provable, and let (Γ_1, Γ_2) and Δ_1, Δ_2 be arbitrary partitions of Γ and Δ, respectively (including the cases that one or more of $\Gamma_1, \Gamma_2, \Delta_1, \Delta_2$ are empty). We denote such a partition by $[\{\Gamma_1; \Delta_1\}, \{\Gamma_2; \Delta_2\}]$ and call it a partition of the sequent $\Gamma \rightarrow \Delta$. Then there exists a formula C of **LK#** (called an **interpolant** of $[\{\Gamma_1; \Delta_1\}, \{\Gamma_2; \Delta_2\}]$) s.t.*

1. *$\Gamma_1 \rightarrow \Delta_1, C$ and $C, \Gamma_2 \rightarrow \Delta_2$ are both **LK#**-provable*
2. *All free variables and individual and predicate constants in C (apart from $\top$) occur both in $\Gamma_1 \cup \Delta_1$ and $\Gamma_2 \cup \Delta_2$*

Theorem 1.34 (Craig's interpolation theorem for **LK**)**.** 1. *Let A and B be two formulas s.t. $A \supset B$ is **LK**-provable. If A and B have at least one predicate constant in common, then there exists a formula C, called an interpolant of $A \supset B$ s.t. C contains only those individual constants, predicate constants and free variables that occur in both A and B and s.t. $A \supset C$ and $C \supset B$ are **LK**-provable. If A and B contain no predicate constant in common, then either $A \rightarrow$ or $\rightarrow B$ is **LK**-provable*

2. *As above, with **LJ** inplace of **LK***

Proof. Assume that $A \supset B$, and hence $A \rightarrow B$ is provable, and A and B have at least one predicate constant in common. Then by Lemma 1.33, taking A as Γ_1 and B as Δ_2 (with Γ_2 and Δ_1 empty), there exists a formula Csatisfying 1 and 2. So $A \rightarrow C$ and $C \rightarrow B$ are **LK#**-provable. Let R be predicate constant which is common to A and B and has k argument places. Let R' be $\forall y_1 \ldots \forall y_k R(y_1, \ldots, y_k)$, where $y_1, \ldots, y_k$ are new bound variables. By replacing $\top$ by $R' \supset R'$ we can transform C into a formula C' of the original language, s.t. $A \rightarrow C'$ and $C' \rightarrow B$ are **LK**-provable. C' is then the desired interpolant.

If there is no predicate common to $\Gamma_1 \cup \Delta_1$ and $\Gamma_2 \cup \Delta_2$ in the partition, then by Lemma 1.33 there is a C s.t. $\Gamma_1 \rightarrow \Delta_1, C$ and $C, \Gamma_2 \rightarrow \Delta_2$ are provable, and C consists of $\top$ and logical symbols only. Then it can easily be shown, by induction on the complexity of C, that either $\rightarrow C$ or $C \rightarrow$ is provable. Hence either $\Gamma_1 \rightarrow \Delta_1$ or $\Gamma_2 \rightarrow \Delta_2$ is provable. $\square$

Lemma [**?**]. The lemma is proved by induction on the number of inferences k, in a cut-free proof of $\Gamma \to \Delta$. At each stage there are several cases to consider; we deal with some examples only.

1. $k = 0, \Gamma \to \Delta$ has the form $D \to D$. There are four cases: 1. $[\{D; D\}, \{;\}]$, 2. $[\{;\}, \{D; D\}]$, 3. $[\{D;\}, \{; D\}]$, 4. $\{; D\}, \{D;\}$. Take for $C : \neg\top$ in 1, $\top$ in 2, D in 3 and $\neg D$ in 4

2. $k > 0$ and the last inference is $\wedge$right:

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

Suppose the partition is $[\{\Gamma_1; \Delta_1, A \wedge B\}, \{\Gamma_2; \Delta_2\}]$. Consider the induced partition of the upper sequents, viz $[\{\Gamma_1; \Delta_1, A\}, \{\Gamma_2; \Delta_2\}]$ and $[\{\Gamma_1; \Delta_1, B\}, \{\Gamma_2; \Delta_2\}]$ respectively. By the induction hypothesis applied to the subproofs of the upper sequents, there exists interpolants C_1 and C_2 so that $\Gamma_1 \to \Delta_1, A, C_1$;$C_1, \Gamma_2 \to \Delta_2$;$\Gamma_1 \to \Delta_1, B, C_2$ and $C_2, \Gamma_2 \to \Delta_2$ are all **LK#**-provable. From these sequents, $\Gamma_1 \to \Delta_1, A \wedge B, C_1 \vee C_2$ and $C_1 \vee C_2, \Gamma_2 \to \Delta_2$

3. $k > 0$ and the last inference is $\forall$left

$$\frac{F(s), \Gamma \to \Delta}{\forall x F(x), \Gamma \to \Delta}$$

Suppose $b_1, \dots, b_n$ are all the free variables and constants which occur in s. Suppose the partition is $[\{\forall x F(x), \Gamma_1; \Delta_1\}, \{\Gamma_2; \Delta_2\}]$. Consider the induced partition of the upper sequent and apply the induction hypothesis. So there exists and interpolant $C(b_1, \dots, b_n)$ so that

$$F(s), \Gamma_1 \to \Delta_1, C(b_1, \dots, b_n)$$
$$C(b_1, \dots, b_n), \Gamma_2 \to \Delta_2$$

are **LK#**-provable. Let $b_{i_1}, \dots, b_{i_m}$ be all the variables and constants among $b_1, \dots, b_n$ which do not occur in $\{F(x), \Gamma_1; \Delta_1\}$. Then

$$\forall y_1 \dots \forall y_m C(b_1, \dots, y_1, \dots, y_m, \dots, b_n)$$

where $b_{i_1}, \dots, b_{i_m}$ are replaced by the bound variables, serve as the required interpolant.

4. $k > 0$ and the last inference is $\forall$right

$$\frac{\Gamma \rightarrow \Delta, F(a)}{\Gamma \rightarrow \Delta, \forall x F(x)}$$

where a doesn't occur in the lower sequent.

Suppose the partition is $[\{\Gamma_1; \Delta_1, \forall x F(x)\}, \{\Gamma_2; \Delta_2\}]$. By the induction hypothesis there exists an interpolant C so that $\Gamma_1 \rightarrow \Delta_1, F(a), C$ and $C, \Gamma_2 \rightarrow \Delta_2$ are provable. Since C doesn't contain a, we can derive

$$\Gamma_1 \rightarrow \Delta_1, \forall x F(x), C$$

and hence C serves as the interpolant

□

Exercise 1.6.1. Let A and B be prenex formulas which have only $\forall$ and $\wedge$ as logical symbols. Assume futhermore that there is at least one predicate constant common to A and B. Suppose $A \supset B$ is provable.

Show that there exists a formula C s.t.

1. $A \supset C$ and $C \supset B$ are provable
2. C is a prenex formula
3. the only logical symbols in C are $\forall$ and $\wedge$
4. the predicate constants in C are common to A and B

Definition 1.35. 1. A **semi-term** is an expression like a term, except that bound variables are allowed in its construction. Let t be a term and s a semi-term. We call s a **sub-semi-term** of t if

   (a) s contain a bound variable (s is not a term)
   (b) s is not a bound variable itself
   (c) some subterm of t is obtained from s by replacing all the bound variables in s by appropriate terms

2. A **semi-formula** is an expression like a formula, except that bound variables are (also) allowed to occur free in it

Theorem 1.36. *Let t be a term and S a provable sequent satisfying*

$$\text{There is no sub-semi-term of } t \text{ in } S \tag{1}$$

Then the sequent which is obtained from S by replacing all the occurrences of t in S by a free variable is also provable

Proof. Consider a cut-free regular proof of S, say P. If 1 holds for the lower sequent of an inference in P then it holds for the upper sequents. The theorem follows by mathematical induction on the number of inferences in P $\square$

Definition 1.37. Let $R_1, \dots, R_m, R$ be predicate constants. Let $A(R, R_1, \dots, R_m)$ be a sentence in which all occurrences of $R, R_1, \dots, R_m$ are indicated. Let R' be a predicate constant with the same number of argument-places as R. Let B be $\forall x_1 \dots \forall x_k (R(x_1, \dots, x_k) \equiv R'(x_1, \dots, x_k))$, where the string of quantifiers is empty if $k = 0$. Let C be $A(R, R_1, \dots, R_m) \wedge A(R', R_1, \dots, R_m)$. We say that $A(R, R_1, \dots, R_m)$ **defines (in LK)** R **implicitly** in terms of $R_1, \dots, R_m$ if $C \supset B$ is (**LK**-)provable and we say that $A(R, R_1, \dots, R_m)$ **defines (in LK)** R **explicitly** in terms of $R_1, \dots, R_m$ and the individual constants in $A(R, R_1, \dots, R_m)$ if there exists a formula $F(a_1, \dots, a_k)$ containing only the predicate constants $R_1, \dots, R_m$ and the individual constants in $A(R, R_1, \dots, R_m)$ s.t.

$$A(R, R_1, \dots, R_m) \to \forall x_1 \dots \forall x_k (R(x_1, \dots, x_k) \equiv F(x_1, \dots, x_k))$$

is **LK**-provable

Proposition 1.38 (Beth's definability theorem for **LK**). *If a predicate constant R is defined implicitly in terms of $R_1, \dots, R_m$ by $A(R, R_1, \dots, R_m)$, then R can be defined explicitly in terms of $R_1, \dots, R_m$ and the individual constants in $A(R, R_1, \dots, R_m)$*

outline. Let $c_1, \dots, c_n$ be free variables not occurring in A. Then

$$A(R, R_1, \dots, R_m), A(R', R_1, \dots, R_m) \to R(c_1, \dots, c_n) \equiv R'(c_1, \dots, c_n)$$

and hence also

$$A(R, R_1, \dots, R_m) \wedge R(c_1, \dots, c_k) \to A(R', R_1, \dots, R_m) \supset R'(c_1, \dots, c_n)$$

are provable. Now apply Craig's theorem to the latter sequent. We get

$$A(R, R_1, \dots, R_m) \wedge R(c_1, \dots, c_k) \supset F(c_1, \dots, c_k)$$
$$F(c_1, \dots, sc_k) \supset A(R', R_1, \dots, R_m) \supset R'(c_1, \dots)$$

First line implies $A(R, R_1, \dots, R_m) \to R(c_1, \dots, c_k) \supset F(c_1, \dots, c_k)$. The second line with the assumption $A(R, R_1, \dots, R_m)$ shows that $A(R, R_1, \dots, R_m) \to F(c_1, \dots, c_k) \supset R(c_1, \dots, c_k)$ $\square$

Proposition 1.39 (Robinson). *Assume that the language contains no function constants. Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two consistent axiom systems. Suppose furthermore that, for any sentence A which is dependent on $\mathcal{A}_1$ and $\mathcal{A}_2$, it is not the case that $\mathcal{A}_1 \to A$ and $\mathcal{A}_2 \to \neg A$ are provable. Then $\mathcal{A}_1 \cup \mathcal{A}_2$ is consistent*

Proof. Suppose $\mathcal{A}_1 \cup \mathcal{A}_2$ is not consistent. Then there are axiom sentences Γ_1 and Γ_2 from $\mathcal{A}_1$ and $\mathcal{A}_2$ respectively s.t. $\Gamma_1, \Gamma_2 \to$ is provable. Since $\mathcal{A}_1$ and $\mathcal{A}_2$ are each consistent, neither Γ_1 nor Γ_2 is empty. Apply Lemma 1.33 to the partition $[\{\Gamma_1;\}, \{\Gamma_2;\}]$ □

Let **LK′** and **LJ′** denote the quantifier-free parts of **LK** and **LJ**

Theorem 1.40. *There exist decision procedures for **LK′** and **LJ′***

Proof. The following decision procedure was given by gentzen. A sequent of **LK′** (or **LJ′**) is said to be **reduced** if in the antecedent the same formula does not occur at more than three places as sequent formulas, and likewise in the succedent. A sequent S' is called a **reduct** of a sequent S is S' is reduced and is obtained from S by deleting some occurrences of formulas. Now given a sequent S of **LK′** (or **LJ′**), let S' be any reduct of S. We note the following

1. S is provable or unprovable according as S' is provable or unprovable

2. The number of all reduced sequents which contain only subformulas of the formula in S is finite

Consider the finite system of sequents as in 2, say $\mathcal{I}$. Collect all initial sequents in the systems. Call this set $\mathcal{I}_0$. Then examine $\mathcal{I} - \mathcal{I}_0$ to see if there is a sequent which can be the lower sequent of an inference whose upper sequent(s) is (are) one (two) sequent(s) from $\mathcal{I}_0$. Call the set of all sequents which satisfy this condition $\mathcal{I}_1$. Now see if there is a sequent in $(\mathcal{I} - \mathcal{I}_0) - \mathcal{I}_1$ which be the lower sequent of an inference whose upper sequent(s) is (are) one (two) of the sequent(s) in $\mathcal{I}_0 \cup \mathcal{I}_1$. Continue this process until either the sequent S' itself is determined as provable, or the process does not give any new sequent as provable. One of the two must happen. (Note that the whole argument is finitary) □

Theorem 1.41 (Harrop). *1. Let Γ be a finite sequence of formulas s.t. in each formula of Γ every occurrence*

of $\vee$ and $\exists$ is either in the scope of a $\neg$ or in the left scope of a sup. *This condition will be referred to as* (*) *in this theorem.*

1. *Then* $\Gamma \to A \vee B$ *is* ***LJ****-provable iff* $\Gamma \to A$ *and* $\Gamma \to B$ *is* ***LJ****-provable*

2. $\Gamma \to \exists x F(x)$ *is* ***LJ****-provable iff for some term* s, $\Gamma \to F(s)$ *is* ***LJ****-provable*

1. *The following sequents (which are* ***LK****-provable) are not* ***LJ****-provable*

$$\neg(\neg A \wedge \neg B) \to A \vee B; \qquad \neg\forall x \neg F(x) \to \exists x F(x)$$
$$A \supset B \to A \vee B; \qquad \neg\forall x F(x) \to \exists x \neg F(x);$$
$$\neg A(\wedge B) \to A \vee \neg B$$

Proof. 1. (a) $\Rightarrow$. Consider a cut-free proof of $\Gamma \to A \vee B$. The proof is carried out by induction on the number of inferences below all the inferences for $\vee$ and $\exists$ in the given proof. If the last inference is $\vee$right, there is nothing to prove. Notice that the last inference cannot be $\vee, \neg$ or $\exists$left

Case 1: The last inference is $\wedge$left

$$\frac{C, \Gamma \to A \vee B}{C \wedge D, \Gamma \to A \vee B}$$

Its obvious that C satisfies the condition (*). Thus the induction hypothesis applies to the upper sequent; hence either $C, \Gamma \to A$ or $C, \Gamma \to B$ is provable. In either case, the end-sequent can be derived in **LJ** Case 2: The last inference is $\supset$left

$$\frac{\Gamma \to C \qquad D, \Gamma \to A \vee B}{C \supset D, \Gamma \to A \vee B}$$

D satisfies the condition; thus by the induction hypothesis applied to the right upper sequent, $D, \Gamma \to A$ or $D, \Gamma \to B$ is provable.

(b) If $\Gamma \to F(s)$ is **LJ**-provable for some term s.

□

1.7 The predicate calculus with equality **PROBLEM**

Definition 1.42. The predicate calculus with equality (denoted $\mathbf{LK_e}$) can be obtained from **LK** by specifying constant of two argument (=: read equals) and adding the following sequents as additional initial sequents ($a = b$ denoting $= (a, b)$)

$$\to s = s$$
$$s_1 = t_1, \dots, s_n = t_n \to f(s_1, \dots, s_n) = f(t_1, \dots, t_n)$$

for every function constant f of n argument-places $(n = 1, 2, ...)$:

$$s_1 = t_1, \dots, s_n = t_n, R(s_1, \dots, s_n) \to R(t_1, \dots, t_n)$$

for every predicate constant R of n argument; where $s, s_1, \dots, s_n, t_1, \dots, t_n$ are arbitrary terms

Each such sequent may be called an equality axiom of $\mathbf{LK_e}$

Proposition 1.43. *Let $A(a_1, \dots, a_n)$ be an arbitrary formula. Then*

$$s_1 = t_1, \dots, s_n = t_n, A(s_1, \dots, s_n) \to A(t_1, \dots, t_n)$$

is provable in $\mathbf{LK_e}$ for any terms s_i, t_i. Furthermore, $s = t \to t = s$ and $s_1 = s_2, s_2 = s_3 \to s_1 = s_3$ are also provable

Definition 1.44. Let Γ_e be the set (axiom system) consisting of the following sentences

$$\forall x(x = x)$$
$$\forall x_1 \dots \forall x_n \forall y_1 \dots \forall y_n [x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset f(x_1, \dots, x_n = f(y_1, \dots, y_n))]$$

for every function constant f with n arguments,

$$\forall x_1 \dots \forall x_n \forall y_1 \dots \forall y_n [x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset R(x_1, \dots, x_n = R(y_1, \dots, y_n))]$$

for every predicate constant R of n arguments. Each such sentence is called an **equality axiom**

Proposition 1.45. *A sequent $\Gamma \to \Delta$ is provable in $\mathbf{LK_e}$ iff $\Gamma, \Gamma_e \to \Delta$ is provable in $\mathbf{LK}$*

Proof. All the initial sequents of $\mathbf{LK_e}$ are provable from Γ_e $\square$

Definition 1.46. If the cut formula of a cut in $\mathbf{LK_e}$ is of the form $s = t$, then the cut is called **inessential**. It's called **essential** otherwise

Theorem 1.47 (the cut-elimination theorem for $\mathbf{LK_e}$)**.** *If a sequent of $\mathbf{LK_e}$ is $\mathbf{LK_e}$-provable, then it is $\mathbf{LK_e}$-provable without an essential cut*

Proof. The theorem is proved by removing essential cuts (mixes as a matter of a fact), following the method used for Theorem 1.24

If the rank is 2, S_2 is an equality axiom and the mix formula is not of the form $s = t$, then the mix formula is of the form $P(t_1, \dots, t_n)$. If S_1 is also an equality axiom, then it has the form

$$s_1 = t_1, \dots, s_n = t_n, P(s_1, \dots, s_n) \to P(t_1, \dots, t_n)$$

From this and S_2, i.e.,

$$t_1 = r_1, \dots, t_n = r_n, P(t_1, \dots, t_n) \to P(r_1, \dots, r_n)$$

we obtain by a mix

$$s_1 = t_1, \dots, s_n = t_n, t_1 = r_1, \dots, t_n = r_n, P(s_1, \dots, s_n) \to P(r_1, \dots, r_n)$$

This may be replaced by

$$s_i = t_i, t_i = r_i \to s_i = r_i \quad (i = 1, 2, \dots, n)$$
$$s_1 = r_1, \dots, s_n = r_n, P(s_1, \dots, s_n) \to P(r_1, \dots, r_n)$$

and then repeated cuts of $s_i = r_i$ to produce the same end-sequent. All cuts introduced here are inessential

If $P(t_1, \dots, t_n)$ in S_2 is a weakening formula, then the mix inference is

$$\frac{s_1 = t_1, \dots, s_n = t_n, P(s_1, \dots, s_n) \to P(t_1, \dots, t_n) \qquad P(t_1, \dots, t_n), \Pi \to \Lambda}{s_1 = t_1, \dots, s_n = t_n, P(s_1, \dots, s_n), \Pi \to \Lambda}$$

Transform this into

$$\frac{\Pi \to \Lambda}{\overline{\text{end-sequent}}}$$

□

Exercise 1.7.1. A sequent of the form

$$s_1 = t_1, \dots, s_n = t_n \to s = t$$

is said to be simple if it is obtained from sequents of the following four forms by applications of exchanges, contractions, cuts, and weakening left.

1. $\to s = s$
2. $s = t \to t = s$
3. $s_1 = s_2, s_2 = s_3 \to s_1 = s_3$
4. $s_1 = t_1, \dots, s_m = t_m \to f(s_1, \dots, s_m) = f(t_1, \dots, t_m)$

Prove that if $s_1 = s_1, \dots, s_m = s_m \to s = t$ is simple, then $s = t$ is of the form $s = s$. As a special case, if $\to s = t$ is simple, then $s = t$ is of the form $s = s$

Let $\mathbf{LK}'_{\mathbf{e}}$ be the system which is obtained from **LK** adding the following sequents as initial sequents

1. simple sequents
2. sequents of the form

$$s_1 = t_1, \dots, s_m = t_m, R(s'_1, \dots, s'_n) \to R(t'_1, \dots, t'_n)$$

   where $s_1 = t_1, \dots, s_m = t_m \to s'_i = t'_i$ is simple for each i

First prove that the initial sequents of $\mathbf{LK}'_{\mathbf{e}}$ are closed under cuts and that if

$$R(s_1, \dots, s_n) \to R(t_1, \dots, t_n)$$

is an initial sequent of $\mathbf{LK}'_{\mathbf{e}}$ (where R is not $=$), then it is of the form $D \to D$. Finally prove that the cut-elimination theorem (without the exception of inessential cuts) holds for $\mathbf{LK}'_{\mathbf{e}}$

Proof. 1. Consider the complexity of s?

If s is a variable, we can only get this by $v_i = v_i$

□

1.8 The completeness theorem

Definition 1.48. 1. Let L be a language. By a **structure** for L we mean a pair $\langle D, \phi \rangle$, where D is a non-empty set and ϕ is a map from the constants of L s.t.

   (a) if k is an individual constant, then ϕk is an element of D

   (b) if f is a function constant of n arguments, then ϕf is a mapping from D^n to D

   (c) if R is a predicate constant of n arguments, then ϕR is a subset of D^n

2. An **interpretation** of L is a structure $\langle D, \phi \rangle$ together with a mapping ϕ_0 from variables into D. We may denote an interpretation $(\langle D, \phi \rangle, \phi_0)$ simply by $\mathfrak{F}$. ϕ_0 is called an assignment from D

3. We say that an interpretation $\mathfrak{F} = (\langle C, \phi \rangle, \phi_0)$ **satisfies** a formula A if this follows from the following inductive definition

   (a) For every semi-term t, $\phi(a) = \phi_0(a)$ and for a$\phi(x) = \phi_0(x)$ll free variables a and bound variables x. next if f is a function constant and t is a semi-term for which ϕt is already defined, then $\phi(f(t))$ is defined to be $(\phi f)(\phi t)$

Theorem 1.49 (Completeness and soundness)**.** *A formula is provable in **LK** iff it is valid*

Lemma 1.50. *Let S be a sequent. Then either there is a cut-free proof of S, or there is an interpretation which does not satisfy S (and hence S is not valid)*

Proof. We will define, for each sequent S, a (possibly infinite) tree, called the reduction tree for S, from which we can obtain either a cut-free proof of S or an interpretation not satisfying S. This reduction tree for S contains a sequent at each node. It is constructed in stages as follows

Stage 0: Write S at the bottom of the tree

Stage k ($k > 0$): This is defined by cases

1. Every topmost sequent has a formula common to its antecedent and succedent. Then stop.

2. This stage is defined according as

$$k \equiv 0, 1, 2, \dots, 12 \mod 13$$

   $k \equiv 0$ and $k \equiv 1$ concern the symbol $\neg$; $k \equiv 2$ and $k \equiv 3$ concern $\wedge$; $k \equiv 4$ and $k \equiv 5$ concern $\vee$; $k \equiv 6$ and $k \equiv 7$ concern $\supset$; $k \equiv 8$ and $k \equiv 9$ concern $\forall$; $k \equiv 10$ and $k \equiv 11$ concern equiv $\exists$

Assume that there are no individual or function constants

All the free variables which occur in any sequent which has been obtained at or before stage k are said to be "available at stage k". In case there is none, pick any free variable and say that it is available

0. $k \equiv 0$. Let $\Pi \to \Lambda$ be any topmost sequent of the tree which has been defined by stage $k-1$. Let $\neg A_1, \dots, \neg A_n$ be all the formulas in Π whose outermost logical symbol is $\neg$, and to which no reduction has been applied in previous stages. Then write down

$$\Pi \to \Lambda, A_1, \dots, A_n$$

   above $\Pi \to \Lambda$. We say that a $\neg$left reduction has been applied to $\neg A_1, \dots, \neg A_n$

1. $k \equiv 1$. Let $\neg A_1, \dots, \neg A_n$ be all the formulas in Λ whose outermost logical symbol is $\neg$ and to which no reduction has been applied so far. Then write down

$$A_1, \dots, A_n, \Pi \to \Lambda$$

   above $\Pi \to \Lambda$. We say that a $\neg$right reduction has been applied to $\neg A_1, \dots, \neg A_n$.

2. $k \equiv 2$. Let $A_1 \wedge B_1, \dots, A_n \wedge B_n$ be all the formulas in Π whose outermost logical symbols is $\wedge$ and to which no reduction has been applied yet. Then write down

$$A_1, B_1, A_2, B_2, \dots, A_n, B_n, \Pi \to \Lambda$$

above $\Pi \to \Lambda$. We say that an $\wedge$left reduction has been applied to

$$A_1 \wedge B_1, \dots, A_n \wedge B_n$$

3. $k \equiv 3$. Let $A_1 \wedge B_1, \dots, A_n \wedge B_n$ be all the formulas in Π whose outermost logical symbols is $\wedge$ and to which no reduction has been applied yet. Then write down

$$\Pi \to \Lambda, C_1, \dots, C_n$$

where C_i is either A_i or B_i, above $\Pi \to \Lambda$. Take all possible combinations of such; so there are 2^n such sequents above $\Pi \to \Lambda$. We say that an $\wedge$right reduction has been applied to $A_1 \wedge B_1, \dots, A_n \wedge B_n$

4. $k \equiv 4$. $\vee$left, similar to 3

5. $k \equiv 5$. $\vee$right, similar to 2.

6. $k \equiv 6$. Let $A_1 \supset B_1, \dots, A_n \supset B_n$ be all the formulas in Π whose outermost symbol is $\supset$ and to which no reduction has been applied yet. Then write down the following sequents above $\Pi \to \Lambda$

$$B_{i_1}, B_{i_2}, \dots, B_{i_k}, \Pi \to \Lambda, A_{j_1}, \dots, A_{j_{n-k}}$$

where $i_1 < i_2 < \cdots < i_k$, $j_1 < j_2 < \cdots < j_{n-k}$ and $(i_1, \dots, i_k, j_1, \dots, j_{n-k})$ is a permutation of $(1, 2, \dots, n)$. Take all possible permutations: so there are 2^n such sequents.

7. $k \equiv 7$. Let $A_1 \supset B_1, \dots, A_n \supset B_n$ be all the formulas in Λ whose outermost logical symbol is $\supset$ and to which no reduction has been applied yet. Then write down

$$A_1, A_2, \dots, A_n, \Pi \to \Lambda, B_1, \dots, B_n$$

above $\Pi \to \Lambda$. We say that an $\supset$right reduction has been applied to

$$A_1 \supset B_1, \dots, A_n \supset B_n$$

8. $k \equiv 8$. Let $\forall x_1 A_1(x_1), \forall x_n A_n(x_n)$ be all the formulas in Π whose outermost logical symbol is $\forall$. let a_i be the first variable available at this stage which has not been used for reduction of $\forall x_i A_i(x$ for $1 \leq i \leq n$. Then write down

$$A_1(a_1), \ldots, A_n(a_n), \Pi \to \Lambda$$

above $\Pi \to \Lambda$. We say that a $\forall$left reduction has been applied to

$$\forall x_1 A_1(x), \ldots, \forall x_n A_n(x_n)$$

9. $k \equiv v$. Let $\forall x_1 A_1(x_1), \ldots, \forall x_n A_n(x_n)$ be all formulas in Λ whose outermost logical symbol is $\forall$ and to which no reduction has been applied so far. Let $a_1, \ldots, a_n$ be the first n free variables which are not available at this stage. Then write down

$$\Pi \to \Lambda, A_1(a_1), \ldots, A_n(a_n)$$

above $\Pi \to \Lambda$. We say that a $\forall$right reduction has been applied to $\forall x_1 A_1(x_1), \ldots, \forall x_n A_n(x_n)$. Notice that $a_1, \ldots, a_n$ are new available free variables

10. $k \equiv 10$. $\exists$left reduction. Similar to 9

11. $k \equiv 11$. $\exists$right reduction. similar to 8

12. If Π and Λ have any formula in common, write nothing above $\Pi \to \Lambda$. If Π and Λ have no formula in common and the reductions described in 0-11 are not applicable, write the same sequent $\Pi \to \Lambda$ again above it.

So the collection of those sequents which are obtained by the above reduction process, together with the partial order obtained by this process, is the reduction tree (for S). It is denoted by $T(S)$. We will construct "reduction trees" like this again

As an example of the case where the reduction process does not terminate, consider a sequent of the form $\forall x \exists y A(x,y) \to$, where A is a predicate

Now a (finite or infinite) sequence $S_0, S_1, \ldots$ of sequents in $T(S)$ is called a branch if

1. $S_0 = S$

2. S_{i+1} stands immediately above S_i

3. if the sequence is finite, say $S_1, \dots, S_n$, then S_n has the form $\Pi \to \Lambda$, where Π and Λ have a formula in common

Now given a sequent S, let T be the reduction tree $T(S)$. If each branch of T ends with a sequent whose antecedent and succedent contain a formula in common, then it is a routine task to write a proof without a cut ending with S by suitably modifying T. Otherwise there is an infinite branch. Consider such a branch consisting of sequents $S = S_0, S_1, \dots, S_n, \dots$

Let S_i be $\Gamma_i \to \Delta_i$. let $\bigcup \Gamma$ be the set of all formulas occurring in Γ_i for some i, and let $\bigcup \Delta$ be the set of all formulas occurring in Δ_j for some j. We shall define an interpretation in which every formula in $\bigcup \Gamma$ holds and no formula in $\bigcup \Delta$ holds. Thus S does not hold in it.

First notice that from the way the branch was chosen, $\bigcup \Gamma$ and $\bigcup \Delta$ have no atomic formula in common. Let D be the set of all the free variables. We consider the interpretation $\mathfrak{F} = (\langle D, \phi \rangle, \phi_0)$ where ϕ and ϕ_0 are defined as follows: $\phi_0(a) = a$ for all free variables a, $\phi_0(x)$ is defined arbitrarily for all bound variables x. For an n-ary predicate constant R, ϕR is any subset of D^n s.t.: if $R(a_1, \dots, a_n) \in \bigcup \Gamma$, then $(a_1, \dots, a_n) \in \phi R$, and $(a_1, \dots, a_n) \notin \phi R$

We claim that this interpretation $\mathfrak{F}$ has the required property: it satisfies every formula in $\bigcup \Gamma$, but no formula in $\bigcup \Delta$. We prove this by induction on the number of logical symbols in the formula A. We consider here only the case where A is of the form $\forall x F(x)$ and assume the induction hypothesis. For the base case, note that $\bigcup \Gamma \cap \bigcup \Delta = \emptyset$.

1. A is in $\bigcup \Gamma$. Let i be the least number s.t. A is in Γ_i. Then A is in Γ_j for all $j > i$. It is sufficient to show that all substitution instances $A(a)$, for $a \in D$, are satisfied by $\mathfrak{F}$.

2. A is in $\bigcup \Delta$. Consider the step at which A was used to define an upper sequent from $\Gamma_i \to \Delta_i$. It looks like

$$\frac{\Gamma_{i+1} \to \Delta^1_{i+1}, F(a), \Delta^2_{i+1}}{\Gamma_1 \to \Delta^1_i, A, \Delta^2_1}$$

Then by the induction hypothesis, $F(a)$ is not satisfied by $\mathfrak{F}$, so A is not satisfied by $\mathfrak{F}$ either.

$\square$

Exercise 1.8.1. Feferman Let J be a non-empty set. Each element of J is called a **sort**. A many-sorted language for the set of sorts J, say $L(J)$, consists of the following

1. Individual constants: $k_0, k_1, \dots, k_i, \dots$, where to each k_i is assigned one sort
2. Predicate constants: $R_0, R_1, \dots, R_i, \dots$, where to each R_i is assigned a number $n \geq 0$ and sorts $j_1, \dots, j_n$. We say that $(n; j_1, \dots, j_n)$ is assigned to R_i
3. Function constants: $f_0, \dots, f_i, \dots$ where to each f_i is assigned a number $n \geq 1$ and sorts $j_1, \dots, j-n, j$. We say that $(n; j_1, \dots, j_n, j)$ is assigned to f_i.
4. Free variables of sort j for each j in J: $a^j_0, a^j_1, \dots, a^j_i, \dots$
5. Bound variables of sort j for each j in J
6. Logical symbols: $\neg, \wedge, \vee, \supset, \forall, \exists$

Terms of sort j for each j are defined as follows. Individual constants and free variables of sort j are terms of sort j; if f is a function constant with $(n; j_1, \dots, j_n, j)$ assigned to it and $t_1, \dots, t_n$ are terms of sort $j_1, \dots, j_n$, respectively, then $f(t_1, \dots, t_n)$ is a term of sort j

If R is a predicate constant with $(n; j_1, \dots, j_n)$ assigned to it and $t_1, \dots, t_n$ are terms of sort $j_1, \dots, j_n$, respectively, then $R(t_1, \dots, t_n)$ is an atomic formula. If $F(a^j)$ is a formula and x^j does not occur in $F(a^j)$ then $\forall x^j F(x_j)$ and $\exists x^j F(x^j)$ are formulas.

The rules of inference are those of **LK**, except that in the rules for $\forall$ and $\exists$, terms and free variables must be replaced by bound variables of the same sort

Prove the following

1. The cut-elimination theorem holds for the system just defined

Sort(A) is the set of j in J s.t. a symbol of sort j occurs in A; Ex(A) and Un(A) are the sets of sorts of bound variables which occur in some essentially existential, respectively universal quantifier in A. (An occurrence of $\exists$, say $\sharp$, is said to be **essentially existential** or **universal** according to the following definition. Count the number of $\neg$ and $\supset$ in A s.t. $\sharp$ is either in the scope of $\neg$, or in the left scope of $\supset$. If this number is even, then $\sharp$ is essentially existential in A, while if it is odd then $\sharp$ is essentially universal. We define dually for $\forall$) . Fr(A) is the set of free variables in A. Pr(A) is the set of predicate constants in A

2. Suppose $A \supset B$ is provable in the above system and at least one of $\mathrm{Sort}(A) \cap \mathrm{Ex}(B)$ and $\mathrm{Sort}(B) \cap \mathrm{Un}(A)$ is not empty. Then there is a formula C s.t. $\sigma(C) \subseteq \sigma(A) \cap \sigma(B)$ where σ sands for Fr, Pr or Sort, and s.t. $\mathrm{Un}(C) \subseteq \mathrm{Un}(A)$ and $\mathrm{Ex}(C) \subseteq \mathrm{Ex}(B)$

Definition 1.51. Let $L(J)$ be a many-sorted language. A structure for $L(J)$ is a pair $\langle D, \phi \rangle$ where D is a set of non-empty sets $\{D_j : j \in J\}$ and ϕ is a map from the constants of $L(J)$ into appropriate objects. We call D_j the domain of the structure of sort j. An individual constant of sort j is a member of D_j. Let $\mathcal{M} = \langle D, \phi \rangle$ and $\mathcal{M}' = \langle D', \phi' \rangle$ be two structures for $L(J)$. We say $\mathcal{M}'$ is an extension of $\mathcal{M}$ and write $\mathcal{M} \subseteq \mathcal{M}'$ if

1. for each $j \in J$, $D_j \subseteq D'_j$
2. for each individual constant k, $\phi' k = \phi k$
3. for each predicate constant R with $(n; j_1, \ldots, j_n)$ assigned to it

$$\phi R = \phi' R \cap (D_{j_1} \times \cdots \times D_{j_n})$$

4. for each constant constant f with $(n; j_1, \ldots, j_n, j)$ assigned to it and $(d_1, \ldots, d_n) \in D_{j_1} \times \cdots \times D_{j_n}$

$$(\phi' f)(d_1, \ldots, d_n) = (\phi f)(d_1, \ldots, d_n)$$

A formula is said to be **existential** if $\mathrm{Un}(A)$ is empty

Corollary 1.52 (Łoś-Tarski)**.** *The following are equivalent: let A be a formula of an ordinary (i.e., single-sorted) language L*

1. *For any structure $\mathcal{M}$ (for L) and extension $\mathcal{M}'$, and any assignments ϕ, ϕ' from the domains of $\mathcal{M}, \mathcal{M}'$, respectively, which agree on the free variables of A, if $(\mathcal{M}, \phi)$ satisfies A, then so does $(\mathcal{M}', \phi')$*
2. *There exists an (essentially) existential formula B s.t. $A \equiv B$ is provable and the free variable of B are among those of A*

Feferman. We assume (for simplicity) that the language has no individual and function constants.

Let $\mathcal{M}$ and $\mathcal{M}'$ be two structures of the form

$$\mathcal{M} = \langle D_1, \{R_i\}_{i \in I} \rangle, \quad \mathcal{M}' = \langle D_2, \{R'_i\}_{i \in I} \rangle$$

Let J be $\{1,2\}$. $(J, I, \langle k_i \rangle_{i \in I})$ will determine a 'type' of structures. Let L^+ be a corresponding language. It contains the original language L as the sublanguage of sort 1. For each bound variable u, the nth bound variable of sort 1, let u' be the nth bound variable of sort 2. If C is an L-formula, then C' denotes the result of replacing each bound variable u in C by u'; hence $\mathrm{Fr}(C) = \mathrm{Fr}(C')$. With this notation, define Ext to be the form $\forall u' \exists u (u' = u)$. Then

$$\mathrm{Ext}, \{\exists u_i'(u_i' = b_i)\}_{i=1}^n, A' \to A$$

□

Definition 1.53. Let R be a set and suppose a set W_p is assigned to every $p \in R$. If $R_1 \subseteq R$ and $f \in \prod_{p \in R_1} W_p$, then f is called a **partial function** (over R) with domain $\mathrm{dom}(f) = R_1$. If $\mathrm{dom}(f) = R$ then f is called a **total function** (over R). If f and g are partial functions and $\mathrm{dom}(f) = D_0 \subseteq \mathrm{dom}(g)$ and $f(x) = g(x)$ for every $x \in D_0$, then we call g an **extension** of f and write $f \prec g$ and $f = g \restriction D_0$

Proposition 1.54 (a generalized Kőnig's lemma)**.** *Let R be any set. Suppose a finite set W_p is assigned to every $p \in R$. Let P be a property of partial functions f over R satisfying the following conditions:*

1. *$P(f)$ holds iff there exists a finite subset N of R satisfying $P(f \restriction N)$*
2. *$P(f)$ holds for every total function f*

Then there exists a finite subset N_0 of R s.t. $P(f)$ holds for every f with $N_0 \subseteq \mathrm{dom}(f)$.

Note that R can have arbitrarily large cardinality. The case that R is the set of natural numbers is the original Kőnig's lemma.

Proof. Let $X = \prod_{p \in R} W_p$, and give each W_p the discrete topology, and X the product topology. Since each W_p is compact, so is X (Tychonoff's theorem). For each g s.t. $\mathrm{dom}(g)$ is finite, let

$$N_g = \{f \mid f \text{ is total and } g \prec f\}$$

Let

$$C = \{N_g \mid \mathrm{dom}(g) \text{ is finite and } P(g)\}$$

C is an open cover of X. Therefore C has a finite subcover, say

$$N_{g_1}, \ldots, N_{g_k}$$

Let $N_0 = \mathrm{dom}(g_1) \cup \cdots \cup \mathrm{dom}(g_k)$. We will show that N_0 satisfies the condition of the theorem. If $N_0 \subseteq \mathrm{dom}(g)$, then let $g \prec f$, f total. Then $P(f)$ and $f \in N_{g_1} \cup \cdots \cup N_{g_k}$. Say $f \in N_{g_i}$. So $g_i \prec f$, $P(g_i)$ and $g_i \prec g$. Therefore $P(g)$. $\square$

To simplify the discussion, we assume that our language does not contain individual or function constants.

We deal with $\mathbf{LJ}'$. $\mathbf{LJ}'$ is defined by restricting $\mathbf{LK}$ as follows: The inferences $\neg$right, $\supset$right and $\forall$right are allowed only when the principal formulas are the only formulas in the succedents of the lower sequents. (these are called the "critical inferences" of $\mathbf{LJ}'$).

By interpreting a sequent of $\mathbf{LJ}'$, say $\Gamma \to B_1 \ldots, B_n$ as $\Gamma \to B_1 \vee \cdots \vee B_n$, its a routine matter to prove that $\mathbf{LJ}'$ and $\mathbf{LJ}$ are equivalent.

Starting with a given $\Gamma \to \Delta$, we can carry out the reduction process which was defined in Lemma 1.50 except that we omit the stages 1,7,9.

The tree obtained by the above reduction process is called the reduction tree for $\Gamma \to \Lambda$

Definition 1.55. Let Γ and Δ be well-ordered sequences of formulas, which may be infinite. We say that $\Gamma \to \Delta$ is **provable** (in $\mathbf{LJ}'$) if there are finite sequences of Γ and Δ, say $\tilde{\Gamma}$ and $\tilde{\Delta}$, respectively, s.t. $\tilde{\Gamma} \to \tilde{\Delta}$ is provable

2 Peano Arithmetic

2.1 A formulation of Peano arithmetic

Definition 2.1. The language of the system, which will be called Ln, contains finitely many constants, as follows

- Individual constant: 0
- Function constants: $'$,+,$\cdot$
- Predicate constant: = where $'$ is unary while the other constants are binary

A **numeral** is an expression of the form $0'^{\cdots}{}'$, i.e., zero followed by n primes for some n, which is denoted by $\bar{n}$. Further, if s is a closed term of Ln denoting a number m (in the intended interpretation), then $\bar{s}$ denotes the numeral $\bar{m}$ (e.g. if s is $\bar{2} + \bar{3}$ then $\bar{s}$ denotes $\bar{5}$)

Definition 2.2. The first axiom system of Peano arithmetic which we consider, CA, consists of Γ_e for Ln in definition 1.44 and the following sentences

$$
\begin{array}{ll}
\text{A1} & \forall x \forall y (x' = y' \supset x = y) \\
\text{A2} & \forall x (\neg x' = 0) \\
\text{A3} & \forall x (x + 0) = x \\
\text{A4} & \forall x \forall y (x + y' = (x + y)') \\
\text{A5} & \forall x (x \cdot 0 = 0) \\
\text{A6} & \forall x \forall y (x \cdot y' = x \cdot y + x)
\end{array}
$$

The second axiom system of Peano arithmetic which we consider VJ, consists of all sentences of the form

$$\forall z_1 \ldots \forall z_n \forall x (F(0, z) \vee \forall y (F(y, z) \supset F(y', z)) \supset F(x, z))$$

where z is an abbreviation for the sentence of variables $z_1, \ldots, z_n$; and all the variables which are free in $F(x, z)$ are among x, z

The basic logical system of Peano arithmetic is **LK**. Then CA $\cup$ VJ is an axiom system with equality. Furthermore $\forall x \forall y (x = y \supset (F(x) \equiv y))$ is provable for every formula of Ln (cf. Proposition 1.43)

Definition 2.3. The system **PA** (Peano arithmetic) is obtained from **LK** (in the language Ln) by adding extra initial sequents (called the **mathematical initial sequents**) and a new rule of inference called "**ind**", stated below

1. Mathematical initial sequents: additional initial sequents of $\mathbf{LK_e}$ for Ln in Definition 1.42 and the following sequents

$$
\begin{array}{l}
s' = t' \rightarrow s = t \\
s' = 0 \rightarrow \\
\rightarrow s + 0 = s \\
\rightarrow s + t' = (s + t)' \\
\rightarrow s \cdot 0 = 0 \\
\rightarrow s \cdot t' = s \cdot t + s
\end{array}
$$

where s, t, r are arbitrary terms of Ln

2. Ind:

$$\frac{F(a), \Gamma \rightarrow \Delta, F(a')}{F(0), \Gamma \rightarrow \Delta, F(s)}$$

where a is not in $F(0)$, Γ or Δ; s is an arbitrary term (which may contain a); and $F(a)$ is an arbitrary formula of Ln

$F(a)$ is called the **induction formula**, and a is called the **eigenvariable** of this inference. Further, we call $F(a)$ and $F(a')$ the **left** and **right**

auxiliary formula, respectively, and $F(0)$ and $F(s)$ the **left** and **right principal formula**, respectively, of this inference.

The initial sequents of the form $D \to D$ are called **logical** initial sequents

A **weak inference** is a structural inference other than cut.

Proposition 2.4. *A sequent is provable from $CA \cup VJ$(in **LK**) iff it is provable in **PA**. Hence the axiom system $CA \cup VJ$ is consistent iff $\to$ is not provable in **PA***

Thus we can restrict out attention to the system **PA**. In the rest of this chapter, "provability" means provability in **PA**.

Proposition 2.5. *Let P be a proof in **PA** of a sequent $S(a)$, where all the occurrence of a in $S(a)$ are indicated. Let s be an arbitrary term. Then we may construct a **PA**-proof P' of $S(s)$ s.t. P' is regular (cf. Lemma 1.12) and P' differs from P only in that some free variables are replaced by some other free variables and some occurrences of a are replaced by s*

Lemma 2.6.

1. *For an arbitrary closed term s, there exists a unique numeral $\bar{n}$ s.t. $s = \bar{n}$ is provable without an essential cut (Definition 1.46) and without ind*
2. *Let s and t be closed terms. Then either $\to s = t$ or $s = t \to$ is provable without an essential cut or ind*
3. *Let s and t be closed terms s.t. $s = t$ is provable without an essential cut or ind and let $q(a)$ and $r(a)$ be two terms with some occurrences of a (possibly none). Then $q(s) = r(s) \to q(t) = r(t)$ is provable without an essential cut or ind*
4. *Let s and t be as in 3. For an arbitrary formula $F(a) : s = t, F(s) \to F(t)$ is provable without an essential cut or*

Definition 2.7. When we consider a formula or a logical symbol together with the place that it occupies in a proof, in a sequent or in a formula, we refer to it as a formula or a logical symbol in the proof, in the sequent or in the formula. A formula in a sequent is also called a **sequent-formula**

1. If a formula E is contained in the upper sequent of an inference using one of the rules of inference in 1 or "ind", then the **successor** of E is defined as follows
   (a) If E is a cut formula, then E has no successor

(b) If E is an auxiliary formula of any inference other than a cut or exchange, then the principal formula is the successor of E

(c) If E is he formula denoted by C (respectively, D) in the upper sequent of an exchange (in Definition 1.8), then the formula C (respectively, D) in the lower sequent is the successor of E

(d) If E is the kth formula of Γ,Π,Δ or Λ in the upper sequent (in Definition 1.8), then the kth formula of Γ,Π,Δ or Λ, respectively, in the lower sequent is the successor of E

2. A sequent formula is called an **initial formula** or an **end-formula** if it occurs, in an initial sequent or an end-sequent

3. A sequent of formulas in a proof with the following properties is called a **bundle**

   (a) The sequence begins with an initial formula or a weakening formula

   (b) The sequence ends with an end-formula or a cut-formula

   (c) Every formula in the sequence except the last is immediately followed by its successor

4. Let A and B be formulas. A is called an **ancestor** of B and B is called a **descenent** of A if there is a bundle containing both A and B in which A appears above B

5. Let A and B be formulas. If A is the successor of B, then B is called a **predecessor** of A

6. A bundle is called **explicit** if it ends with an end formula

   It is called **implicit** if it ends with a cut-formula

   A formula in a proof is called explicit or implicit according as the bundles containing the formula are explicit or implicit

   A sequent in a proof is called explicit or implicit according as this sequent contains an implicit formula or not

   A logical inference in a proof is called explicit or implicit according as the principal formula of this inference is explicit or implicit

7. The **end-piece** of a proof is defined as follows

   (a) The end-sequent of the proof is contained in the end-piece

(b) The upper sequent of an inference other than an implicit logical inference is contained in the end-piece iff the lower sequent is contained in it

(c) The upper sequent of an implicit logical inference is not contained in it

We can rephrase this definition as follows: A sequent in a proof is in the end-piece of the proof iff there is no implicit inference below this sequent

8. An inference of a proof is said to be **in the end-piece** of the proof if the lower sequent of the inference is in the end-piece

9. Let J be an inference in a proof. We say J **belongs to the boundary** (or J is a **boundary inference**) if the lower seuqent of J is in the end-piece and the upper sequent is not. It should be noted that if *J belongs to the boundary, then it is an implicit logical inference*.

10. A cut in the end-piece is called **suitable** if each cut formula of this cut has an ancestor which is the principal formula of a boundary inference

11. A cut is called **inessential** if the cut formula contains no logical symbol; otherwise it is called **essential**

    In **PA**, the cut formulas of inessential cuts are of the form $s = t$

12. A proof P is **regular** if: 1. the eigenvariables of any two distinct inferences ($\forall$right, $\exists$left or induction) in P are disctinct from each other 2. if a free variable a occurs as an eigenvariable of a sequent S of P, then a only occurs in sequents above S

Proposition 2.8. *For an arbitrary proof of **PA**, there exists a regular proof of the same end-sequent, which can be obtained from the original proof by simply replacing free variables*

Proof. Lemma **??** □

2.2 The Incompleteness Theorem

Definition 2.9. An axiom system $\mathcal{A}$ is said to be **axiomatizable** if there is a finite set of schemata s.t. $\mathcal{A}$ consists of all the instances of these schemata. A formal system **S** is called axiomatizable if there is an axiomatizable axiom system $\mathcal{A}$ s.t. **S** is equivalent to $\mathbf{LK}_{\mathcal{A}}$

A system **S** is called an extension of **PA** if every theorem of **PA** is provable in **S**.

Definition 2.10. The class of primitive recursive functions is the smallest class of functions generated by the following schemata

1. $f(x) = x'$, where $'$ is the successor function
2. $f(x_1, \dots, x_n) = k$, where $n \geq 1$ and k is a natural number
3. $f(x_1, \dots, x_n) = x_i$, where $1 \leq i \leq n$
4. $f(x_1, \dots, x_n) = g(h_1(x_1, \dots, x_n), \dots, h_m(x_1, \dots, x_n))$, where $g, h_1, \dots, h_m$ are primitive recursive functions
5. $f(0) = k, f(x') = g(x, f(x))$ where k is a natural number and g is a primitive recursive function
6. $f(0, x_2, \dots, x_n) = g(x_2, \dots, x_n), f(x', x_2, \dots, x_n) = h(x, f(x, x_2, \dots, x_n), x_2, \dots, x_n)$, where g and h are primitive recursive functions

   An n-ary relation R is said to be primitive recursive if there is a primitive recursive function f which assumes values 0 and 1 only s.t. $R(a_1, \dots, a_n)$ is true iff $f(a_1, \dots, a_n) = 0$

Lemma 2.11. *The consistency of **S** (i.e., **S**-unprovability of* $\rightarrow$*) is equivalent to the **S**-unprovability of* $0 = 1$ *(cf. Proposition 1.21)*

Proposition 2.12 (Gödel)**.** *1. The graphs of all the primitive recursive functions can be expressed in Ln, so that their defining equations are provable in **PA***

*Thus the theory of primitive recursive functions can be translated into our formal system of arithmetic. We may therefore assume that **PA** (or any of its extensions) actually contains the function symbols for primitive recursive functions and their defining equations, as well as predicate symbols for the primitive recursive relations*

2. *Let R be a primitive recursive relation of n arguments. It can be represented in **PA** by a formula* $\bar{R}(a_1, \dots, a_n)$*, namely* $\bar{f}(a_1, \dots, a_n) = \bar{0}$*, where f is the characteristic function of R. Then for any n-tuple of numbers* $(m_1, \dots, m_n)$*, if* $R(m_1, \dots, m_n)$ *is true, then* $\bar{R}(\bar{m}_1, \dots, \bar{m}_n)$ *is **PA**-provable*

Proof. Follow this note.

2. We prove that for any primitive recursive function f (of n arguments) and any

numbers $m_1, \ldots, m_n$, p, if $f(n_1, \ldots, m_n) = p$, then $\bar{f}(\bar{m}_1, \ldots, \bar{m}_n) = \bar{p}$ is **PA**-provable. The proof is by induction on the construction of f. $\square$

The converse proposition (i.e. for primitive recursive R, if $\bar{R}(\bar{m}_1, \ldots, \bar{m}_n)$ is **PA**-provable, then $R(m_1, \ldots, m_n)$ is true) follows from the consistency of **PA**

Definition 2.13 (Gödel numbering)**.** For an expression X, we use $\ulcorner X \urcorner$ to denote the corresponding number, which we call the Gödel number of X

1. First assign different odd numbers to the symbols of Ln (We include $\rightarrow$ and $-$ among the symbols of the language here)

2. Let X be a formal expression $X_0 X_1 \ldots X_n$, where each X_i, $0 \leq i \leq n$ is a symbol of L. Then $\ulcorner X \urcorner$ is defined to be $2^{\ulcorner X_1 \urcorner} 3^{\ulcorner X_1 \urcorner} \ldots p_n^{\ulcorner X_n \urcorner}$, where p_n is the nth prime number

3. If P is a proof of the form

$$\frac{Q}{S} \quad \text{or} \quad \frac{Q_1 \qquad Q_2}{S}$$

hen $\ulcorner P \urcorner$ is $2^{\ulcorner Q \urcorner} 3^{\ulcorner - \urcorner} 5^{\ulcorner S \urcorner}$ or $2^{\ulcorner Q_1 \urcorner} 3^{\ulcorner Q_2 \urcorner} 5^{\ulcorner - \urcorner} 7^{\ulcorner S \urcorner}$ respectively

If an operation or relation defined on a class of formal objects is thought of in terms of the corresponding number-theoretic operation or relation on their Gödel numbers, we say that the operation or relation has been **arithmetized**. More precisely, suppose ψ is an operation defined on n-tuples of formal objects of a certain class, and f is a number-theoretic function s.t. for all formal objects $X_1, \ldots, X_n, X$ if ψ applied to $X_1, \ldots, X_n$ produces X, then $f(\ulcorner X_1 \urcorner, \ldots, \ulcorner X_n \urcorner) = \ulcorner X \urcorner$. Then f is called the **arithmetization** of ψ

Lemma 2.14. *1. The operation of substitution can be arithmetized primitive recursively, i.e., there is a primitive recursive function sb of two arguments s.t. if $X(a_0)$ is an expression of L (where all occurences of a_0 in X are indicated), and Y is another expression, then $sb(\ulcorner X(a_0) \urcorner, \ulcorner Y \urcorner) = \ulcorner X(Y) \urcorner$ where $X(Y)$ is the result of substituting Y for a_0 and X*

2. There is a primitive recursive function ν s.t. $v(m) = \ulcorner$the mth numeral$\urcorner$. That is, $\nu(m) = \ulcorner \bar{m} \urcorner$.

3. *The notion that P is a proof (of the system $\boldsymbol{S}$) of a formula A (or a sequent S) is arithmetized primitive recursively; i.e. there is a primitive recursive relation $Prov(p, a)$ s.t. $Prov(p, a)$ is true iff there is a proof P and a formula A (or a sequent S) s.t. $p = \ulcorner P \urcorner$, $a = \ulcorner A \urcorner$ (or $a = \ulcorner S \urcorner$) and P is a proof of A (or S)*

4. *Prov may be written as $Prov_S$ to emphasize the system $\boldsymbol{S}$*

5. *the formal expression for Prov will be denoted by $\overline{Prov}$*

$\exists x \overline{\text{Prov}}(x, \overline{\ulcorner A \urcorner})$ is often abbreviated to $\overline{\text{Pr}}(\overline{\ulcorner A \urcorner})$ or $\vdash \overline{\ulcorner A \urcorner}$

Proposition 2.15. 1. *If A is $\boldsymbol{S}$-provable, then $\vdash \overline{\ulcorner A \urcorner}$ is $\boldsymbol{S}$-provable*

2. *If $A \leftrightarrow B$ is $\boldsymbol{S}$-provable, then $\overline{Pr}(\overline{\ulcorner A \urcorner}) \leftrightarrow \overline{Pr}(\overline{\ulcorner B \urcorner})$ is $\boldsymbol{S}$-provable*

3. *$\vdash \overline{\ulcorner A \urcorner} \to (\vdash \overline{\ulcorner \vdash \overline{\ulcorner A \urcorner} \urcorner})$ is $\boldsymbol{S}$-provable*

Proof. 1. Suppose A is provable with a proof P. Then by 3 of Lemma **??**, $\text{Prov}(\ulcorner P \urcorner, \ulcorner A \urcorner)$ is true, which by 2 of Proposition 2.12, that $\exists x \overline{\text{Prov}}(x, \overline{\ulcorner A \urcorner})$, i.e., $\vdash \overline{\ulcorner A \urcorner}$ is **S**-provable.

2. Suppose $A \equiv B$ is provable with a proof P and A is provable with a proof Q. There is a prescription for constructing a proof of B from P and Q, which can be arithmetized by a primitive recursive fucntion f. Thus $\text{Prov}(q, \ulcorner A \urcorner) \to \text{Prov}(f(p, q), \ulcorner B \urcorner)$ is true, from which it follows by Proposition 2.12 that $\vdash \overline{\ulcorner A \urcorner} \to \vdash \overline{\ulcorner B \urcorner}$ is provable.

3. If P is a proof of A, then we can construct a proof Q of $\vdash \overline{\ulcorner A \urcorner}$ by 1. This process is uniform in P; in other words, there is a uniform prescription for obtaining Q from P. Thus

$$\text{Prov}(p, \ulcorner A \urcorner) \to \text{Prov}(f(p), \ulcorner \overline{\text{Pr}}(\overline{\ulcorner A \urcorner}) \urcorner)$$

is true for some primitive recursive function f, from which it follows that $\vdash \overline{\ulcorner A \urcorner} \to \vdash \overline{\ulcorner \vdash \overline{\ulcorner A \urcorner} \urcorner}$

□

Definition 2.16. A formula of L (the language of **S**) with one free variable, say $T(a_0)$, is called a **truth definition** for **S** if for every sentence of A of L

$$T(\overline{\ulcorner A \urcorner}) \equiv A$$

is **S**-provable

Theorem 2.17 (Tarski)**.** *If **S** is consistent, then it has no truth definition*

Proof. Suppose otherwise. Consider the formula $F(a_0)$, with sole free variable a_0, defined as: $\neg T(\overline{\text{sb}}(a_0, \bar{\nu}(a_0)))$. Put $p = \ulcorner F(a_0) \urcorner$, and let A_T be the sentecne $F(\bar{p})$. Then by definition

$$A_T \equiv \neg T(\overline{\text{sb}}(\bar{p}, \bar{\nu}(\bar{p})))$$

Also since $\ulcorner A_T \urcorner = \text{sb}(p, \nu(p))$ by definition, (note that $\ulcorner F(\bar{p}) \urcorner = \text{sb}(p, \nu(p)))$ we can prove in **S** the equivalences

$$\begin{aligned} A_T &\equiv T(\overline{\ulcorner A_T \urcorner}) \quad \text{(by assumed property of } T\text{)} \\ &\equiv T(\overline{\text{sb}}(\bar{p}, \bar{\nu}(\bar{p}))) \end{aligned}$$

□

In the proof of Theorem 2.17 we need *not* asume that **S** is axiomatizable. So we may take as the axioms of **S** the set of all sentences of Ln which are *true* in the intended interpretation $\mathfrak{M}$ of **PA**. We then obtain that there is no formula $T(a_0)$ of Ln s.t. for any sentence $A$of Ln

$$A \text{ is true } \Leftrightarrow T(\overline{\ulcorner A \urcorner}) \text{ is true}$$

The corollary of Theorem 2.17 can be stated in the form: "The notion of arithmetical truth is not arithmetical"

Definition 2.18. **S** is called **incomplete** if for some sentence A, neither A nor $\neg A$ is provable in **S**

Definition 2.19. Consider a formula $F(\alpha)$ with a metavariable α (i.e. a new predicate variable, not in L, which we only use temporarily for notational convenience), where α is regared as an atomic formula in $F(\alpha)$ and $F(\alpha)$ is closed. $F(\vdash \overline{\text{sb}}(a_0, \bar{\nu}(a_0)))$ is a formula with a_0 as its sole free variable. Define $p = \ulcorner F(\vdash \overline{\text{sb}}(a_0, \bar{\nu}(a_0))) \urcorner$ and A_F as $F(\vdash \overline{\text{sb}}(\bar{p}, \bar{\nu}(\bar{p})))$. Note that A_F is a sentence of L

Lemma 2.20. $A_F \equiv F(\vdash \overline{\ulcorner A_F \urcorner})$ *is provable in **S***

Proof. Since $\ulcorner A_F \urcorner = \text{sb}(p, \nu(p))$ by definition

$$\overline{\ulcorner A_F \urcorner} = \overline{\text{sb}}(\bar{p}, \bar{\nu}(\bar{p})) \text{ is provable in } \mathbf{S}$$

Hence $A_F \equiv F(\vdash \overline{\ulcorner A_F \urcorner})$ is provable in **S** □

From now no we shall use the abbreviation $\vdash A$ for $\vdash \overline{\ulcorner A \urcorner}$

Definition 2.21. **S** is called ω**-consistent** if the following condition is satisfied. For every formula $A(a_0)$, if $\neg A(\bar{n})$ is provable in **S** for every $n \in \mathbb{N}$, then $\exists x A(x)$ is not provable in **S**. Note that ω-consistency of **S** implies consistency of **S**

Theorem 2.22 (Gödel's first incompleteness theorem)**.** *If **S** is ω-consistent, then **S** is incomplete*

Proof. There exists a sentence A_G of L s.t. $A_G \equiv \neg \vdash A_G$ is provable in **S**. (Any such sentence will be called a Gödel sentence for **S**.) This follows from Lemma 2.20 by taking $F(\alpha)$ to be $\neg\alpha$.

First we shall show that A_G is not provable in **S**, assuming only the consistency of **S** (but without assuming the ω-consistency of **S**). Suppose that A_G where provable in **S**. Then by 1 of Proposition 2.15, $\vdash A_G$ is provable in **S**; thus by the definition of Gödel sentence, $\neg A_G$ is provable in **S**

Next we shall show that $\neg A_G$ is not provable in **S**, assuming the ω-consistency of **S**. Since we have proved that A_G is not provable in **S**, for each $n = 0, 1, 2, ...,$ $\neg\overline{\mathrm{Prov}}(\bar{n}, \overline{\ulcorner A_G \urcorner})$ is provable in **S**. By the ω-consistency of **S**, $\exists x \overline{\mathrm{Prov}}(x, \overline{\ulcorner A_G \urcorner})$ is not provable in **S**. Since $\neg A_G \equiv \vdash A_G$ is provable in **S**, $\neg A_G$ is not provable in **S** ☐

In fact ,A_G, although unprovable, is (intuitively) true, since it asserts its own unprovability

Definition 2.23. $\overline{\mathrm{Consis}_\mathbf{S}}$ is the sentence $\neg \vdash 0 = 1$ (So $\overline{\mathrm{Consis}_\mathbf{S}}$ asserts the consistency of **S**)

Theorem 2.24 (Gödel's second incompleteness theorem)**.** *If **S** is consistent, then $\overline{\mathrm{Consis}_\mathbf{S}}$ is not provable in **S***

Proof. Let A_G be a Gödel sentence. In the proof of Theorem 2.22, we proved that A_G is not provable, assuming only consistency of **S**. Now we shall prove a stronger theorem: that $A_G \equiv \overline{\mathrm{Consis}_\mathbf{S}}$ is provable in **S**

1. To show $A_G \to \overline{\mathrm{Consis}_\mathbf{S}}$ is provable in **S**. By Lemma 2.11, $\neg\overline{\mathrm{Consis}_\mathbf{S}} \equiv \forall \ulcorner A \urcorner (\vdash A)$ is provable. Therefore

$$A_G \to \neg \vdash A_G \to \neg \forall \ulcorner A \urcorner (\vdash A) \to \overline{\mathrm{Consis}_\mathbf{S}}$$

2. To show $\overline{\mathrm{Consis}_\mathbf{S}} \to A_G$ is provable in **S**. Again by Lemma 2.11, $\overline{\mathrm{Consis}_\mathbf{S}}, \vdash A_G \to \neg \vdash \neg A_G \to \neg \vdash\vdash A_G$ since $\neg A_G \equiv \vdash A_G$. But $\vdash A_G \to \vdash\vdash A_G$, by Proposition 2.15. So $\overline{\mathrm{Consis}_\mathbf{S}}, \vdash A_G \to \neg \vdash\vdash A_G \wedge \vdash\vdash A_G$ and so $\overline{\mathrm{Consis}_\mathbf{S}} \to \neg \vdash A_G \to A_G$

$\square$

2.3 A Discussion of Ordinals from a Finitist Standpoint

O1 0 is an ordinal

O2 Let μ and $\mu_1, \dots, \mu_n$ be ordinals. Then $\mu_1 + \cdots + \mu_n$ and ω^μ are ordinals

O3 Only those objects obtained by O1 and O2 are ordinals

ω^0 will be dnoted by 1.

1. $<$ is a linear ordering and 0 is its least element

2. $\omega^\mu < \omega^\nu$ iff $\mu < \nu$

3. Let μ be an ordinal containing an occurrence of the symbol 0 but not 0 itself, and let μ' be the ordinal obtained from μ by eliminating this occurrence of 0 as well as excessive occurrence of $+$. Then $\mu = \mu'$

As a consequence of 3 it can be easily shown that

4. Every ordinal which is not 0 can be expressed in the form

$$\omega^{\mu_1} + \omega^{\mu_2} + \cdots + \omega^{\mu_n}$$

   where each of $\mu_1, \dots, \mu_n$ which is not 0 has the same property. (Each term ω^{μ_i} is called a monomial of this ordinal)

5. Let μ and ν be of the forms

$$\omega^{\mu_1} + \cdots + \omega^{\mu_k} \quad \text{and} \quad \omega^{\nu_1} + \cdots + \omega^{\nu_l}$$

   respectively. Then $\mu + \nu$ is defined as

$$\omega^{\mu_1} + \cdots + \omega^{\mu_k} + \omega^{\nu_1} + \cdots + \omega^{\nu_l}$$

6. Let μ be an ordinal which is written in the form of 4 and contains two consecutive terms ω^{μ_j} and $\omega^{\mu_{j+1}}$ with $\mu_j < \mu_{j+1}$, i.e., μ is of the form

$$\cdots + \omega^{\mu_j} + \omega^{\mu_{j+1}} + \dots$$

   and let μ' be an ordinal obtained from μ by deleting "$\omega^{\mu_j}+$", so that μ' is of the form

$$\cdots + \omega^{\mu_{j+1}} + \dots$$

   Then $\mu = \mu'$

As a consequence of 6 we can show that

7. For every ordinal μ (which is not 0) there is an ordinal of the form

$$\omega^{\mu_1} + \cdots + \omega^{\mu_n}$$

where $\mu_1 \geq ... \geq \mu_n$ s.t. $\mu = \omega^{\mu_1} + \cdots + \omega^{\mu_n}$. This is called the normal form of μ

8. Let μ have the normal form

$$\omega^{\mu_1} + \cdots + \omega^{\mu_n}$$

and ν be > 0. Then $\mu \cdot \omega^\nu = \omega^{\mu_1 + \nu}$

9. Let μ and ν be as in 5. Then

$$\mu \cdot \nu = \nu \cdot \omega^{\nu_1} + \cdots + \mu \cdot \omega^{\nu_l}$$

10. $(\omega^\mu)^n$ is defiend as $\omega^\mu \cdots \omega^\mu$ (n times) for any natural number n. Then $(\omega^\mu)^n = \omega^{\mu \cdot n}$

(*) Whenever a concrete method of constructing decreasing sequences of ordinals is given, any such decreasing sequence must be finite

Suppose $a_0 > a_1 > ...$ is a decreasing sequence concretely given

1. Assume $a_0 < \omega$, or a_0 is a natural number

2. Suppose each a_i in $a_0 > a_1 > ...$ is written in the canonical form; a_i has the form

$$\omega^{\mu_1^i} + \omega^{\mu_2^i} + \ldots \cdots + \omega^{\mu_{n_i}^i} + k_i$$

where $\mu_j^i > 0$ and k_i is a natural number. A sequence in which k_i does not appear for any a_i will be called 1-sequence. We call $\omega^{\mu_1^i} + \cdots + \omega^{\mu_{n_i}^i}$ in a_i the 1-major part of a_i. We shall give a concrete method (M_1) which enables us to do the following: given a decreasing sequence $a_0 > a_1 > ...$, where each a_i is written in its canonical form, the method M_1 concretely produces a (decreasing) 1-sequence $b_0 > b_1 > ...$ so as to satisfy the condition

(C_1) b_0 is the 1-major part of a_0, and we can concretely show that if $b_0 > b_1 > ...$ is a finite sequence, then so is $a_0 > a_1 > ...$

This method M_1 (a 1-eliminator) is defined as follows. Put $a_i = a_i' + k_i$, where a_i' is the 1-major part of a_i. Then $a_0 > a_1 > \dots$ can be expressed by $a_0' + k_0 > a_1' + k_1 > \dots$

Put $b_0 = a_0'$. Suppose $b_0 > b_1 > \cdots > b_m$ has been constructed in such a manner that b_m is a_j' for some j. Then either $a_j' = a_{j+1}' = \cdots = a_{j+p}'$ for some p and a_{j+p} is the last term in the sequence, or $a_j' = a_{j=1}' = \cdots = a_{j+p}' > a_{j+p+1}'$. This is so, since $a_j' = a_{j+1}' = \cdots = a_{j+p}' = \dots$ implies $k_j > k_{j+1} > \cdots > k_{j+p} > \dots$, but since a sequence (of natural numbers) must stop. Therefore either the whole sequence stops or $a_{j+p}' > a_{j+p+1}'$ for some p. If the former is the case, then stop. If the latter holds then put $b_{m+1} = a_{j+p+1}'$

From the definition, it is obvious that $b_0 > b_1 > \cdots > b_m > \dots$. Suppose this sequence is finite, say $b_0 > b_1 > \cdots > b_m$. Then according to the prescribed construction of b_{m+1} the original sequence is finite. Then (C_1) is satisfied.

3. Suppose we are given a decreasing sequence $a_0 > a_1 > \dots$ in which $a_0 < \omega^2$. Then by a 1-eliminator M_1 applied to this sequence we can construct a 1-sequence $b_0 > b_1 > \dots$ where $b_0 \leq a_0$. Then $b_0 > b_1 > \dots$ can be written in the form $\omega \cdot k_0 > \omega \cdot k_1 > \dots$, which implies $k_0 > k_1 > \dots$. Then by 1, $k_0 > k_1 > \dots$ must be finite, which implies $b_0 > b_1 > \dots$ and $a_0 > a_1 > \dots$ are finite.

4. We now define "n-sequences" as follows. Let $a_0 > a_1 > \dots$ be a descending sequence which is written in the form $a_0' + c_0 > a_1' + c_1 > \dots$ where if $a_i = a_i' + c_i$, then each monomial in a_i' is $\geq \omega^n$ and each monomial in c_i is $< \omega^n$ (a_i' is called the n-major part of a_i) Such a sequence is called an n-sequence if every c_i is empty

   Now assume (as an inductive hypothesis) that any descending sequence $d_0 > d_1 > \dots$, with $d_0 < \omega^n$ is finite. We shall define a concrete method M_n (an n-eliminator) s.t., given a decreasing sequence $a_0 > a_1 > \dots$, M_n concretely produces an n-sequence, say $b_0 > b_1 > \dots$, which satisfies

   (C_n) b_0 is the n-major part of a_0, and if $b_0 > b_1 > \dots$ is finite then we can concretely show that $a_0 > a_1 > \dots$ is finite

   The prescription for M_n is as follows. Write each a_i as $a_i' + c_i$ where a_i' is the n-major part of a_i. Put $b_0 = a_0'$. Suppose $b_0 > b_1 > \cdots > b_m$

has been constructed and b_m is a'_j. If $a'_j = a'_{j+1} = \cdots = a'_{j+p}$ and a'_{j+p} is the last term in the given sequence, then stop. Otherwise $a'_j = a'_{j+1} = \cdots = a'_{j+p} > a'_{j+p+1}$ for some p, since $a'_j = a'_{j+1} = \cdots = a'_{j+p}$ implies that $c_j > c_{j+1} > \cdots > c_{j+p}$, which, by the induction hypothesis, is finite; hence for some p, $c_{j+p+1} \geq c_{j+p}$, which implies $a'_{j+p} > a'_{j+p+1}$. Then define $b_m = a'_{j+p+1}$. Then the sequence $b_0 > b_1 > ...$ satisfies (C_n).

5. By means of the n-eliminator M_n we shall prove that a decreasing sequence $a_0 > a_1 > ...$, where $a_0 < \omega^{n+1}$ must be finite. By applying M_n to $a_0 > a_1 > ...$ we can construct concretely an n-sequence, say $b_0 > b_1 > ...$ where $b_0 \leq a_0$. Moreover b_i can be written as $\omega^n \cdot k_i$, where k_i is a natural number. So $\omega^n \cdot k_0 > \omega^n \cdot k_1 > ...$ and this implies $k_0 > k_1 > ...$, which is a finite sequence by 1, hence $b_0 > b_1 > ...$ is finite, which in turn implies that $a_0 > a_1 > ...$ with $a_0 < \omega^n$ is finite

6. From 3 and 5 we conclude: given (concretely) any natural number n, we can concretely demonstrate that any decreasing sequence $a_0 > a_1 > ...$ with $a_0 < \omega^n$

7. Any decreasing sequence $a_0 > a_1 > ...$ is finite if $a_0 < \omega^\omega$

8. Now the general theory of α-sequences and (α, n)-eliminators will be developed, where α ranges over all ordinals $< \epsilon$ and n ranges over natural numbers > 0. A descending sequence $d_0 > d_1 > ...$ is called an **α-sequence** if in each d_i all the monomials are $\geq \omega^\alpha$. If $a = a' + c$ where each monomial in a' is $\geq \omega^\alpha$ and each monomial in c is $< \omega^\alpha$, then we say that a' is the α-major part of a. An **α-eliminator** has the property that given any concrete descending sequence, say $a_0 > a_1 > ...$, it concretely produces an α-sequence $b_0 > b_1 > ...$ s.t.

   (a) b_0 is the α-major part of a_0

   (b) if $b_0 > b_1 > ...$ is a finite sequence then we can concretely demonstrate that $a_0 > a_1 > ...$ is finite

   Assuming that an α-eliminator has been defined for every α, we can show that any decreasing sequence is finite. For consider $a_0 > a_1 > ....$ There exists an α s.t. $a_0 < \omega^{\alpha+1}$. An α-eliminator concretely gives an α-sequence $b_0 > b_1 > ...$ satisfying 1 and 2 above. Since $b_0 \leq a_0$, each b_i can be written in the form $\omega^\alpha \cdot k_i$; thus $a_0 > a_1 > ...$ is finite. This proves our objective (*). Therefore what must be done is to define α-eliminators for all $\alpha < \epsilon$

9. We rename an α-eliminator to be an $(\alpha, 1)$-eliminator. Suppose that (α, n)-eliminators have benn defined. A $(\beta, n+1)$-eliminator is a **concrete** method for constructing an $(\alpha \cdot \omega^{\beta}, n)$-eliminator from any given (α, n)-eliminator.

10. Suppose $\{\mu_m\}_{m<\omega}$ is an increasing sequence of ordinals whose limit is μ (where there is a concrete method for obtaining μ_m for each m), and suppose g_m is an μ_m-eliminator. Then the g defined as follows is a μ-eliminator.

    Suppose $a_0 > a_1 > \ldots$ is a concretely given sequence. If a_0 is written as $a_0' + c_0$, where a_0' is the μ-major part of a_0 then there exists an m for which $c_0 < \omega^{\mu_m}$, so we may assume that each a_i is written as $a_i' + c_i$, where a_i' is the μ_m-major part of a_i. Then g_m can be applied to the sequence $a_0 > a_1 > \ldots$ and hence it concretely produces a μ_m-sequence

    $$b_{10} > b_{11} > b_{12} > \ldots \tag{2}$$

    satisfying (a) and (b) above (with μ_m in place of α), with $b_{10} = a_0'$ so that in fact B_{10} is the μ-major part of a_0. Write $b_0 = b_{10}$.

    Now consider the sequence $b_{11} > b_{12} > \ldots$. Suppose $b_{11} \geq \omega^{\mu}$. Then repeat the above procedure: i.e., for the sequence (2), write $b_{10} = b_{10}' + c_{10}$, where b_{10}' is the μ-major part of b_{10}. Then there exists an m_1 s.t. $c_{10} < \omega^{m_1}$. So apply g_{m_1} to the sequence $b_{11} > b_{12} > b_{13} > \ldots$ to obtain a μ_{m_1}-sequence

    $$b_{21} > b_{22} > b_{23} > \ldots$$

    satisfying (a) and (b), with b_{21} the μ-major part of b_{10}. Put $b_1 = b_{21}$. Suppose $b_{22} \geq \omega^{\mu}$. Then repeat this procedure with the sequence $b_{22} > b_{23} > \ldots$ to obtain a sequence

    $$b_{32} > b_{33} > b_{34} > \ldots$$

    and put $b_2 = b_{32}$. Continuing in this way, we obtain a μ-sequence

    $$b_0 > b_1 > b_2 > \ldots$$

    If this sequence is finite with last term $b_l = b_{l+1,l}$, then it follows that in this sequence

    $$b_{l+1,l} > b_{l+1,l+1} > b_{l+1,l+2} > \ldots$$

We must have $b_{l+1,l+1} < \omega^{\mu}$. So $b_{l+1,l+1} < \omega^{\mu_{m'}}$ for some m'. Applying $g_{m'}$ to this sequence (10), we obtain a finite $\mu_{m'}$-sequence with only the term 0; hence the sequence $b_{l,l-1} > b_{l,l} > \ldots$ is finite; then $b_{l-1,l-2} > b_{l-1,l-1} > \ldots$ is finite; hence $a_0 > a_1 > \ldots$ is finite

11. Suppose $\{\mu_m\}_{m<\omega}$ is a sequence of ordinals whose limit is μ and suppose for each m, a $(\mu_m, n+1)$-eliminator is concretely given. Then we can define a $(\mu, n+1)$-eliminator g as follows. The difinition is by induction on n.

    For $n = 0$, 10 applies.

    Assume 11 for n; so there is an operation k_n s.t. for any sequence $\{\gamma_m\}_{m<\omega}$ with limit γ and (γ_m, n)-eliminator g'_m, k_n applied to g'_m concretely produces a (γ, n)-eliminator. Now for $n+1$, suppose a sequence $\{\beta_m\}_{m<\omega}$ with limit β and an (α, n)-eliminator p are given. Since g_m is a $(\beta_m, n+1)$-eliminator, it produces concretely an $(\alpha \cdot \omega^{\beta_m}, n)$-eliminator from p, which we denote by $g_m(p)$. So by taking $\alpha \cdot \omega^{\beta_m}$ for γ_m, $g_m(p)$ for g'_m and $\alpha \cdot \omega^{\beta}$ for γ, we can apply the inductive hypothesis; thus k_n applied to $\{g_m\}'$ defines an $(\alpha \cdot \omega^{\beta}, n)$-eliminator q. This procedure for defining q from p is concrete, and so serves as a $(\beta, n+1)$-eliminator.

12. Suppose g is a $(\mu, n+1)$-eliminator. Then we will construct a $(\mu \cdot \omega, n+1)$-eliminator. In virtue of 11 it suffices to show that we can concretely construct (from g) a $(\mu \cdot m, n+1)$-eliminator for every $m < \omega$. Suppose an (α, n)-eliminator, say f, is given. Note that

$$\alpha \cdot \omega^{\mu \cdot m} = \alpha \cdot \underbrace{\omega^{\mu} \cdot \omega^{\mu} \ldots \omega^{\mu}}_{m}$$

    Since g is a $(\mu, n+1)$-eliminator, g concretely constructs an $(\alpha \cdot \omega^{\mu}, n)$-eliminator from f, which we denote by $g(f)$. Now apply g to this, to obtain an $(\alpha \cdot \omega^{\mu} \cdot \omega^{\mu}, n)$-eliminator $g(g(f))$. Repeating this procedure m times, we obtain the $(\alpha \cdot \omega^{\mu \cdot m}), n$-eliminator $g(g(\cdots g(f)\cdots))$

13. We can construct a $(1, m+1)$-eliminator for every $m \geq 0$. The induction is by induction on m. We may take M_1 as a $(1,1)$-eliminator.

    For $m = 1$, the construction of a $(1,2)$-eliminator is reduced to the construction of an $(\alpha + \alpha)$-eliminator from an α-eliminator. Given $a_0 > a_1 > \ldots$, apply an α-eliminator to obtain $b_0 > b_1 > \ldots$, where $\{b_i\}$ is an α-sequence, b_0 is the α-major part of a_0, and if $\{b_i\}$ is finite, then so is $\{a_i\}$. Each b_i can be written in the form $\omega^{\alpha} \cdot c_i$, where $\{c_i\}$ is decreasing and if $\{c_i\}$ is finite ,then so it $\{b_i\}$. $a_0 = b_0 + e_0$ where $e_0 < \omega^\alpha$. Apply α-eliminator to $\{c_i\}$ to obtain $d_0 > d_1 > \ldots$, where $\{d_i\}$ is an α-sequence, d_0 is the α-major part of c_0 and if $\{d_i\}$ is finite, then so is $\{c_i\}$.

$\{\omega^\alpha \cdot d_i\}$ is an $(\alpha+\alpha)$-sequence and decresing. If $\{\omega^\alpha d_i\}$ is finite, then so are $\{d_i\}$,$\{c_i\}$, $\{b_i\}$, $\{a_i\}$ successively, and

$$\begin{aligned}
\omega^\alpha \cdot d_0 &= \omega^\alpha \cdot (\text{the } \alpha\text{-major part of } c_0) \\
&= (\alpha+\alpha)\text{-major part of } b_0 \\
&= (\alpha+\alpha)\text{-major part of } a_0
\end{aligned}$$

So $\{\omega^\alpha d_i\}$ is the $(\alpha+\alpha)$-sequence which was desired for $\{a_i\}$

For $m > 1$, suppose f is an (α, m)-eliminator. Then by 12 we can construct an $(\alpha \cdot \omega, m)$-eliminator concretely from f. Hence we have a $(1, m+1)$-eliminator

14. Conclusion: An (α, n)-eliminator can be constructed for every α of the form ω_m, i.e.,

$$\left.\omega^{\cdot^{\cdot^{\omega}}}\right\} m$$

The construction is by induction on m. If $m = 0$ then we define α to be $1 = \omega^0$. Then an (α, n)-eliminator has been defined in 13 for every n. Suppose f is $(1, n)$-eliminator, and g is an $(\alpha, n+1)$-eliminator, which we assume to have been defined. Then g operates on f and produces the required $(1 \cdot \omega^\alpha, n) = (\omega^\alpha, n)$-eliminator.

An ordinal μ is **accessible** if it has been demonstrated that every strictly decreasing sequence starting with μ is finite. More precisely, we consider the notion of accessibility only when we have actually seen, or demonstrated constructively, that a given ordinal is accessible.

First, we assume we have arithmetized the construction of the ordinals (less than ϵ_0) given by clauses O1-O3. In other words, we assume a Gödel numbering of these ordinals, with certain nice properties: namely, the induced number-theoretic relations and functions corresponding to the ordinal relations and functions $=$,$<$,$+$,$\cdot$ and exponentiation by ω are primitive recursive. also we can primitive recursively represent any (Gödel number of an) ordinal in its normal form, and hence decide primitive recursively whether it represents a limit or successor ordinal, etc. The ordering of the natural numbers corresponding to $<$ will be called a "standard well-ordering of type ϵ_0".

Our method for proving the accessibility of ordinals will be as follows

1. when it is known that $\mu_1 < \mu_2 < \ldots \to \nu$ (i.e., ν is the limit of the increasing sequence $\{\mu_i\}$) and that every μ_i is accessible, then ν is also accessible
2. A method is given by which, from the accessibility of a subsystem, one can deduce the accessibility of a larger system
3. by repeating 1 and 2, we show that every initial segment of our ordering is accessible, and hence so is the whole ordering

Consider the decreasing sequences of ordinals less than $\omega+\omega$. Here we can again see that every decreasing sequence terminates. Consider the first term μ_0 of such a sequence. We can effectively decide whether it is of the form n or of the form $\omega+n$, where n is a natural number. If it is of the form $\omega+n$, consider the first $n+2$ terms of the sequence

$$\mu_{n+1} < \cdots < \mu_2 < \mu_1 < \mu_0$$

It is easily seen that μ_{n+1} cannot be of the form $\omega+m$ for any natural number m and hence must be a natural number. This method can be extended to the cases of decreasing sequences of ordinals less than $\omega \cdot n$, less than ω^2, less than ω^ω, etc

Lemma 2.25. *If μ and ν are accessible, then so is $\mu+\nu$*

Proof. given ordinals μ,ξ,ν s.t. $\mu \leq \xi < \nu$, we can effectively find a ν_0 s.t. $\nu_0 < \nu$ and $\xi = \mu + \nu_0$ □

Lemma 2.26. *if μ is accessible, then so is $\mu \cdot \omega$*

Proof. If $\nu < \mu \cdot \omega$, then we can find an n s.t. $\nu < \mu \cdot n$ □

Definition 2.27. μ is said to be **1-accessible** if μ is accessible, μ is said to be $(n+1)$**-accessible** if for every ν which is n-accessible, $\nu \cdot \omega^\mu$ is n-accessible

Lemma 2.28. *If μ is n-accessible and $\nu < \mu$, then ν is n-accessible*

Lemma 2.29. *Suppose $\{\mu_m\}$ is an increasing sequence of ordinals with limit μ. If each μ_m is n-accessible, then so is μ*

Lemma 2.30. *If ν is $(n+1)$-accessible, then so is $\nu \cdot \omega$*

Proof. We must show that for any n-accessible μ, $\mu \cdot^{\nu \cdot \omega}$ is n-accessible. For this purpose it suffices to show that $\mu \cdot \omega^{\nu \cdot m}$ is n-accessible for each m (cf. Lemma 2.29). This is obvious since

$$\mu \cdot \omega^{\nu \cdot m} = \mu \cdot \omega^{\nu} \ldots \omega^{\nu}$$

and ν is $(n+1)$-accessible □

Proposition 2.31. *1 is $(n+1)$-accessible*

Proof. Suppose μ is n-accessible. Then by Lemma 2.30 $\mu \cdot \omega = \mu \cdot \omega^1$ is n-accessible □

Definition 2.32. $\omega_0 = 1$; $\omega_{n+1} = \omega^{\omega_n}$

Proposition 2.33. *ω_k is $(n-k)$-accessible for an arbitrary $n > k$*

Proof. By induction on k. If $k = 0$, then $\omega_k = 1$ and hence is n-accessible for all n. Suppose ω_k is $(n-k)$-accessible. Since 1 is $[n-(k+1)]$-accessible, $1 \cdot \omega^{\omega_k}$ is $[n-(k+1)]$-accessible, i.e., ω_{k+1} is $[n-(k+1)]$-accessible □

Proposition 2.34. *ω_k is accessible for every k*

Given any decreasing sequence of ordinals (less than ϵ_0) there is an ω_k s.t. all ordinals in the sequence are less than ω_k. Therefore the sequence must be finite by Proposition 2.34. Thus we can conclude:

Proposition 2.35. *ϵ_0 is accessible*

A Gentzen-style consistency proof is carried out as follows

1. Construct a suitable standard ordering, in the strictly finitist standpoint
2. Convince oneself, in the Hilbert-Gentzen standpoint, that it is indeed a well-ordering
3. othewise use only strictly finitist means in the consistency proof

2.4 A Consistency Proof of PA

We assume from now on that **PA** is formalized in a language which includes a constant f for every primitive recursive function f. We call this language L.

As initial sequents of **PA** we will also take from now on the defining equations for all primitive recursive functions, as well as all sequents $\rightarrow s = t$ where s, t are closed terms of L denoting the same number, and all sequents $s = t \rightarrow$ where s, t are closed terms of L denoting different numbers.

Let R be a property of proofs s.t.

(*) For any proof P satisfying R, we can find (effectively from P) a proof P' satisfying R s.t. P' has a smaller ordinal than P

We can then infer from (*) and the accessibility of ϵ_0:

(**) No proof satisfies R

The procedure of finding P' from P in () **is called a *reduction of P to** P' (for the property R)

The property R of proofs that we will be interested in is the property of having $\rightarrow$ as an end-sequent

By giving a uniform reduction procedure for this property (Lemma **??**) we will have shown that no proof of **PA** ends with $\rightarrow$, in other words

Theorem 2.36. *The system **PA** is consistent*

Definition 2.37. A proof in **PA** is **simple** if no free variables occur in it, and it contains only mathematical initial sequents, weak inferences and inessential cuts

Lemma 2.38. *There is no simple proof of* $\rightarrow$

Proof. Let P be any simple proof. All the formulas in P are of the form $s = t$ with s and t closed.

A sequent in P is then given the value T if at least one formula in the antecedent is false or at least one formula in the succedent is true, and it is given the value F. It is easy to see that all mathematical initial sequents take the value T, and weak inferences and inessential cuts preserve the value T downward for sequents. So all sequents of P have the value T. But $\rightarrow$ has the value F. □

Definition 2.39. 1. The **grade of a formula**, is the number of logical symbols it contains. The **grade of a cut** is the grade of the cut formula; the **grade of an ind inferene** is the grade of the induction formula

2. The **height of a sequent** S in a proof P (denoted by $h(S;P)$ or for short $h(S)$) is the maximum of the grades of the cuts and ind's which occur in P below S

Proposition 2.40. *1. The height of the end-sequent of a proof is 0*

2. If S_1 is above S_2 in a proof, then $h(S_1) \geq h(S_2)$; if S_1 and S_2 are the upper sequent of an inference, then $h(S_1) = h(S_2)$

For any ordinal α and natural number n, $\omega_n(\alpha)$ is defined by induction on n; $\omega_0(\alpha) = \alpha$, $\omega_{n+1} = \omega^{\omega_n(\alpha)}$. So

$$\omega_n(\alpha) = \underbrace{\omega^{\cdot^{\cdot^{\omega}}}}_{n}{}^{\alpha}$$

Definition 2.41. Assignment of ordinals (less than ϵ_0) to the proofs of **PA**. First we assign ordinals to the sequents in a proof. The ordinal assigned to a sequent S in a proof P is denoted by $o(S;P)$ or $o(S)$. Now suppose a proof P is given. We shall define $o(S) = o(S;P)$ for all sequents

We assume that the ordinals are expressed in normal form. If μ and ν are ordinals of the form $\omega^{\mu_1} + \cdots + \omega^{\mu_m}$ and $\omega^{\nu_1} + \cdots + \omega^{\nu_n}$ respectively, then $\mu\#\nu$ denotes the ordinal $\omega^{\lambda_1} + \cdots + \omega^{\lambda_{m+n}}$ where $\{\lambda_1, \ldots, \lambda_{m+n}\} = \{\mu_1, \ldots, \mu_m, \nu_1, \ldots, \nu_n\}$ and $\lambda_1 \geq \ldots \geq \lambda_{m+n}$. $\mu\#\nu$ is called the **natural sum** of μ and ν

1. An initial sequent is assigned the ordinal 1
2. If S is the lower sequent of a weak inference, then $o(S)$ is the same as the ordinal of its upper sequent.
3. If S is the lower sequent of $\wedge$left,$\vee$right,$\supset$right,$\neg$right,$\neg$left or an inference involving a quantifier, and tnhe upper sequent has the ordinal μ, then $o(S) = \mu + 1$
4. If S is the lower sequent of $\wedge$right,$\vee$left or $\supset$left and the upper sequents have ordinals μ and ν, then $o(S) = \mu\#\nu$
5. If S is the lower sequent of a cut and its upper sequents have the ordinals μ and ν, then $o(S) = \omega_{k-l}(\mu\#\nu)$, i.e.

$$\left.\omega^{\cdot^{\cdot^{\omega^{\mu\#\nu}}}}\right\} k-l$$

   where k and l are the heights of the upper sequents and of S, respectively

6. If S is the lower sequent of an ind and its upper sequent has the ordinal μ, then $o(S)$ is $\omega_{k-l+1}(\mu_1+1)$

$$\left.\omega^{\cdot^{\cdot^{\cdot^{\omega^{\mu_1+1}}}}}\right\}(k-l)+1$$

where μ has the normal form $\omega^{\mu_1}+\cdots+\omega^{\mu_n}$ (so that $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n$) and k and l are the heights of the upper sequent and of S respectively

7. The ordinal of a proof P, $o(P)$ is the ordinal of its end-sequent. We use the notation

$$P: \quad \begin{array}{c}\vdots \\ \Gamma \xrightarrow{\mu} \Delta\end{array}$$

to denote a proof P of $\Gamma \rightarrow \Delta$ s.t. $o(\Gamma \rightarrow \Delta; P) = o(P) = \mu$

Lemma 2.42. *Suppose P is a proof containing a sequent S_1 , there is no ind below S_1, P_1 is the subproof of P ending with S_1, P_1' is any other proof of S_1 and P' is the proof formed from P by replacing P_1 by P_1'*

$$P: \begin{array}{c}\vdots\, P_1 \\ S_1 \\ \vdots\end{array} \qquad P': \begin{array}{c}\vdots\, P_1' \\ S_1 \\ \vdots\end{array}$$

Suppose also that $o(S_1; P') < o(S_1; P)$. Then $o(P') < o(P)$

Proof. Consider a thread of P passing through S_1. We show that for any sequent S of this thread at or below S_1: if S' is the sequent "corresponding to" S in P', then

$$o(S'; P') < o(S; P)$$

This is true for $S = S_1$ by assumption, and this property is preserves downwards by all the inference rules (We use the fact that the natural sum is strictly monotonic in each argument, i.e., $\alpha < \beta \Rightarrow \alpha\#\gamma < \beta\#\gamma$). Finally let S be the end-sequent □

Now let R be the property of proofs of ending with the sequent $\rightarrow$; i.e., for any proof P, $R(P)$ holds iff P is a proof of $\rightarrow$

Notice first that if P is a proof of $\rightarrow$, then every logical inference of P is implicit (cf. Definition 2.7). Hence the definition of end-piece for such proofs can be simply stated as follows.

The end-piece of a proof of $\rightarrow$ consists of all those sequents that are encountered as we acsend each thread from the end-sequent and stop as soon as we arrive at a logical inference. This inference belongs to the boundary

Lemma 2.43. *If P is a proof of $\rightarrow$, then there is another proof P' of $\rightarrow$ s.t. $o(P') < o(P)$*

Proof. Let P be a proof of $\rightarrow$. We can assume, by Proposition 2.8 that P is regular. We describe a "reduction" of P to obtain the desired P'.

At each step, the ordinal of the resulting proof does not increase, and at least at one step, the ordinal decreases

Step 1. Suppose the end-piece of P contains a free variable, say a, which is not used as an eigenvariable. Then replace a by constant 0. This results in a proof of $\rightarrow$, with the same ordinal.

Step 1 is performed repeated until there is no free variable in the end-piece which is not used as an eigenvariable

Step 2. Suppose the end-sequence of P contains an ind. Then take a lowermost one, say I. Suppose I is of the following form

$$\dfrac{\begin{matrix} \vdots\, P_0(a) \\ F(a), \Gamma \overset{\mu}{\rightarrow} \Delta, F(a') \end{matrix}}{\begin{matrix} F(0), \Gamma \rightarrow \Delta, F(s) \\ \vdots \\ \rightarrow \end{matrix}}\ I$$

where P_0 is the subproof ending with $F(a), \Gamma \rightarrow \Delta, F(a')$ and let l and k be the heights of the upper sequent (call it S) and the lower sequent (calll it S_0) of I, respectively. Then

$$o(S_0) = \omega_{l-k+1}(\mu_1 + 1)$$

where $\mu = o(S) = \omega^{\mu_1} + \cdots + \omega^{\mu_n}$ and $\mu_n \leq ... \leq \mu_1$. Since no free variable occurs below I, s is a closed term and hence there is a number m s..t $\rightarrow s = \bar{m}$ is **PA**-provable without an essential cut or ind (cf. Lemma

2.6). Hence there is a proof Q of $F(\bar{m}) \to F(s)$ without an essential cut or ind (cf. Lemma 2.6). Let $P_0(\bar{n})$ be the proof which is obtained from P_0 by replacing a by $\bar{n}$ throughout. Consider the following proof P'

$$\dfrac{\dfrac{\begin{matrix} P_0(\bar{0}) \\ \vdots \\ S_1 : F(0), \Gamma \to \Delta, F(0') \end{matrix} \qquad \begin{matrix} P_0(\bar{1}) \\ \vdots \\ F(0'), \Gamma \to \Delta, F(0'') \end{matrix}}{S_2 : F(0), \Gamma \to \Delta, F(0'')} \qquad \begin{matrix} P_0(\bar{2}) \\ \vdots \\ F(0''), \Gamma \to \Delta, F(0''') \end{matrix}}{S_3 : F(0), \Gamma \to \Delta, F(0''')}$$

$$\dfrac{\begin{matrix} \vdots \\ S_m : F(0), \Gamma \to \Delta, F(\bar{m}) \end{matrix} \qquad \begin{matrix} Q \\ \vdots \\ F(\bar{m}) \to F(s) \end{matrix}}{\begin{matrix} S_0 : F(0), \Gamma \to \Delta, F(s) \\ \vdots \\ \to \end{matrix}}$$

where $S_1, S_2, \ldots, S_0$ denotes the sequents shown on their right, $S_1, \ldots, S_m$ all have height l, since the formulas $F(\bar{n})$, $n = 0, \ldots, m$ all have the same grade. Therefore

$$o(F(\bar{n}), \Gamma \to \Delta, F(\bar{n}'); P') = \mu \qquad \text{for } n = 0, 1, \ldots, m$$

Since Q has no essential cut or ind, $o(F(\bar{m}) \to F(s); P') = q < \omega$, $o(S_2) = \mu \# \mu$, $o(S_3) = \mu \# \mu \# \mu$; ..., and in general, writing $\mu * n = \mu \# \ldots \# \mu$ (n times), $o(S_n) = \mu * n$ for $n = 1, 2, \ldots, m$. Thus

$$o(S_0) = \omega_{l-k}(\mu * n + q)$$

and $\mu * m + q < \omega^{\mu_1 + 1}$, since $q < \omega$. Therefore

$$o(S_0; P') = \omega_{l-k}(\mu * m + q) < \omega_{l-k+1}(\mu_1) = o(S_0; P)$$

Thus $o(S_0; P') < o(S_0; P)$. Hence by Lemma 2.42 $o(P') < o(P)$

Thus if P has an ind the end-piece, we are done; we have reduced P to a proof P' of $\to$ with $o(P') < o(P)$. Otherwise we assume from now on that P has no ind in its end-piece, and go to Step 3.

Step 3. Suppose the end-piece of P contains a logical initial sequent $D \to D$. Since the end-sequent is empty, both D's (or more strictly, descendants of both D's) must disappear by cuts. Suppose that (a descendant of) the D in the antecedent is a cut formula first (viz. in the following figure a descendant of the D in the succedent of $D \to D$ occurs in Ξ)

$$\dfrac{\begin{matrix} \\ \vdots \\ \Gamma \to \Delta, D \end{matrix} \qquad \begin{matrix} D \to D \\ \vdots \\ D, \Pi \to \Xi \end{matrix}}{S : \Gamma, \Pi \to \Delta, \Xi} \\ \vdots \\ \to$$

P is reduced to the following P':

$$\begin{matrix} \vdots \\ \Gamma \to \Delta, D \\ \hline \text{weakenings and exchanges} \\ \hline S' : \Gamma, \Pi \to \Delta, \Xi \\ \vdots \\ \to \end{matrix}$$

Note that there is a cut whose cut formula is D below S since both Ds in $D \to D$ must disappear by cuts. Hence the height of $\Gamma \to \Delta, D$ does not change when we transform P into P': $o(S'; P') < o(S; P)$.

Hence by Lemma 2.42, $o(P') < o(P)$

We assume from now on that the end-piece of P contains no logical initial sequents

Step 4. Suppose there is a weakening in the end-piece. Let I be the lower most weakening inference in the end-piece. Since the end-sequent is empty, there must exists a cut, J, below I and the cut formula is the descendent of the principal formula of I.

$$\dfrac{\begin{matrix} \\ \\ \vdots \\ \Gamma \to \Delta, D \end{matrix} \qquad \begin{matrix} \vdots \\ \dfrac{\Pi' \to \Xi'}{D, \Pi' \to \Xi'}\, I \\ \vdots \\ D, \Pi \to \Xi \quad (k) \end{matrix}}{\Gamma, \Pi \to \Delta, \Xi \quad (l)}\, J \\ \vdots \\ \to$$

(a) If no contraction is applied to D from the inference I through J, by deleting some exchanges from P if necessary, reduce P into the following proof P'

$$
\dfrac{\dfrac{\begin{matrix} \vdots \\ \Pi' \to \Xi' \\ \vdots \\ \Pi \to \Xi \quad (l) \end{matrix}}{\text{weakenings and exchanges}}}{\begin{matrix} \Gamma, \Pi \to \Delta, \Xi \quad (l) \\ \vdots \\ \to \end{matrix}}
$$

Let $h(\Gamma, \Pi \to \Delta, \Xi; P) = l$ and $h(D, \Pi \to \Xi; P) = k$. Then $l \leq k$ and $h(\Pi \to \Xi; P') = h(\Gamma, \Pi \to \Delta, \Xi; P') = l$. Let S be a sequent in P above $D, \Pi \to \Xi$ and let S' be the corresponding sequent in P'. Then by the induction on number of inferences up to $D, \Pi \to \Xi$, we can show that

$$\omega_{k_1 - k_2}(o(S; P)) \geq o(S'; P')$$

where $k_1 = h(S; P)$ and $k_2 = h(S'; P)$ (). Hence if $o(\Gamma \to \Delta, D; P) = \mu_1$, $o(D, \Pi \to \Xi; P) = \mu_2$, $o(\Gamma, \Pi \to \Delta, \Xi) = \nu$, $o(\Pi \to \Xi; P) = \mu_2'$ and $o(\Gamma, \Pi \to \Delta, \Xi) = \nu'$, then

$$\omega_{k-l}(\mu_2) \geq \mu_2'$$

and further

$$\nu = \omega_{k-l}(\mu_2 \# \mu_1) > \omega_{k-l}(\mu_2) \geq \mu_2' = \nu'$$

Thus $o(P) > o(P')$

(b) If not the case 1, let the uppermost contraction applied to D be I'. Reduce P into the following proof Q:

$$
P : \begin{matrix} \vdots \\ \dfrac{\Pi' \to \Xi'}{D, \Pi' \to \Xi'} \\ \vdots \\ \dfrac{D, D, \Pi'' \to \Xi''}{D, \Pi'' \to \Xi''} \\ \vdots \\ D, \Pi \to \Xi \end{matrix}
\qquad
Q : \begin{matrix} \vdots \\ \Pi' \to \Xi' \\ \vdots \\ D, \Pi'' \to \Xi'' \\ \vdots \\ D, \Pi \to \Xi \end{matrix}
$$

Apparently, $o(P) = o(Q)$. Hence we can assume that the end-piece of P, contains no weakening.

Step 5. We can now assume that P is not its own end-piece, since otherwise it would be simple, and hence by Lemma 2.38 could not end with $\rightarrow$ $\square$

Lemma 2.44 (sublemma)**.** *Suppose that a proof in* ***PA****, say P, satisfies the following*

(a) *P is not its own end-piece*

(b) *The end-piece of P does not contain any logical inference, ind or weakening*

(c) *If an initial sequent belongs to the end-piece of P, then it does not contain any logical symbol*

Then there exists a suitable cut in the end-piece of P

Proof. Induction on the number of essential cuts in the end-piece of P

The end-piece of P contains an essential cut (), since P is not its own end-piece. Take a lowermost such cut, say I. If I is a suitable cut, then the lemma is proved. Otherwise, let P be of the form

$$I \; \dfrac{\begin{array}{c} \vdots P_1 \\ \Gamma \rightarrow \Delta, D \end{array} \qquad \begin{array}{c} \vdots P_2 \\ D, \Pi \rightarrow \Lambda \end{array}}{\Gamma, \Pi \rightarrow \Delta, \Lambda}$$

Since I is not a suitable cut, one of two cut formulas of I is not a descendant of the principal formula of a boundary inference. Suppose that D in $\Gamma \rightarrow \Delta, D$ is not a descendant of the principal formula of a boundary inference. Now we prove:

(a) P_1 contains a boundary inference of P

Suppose otherwise. Then D in $\Gamma \rightarrow \Delta, D$ is a descendant of D in an initial sequent in the end-piece of P by 2. This contradicts the assumption that I is an essential cut by 3.

(b) If an inference J in P_1 is a boundary inference of P, then J is a boundary inference of P_1

(c) P_1 is not its own end-piece and the end-piece of P_1 is the intersection of P_1 and the end-piece of P
this follows from 1 and 1,2 (condition first)

Then end-piece of P_1 has a suitable cut. This cut is suitable cut in the end-piece of P. $\square$

[**?**]. Returning to our proof P of $\rightarrow$ which satisfies the conclusion of steps 1-4, we have, as an immediate consequence of Sublemma 2.43 that the end-piece of P contains a suitable cut. We now define an **essential reduction** of P

Take a lowermost suitable cut in the end-piece of P, say I

Case 1. The cut formula of I is of the form $A \wedge B$. Suppose P is of the form

$$
I \; \dfrac{\begin{array}{c}\vdots \\ I_1\;\dfrac{\Gamma \rightarrow \Theta', A \qquad \Gamma \rightarrow \Theta', B}{\Gamma' \rightarrow \Theta', A \wedge B} \\ \vdots \\ \Gamma \xrightarrow{\mu} \Theta, A \wedge B\end{array} \qquad \begin{array}{c} I_2\;\dfrac{A, \Pi' \rightarrow \Lambda'}{A \wedge B, \Pi' \rightarrow \Lambda'} \\ \vdots \\ A \wedge B, \Pi \xrightarrow{\nu} \Lambda \quad (l)\end{array}}{\begin{array}{c}\Gamma, \Pi \rightarrow \Theta, \Lambda \\ \vdots \\ \Delta \xrightarrow{\lambda} \Xi \quad (k) \\ \vdots \\ \rightarrow \end{array}}
$$

where $\Delta \rightarrow \Xi$ denotes the uppermost sequent below I whose height is less than that of the upper sequents of I. Let l be the height of each upper sequent of I, and k that of $\Delta \rightarrow \Xi$. Then $k < l$. Note that $\Delta \rightarrow \Xi$ may be the lower sequent of I, or the end-sequent. The existence of such a sequent follows from Proposition **??**

$\Delta \rightarrow \Xi$ must be the lower sequent of a cut J (since there is no ind below I). Consider

$$
P_1 : \quad \begin{array}{c} J_1\;\dfrac{A \wedge B, \Pi \xrightarrow{\nu_1} \Lambda \quad (l)}{\Gamma, \Pi \rightarrow A, \Theta, \Lambda} \\ \vdots \\ \dfrac{\Delta \xrightarrow{\lambda_1} A, \Xi}{\Delta \rightarrow \Xi, A \quad (m)} \end{array}
$$

$\square$

3 TODO ALL the problems

1.41 2.17 7a

7 it has a cut, but why this cut is essential