Mathematisches Institut
der Ludwig-Maximilians-Universität München

Diplomarbeit

Minimal from Classical Proofs

vorgelegt von Christoph-Simon Senjak

betreut von Prof. Dr. H. Schwichtenberg

17. Oktober 2011

Hiermit erkläre ich, dass die vorliegende Arbeit von mir selbstständig verfasst wurde und dass keine anderen als die angegebenen Hilfsmittel benutzt wurden. Die Stellen der Arbeit, die anderen Werken dem Wortlaut oder Sinn nach entnommen sind, sind in jedem einzelnen Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht.

Preface

In this diploma thesis we want to have a practical view on the several approaches to get minimal or intuitionistic proofs from classical proofs satisfying some constraints. Those theorems are usually referred to as “Glivenko style” due to Glivenko's theorem, which states that if $P \to Q$ is a classical theorem, then $\neg\neg P \to \neg\neg Q$ is an intuitionistic theorem, but restricted to propositional logic.

Several theorems, like Barr's theorem, are going in this direction and are already widely used in constructive mathematics. On the other hand, Orevkov's theorem with its proofs appears to be widely unknown, though it seems to be a more general theory.

We will concentrate on one case of Orevkov's “complete Glivenko classes”, which are classes of sequents in which classical and intuitionistic derivability are equivalent, sketch Orevkov's and Nadathur's proof for this complete Glivenko class, and give an own proof of its sufficiency using only elementary transformations of proofs given in natural deduction.

We will also give an example from actual constructive mathematics, a constructive version of Heitmann's theorem for rings, by Coquand and Lombardi. Even though this example has already been researched, it shows what type of problems Orevkov's theorem could be applied to.

One should keep in mind that this might very well be the start of some useful theory, not the end.

At this point many thanks go to Prof. Dr. Helmut Schwichtenberg for supervising this diploma thesis. Furthermore I am indebted to Matthias Benkard, Daniel El-Baz, Nicolai Kraus, Christian Sattler and Matvey Soloviev for reading over this thesis and assisting me with improving the exposition.

Contents

1 Preliminary Definitions

We deal with different approaches for the same or similar things. Therefore we aim to point out the similarities as well as the differences between them. We will try to unify the concepts and notations of all the references we give. This is not always as simple as one would expect, especially since "portability" does not seem to be a concept that has reached the mathematical world yet.

This list of definitions has the purpose of clarifying what we mean by the several concepts and notations we will use. However, it does not have the purpose of explaining them in detail in a didactic way, as this is outside the scope of this thesis. For more detailed discussions, the references should be considered.

1.1 The Language

We use the usual notation of first order logic. We have the logical symbols $\{\forall, \exists, \wedge, \vee, \rightarrow\}$, infinitely many variable symbols $\{x_0, x_1, x_2, x_3, \ldots\}$, for every $n \in \mathbb{N}_0$ infinitely many n-ary function symbols $\{f_0^n, f_1^n, f_2^n, f_3^n, \ldots\}$ and infinitely many n-ary relation symbols $\{R_0^n, R_1^n, R_2^n, R_3^n, \ldots\}$. Every variable is a term, and if $t_1, \ldots, t_n$ are terms, then $f_i^n t_1 \ldots t_n$ is a term for suitable i. If $t_1, \ldots, t_n$ are terms, then $R_i^n t_1 \ldots t_n$ is an atomic formula (or atom) for suitable i. As usual, we use other symbols than R_i^n and f_i^n and x_i to denote relations, functions and variables. Their arity should always be clear. Every atomic formula is a formula, and if A and B are formulae and x is a variable, then $A \rightarrow B$, $A \wedge B$, $A \vee B$, $\forall_x A$ and $\exists_x A$ are formulae.

In some of the systems, we have an additional connective $\neg$, such that $\neg A$ is also a formula if A is a formula. But mostly, we define a special nullary relation $\bot$, the *falsum* (with the intuition that this is a formula which is always wrong) and define $\neg A$ by $A \rightarrow \bot$. In the definitions below, we therefore include the case for $\neg$. Furthermore, some of the systems have an additional nullary relation symbol $\top$, the *verum*, with the intuition that this is a formula that is always true. Some systems, for example in (5), do not regard $\bot$ or $\top$ as atomic formulae. We will however always do this, as it does not make a difference anywhere.

The function *var* maps terms to the set of the variables they contain, which can be recursively defined by $\mathit{var}(x_i) = \{x_i\}$, $\mathit{var}(f_i^n t_1 \ldots t_n) = \bigcup_j \mathit{var}(t_j)$. The function *FV* maps formulae to the set of their free variables, which can be recursively defined by $\mathit{FV}(R_i^n t_1 \ldots t_n) = \bigcup_j \mathit{var}(t_j)$, $\mathit{FV}(A \rightarrow B) = \mathit{FV}(A \wedge B) = \mathit{FV}(A \vee B) = \mathit{FV}(A) \cup \mathit{FV}(B)$, $\mathit{FV}(\forall_x A) = \mathit{FV}(\exists_x A) = \mathit{FV}(A) \backslash \{x\}$, $\mathit{FV}(\neg A) = \mathit{FV}(A)$.

The substitution $A[x_1 := t_1, \ldots, x_n := t_n] = A[(x_i)_{1 \leq i \leq n} := (t_i)_{1 \leq i \leq n}]$ for terms and formulae A is recursively defined by

- $x_j[(x_i)_{1 \leq i \leq n} := (t_i)_{1 \leq i \leq n}] = t_j$ for $1 \leq j \leq n$
- $y[(x_i)_{1 \leq i \leq n} := (t_i)_{1 \leq i \leq n}] = y$ for variables y not occurring in $(x_i)_{1 \leq i \leq n}$

- $(f^n_i t_1 \dots t_n)[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}] = f^n_i(t_1[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}]) \dots (t_n[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}])$
- $(R^n_i t_1 \dots t_n)[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}] = R^n_i(t_1[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}]) \dots (t_n[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}])$
- $(\neg A)[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}] := \neg(A[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}])$
- $(A \circ B)[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}] = A[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}] \circ B[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}]$ for $\circ \in \{\wedge, \vee, \rightarrow\}$
- $(\mathfrak{Q}_x A)[(x_i)_{1\le i\le n} := (t_i)_{1\le i\le n}] = \mathfrak{Q}(A[(x_i)_{1\le i\le n} := (u_i)_{1\le i\le n}]$ for $\mathfrak{Q} \in \{\forall, \exists\}$ where $u_i = \begin{cases} t_i & \text{for } x \neq x_i \\ x & \text{otherwise} \end{cases}$

where we assume that the t_i do not contain variables that are bound in a context where the x_i is unbound, saying t_i *is free for* x_i, and if so, we rename these bound variables appropriately. If we want to express that a variable may occur freely in multiple elements of a theory Γ, we call that variable a *parameter*, and if we want to stress that it has a special meaning in the theory (and thus usually not bind it), *constant* of a theory Γ. Usually, parameters will be disjoint from all variables that are bound in proofs.

1.2 Subformulae

Definition We need the concept of *subformulae SF*, *positive* subformulae *PSF* and *negative* subformulae *NSF* of a formula, which we define recursively by

- $SF(\bot) = PSF(\bot) = NSF(\bot) = \{\bot\}$.
- $SF(R^n_i t_1 \dots t_n) = PSF(R^n_i t_1 \dots t_n) = \{R^n_i t_1 \dots t_n\}$, $NSF(R^n_i t_1 \dots t_n) = \emptyset$.
- $f(A \circ B) = f(A) \cup f(B) \cup \{A \circ B\}$ for $f \in \{SF, PSF\}$, $NSF(A \circ B) = NSF(A) \cup NSF(B)$, for $\circ \in \{\wedge, \vee\}$.
- $SF(\neg A) = \{\neg A\} \cup SF(A)$, $PSF(\neg A) = \{\neg A\} \cup NSF(A)$, $NSF(\neg A) = PSF(A)$ if we have $\neg$ as own connective (otherwise like $A \rightarrow \bot$).
- $SF(A \rightarrow B) = \{A \rightarrow B\} \cup SF(A) \cup SF(B)$, $PSF(A \rightarrow B) = \{A \rightarrow B\} \cup NSF(A) \cup PSF(B)$, $NSF(A \rightarrow B) = PSF(A) \cup NSF(B)$.
- $SF(\mathfrak{Q}_x A) = \{\mathfrak{Q}_x A\} \cup \bigcup_{t \text{ term}} SF(A[x := t])$, $NSF(\mathfrak{Q}_x A) = \bigcup_{t \text{ term}} NSF(A[x := t])$, $PSF(\mathfrak{Q}_x A) = \{\mathfrak{Q}_x A\} \cup \bigcup_{t \text{ term}} PSF(A[x := t])$, for $\mathfrak{Q} \in \{\forall, \exists\}$.

1.3 Height, Length

The height of a tree is the length of its longest branch, the length of a tree is the number of its nodes (with the intuition that this is about the length of a string representation of this tree). As formulae are trees, the following definitions can be given.

$$len(A) = \begin{cases} 1 & \text{for atomic } A \\ 1 + len(B) + len(C) & \text{for } A = B \circ C \text{ with } \circ \in \{\rightarrow, \vee, \wedge\} \\ 1 + len(B) & \text{for } A = \mathfrak{Q}xB \text{ with } \mathfrak{Q} \in \{\forall, \exists\} \end{cases}$$

$$hgt(A) = \begin{cases} 1 & \text{for atomic } A \\ 1 + max(hgt(B), hgt(C)) & \text{for } A = B \circ C \text{ with } \circ \in \{\rightarrow, \vee, \wedge\} \\ 1 + hgt(B) & \text{for } A = \mathfrak{Q}xB \text{ with } \mathfrak{Q} \in \{\forall, \exists\} \end{cases}$$

Similarly for proof trees which will be defined in Section 1.6

$$len(t) = \begin{cases} 1 & \text{for variables } t \\ 1 + len(q) & \text{for } t = \vee_+ q, t = \lambda_x q, t = \exists_+ uq, t = q^{\forall_x A} u \\ 1 + len(q) + len(r) & \text{for } t = q^{A \rightarrow B} r^A, t = \langle q, r \rangle, t = q(r) \\ 1 + len(p) + len(q) + len(r) & \text{for } t = p(q, r) \end{cases}$$

$$hgt(t) = \\ \begin{cases} 1 & \text{for variables } t \\ 1 + hgt(q) & \text{for } t = \vee_+ q, t = \lambda_x q, t = \exists_+ uq, t = q^{\forall_x A} u \\ 1 + max(hgt(q), hgt(r)) & \text{for } t = q^{A \rightarrow B} r^A, t = \langle q, r \rangle, t = q(r) \\ 1 + max(hgt(p), hgt(q), hgt(r)) & \text{for } t = p(q, r) \end{cases}$$

1.4 Entailment

We will use the $\vdash$-character to denote several different *entailment relations*. For example, $\vdash_m$, $\vdash_i$, $\vdash_{tc}$ and $\vdash_{ec}$ shall denote minimal, intuitionistic, traditional classical and extended classical derivability. These concepts are well-known and have several equivalent definitions. We will give one definition in Section 1.6.

$\vdash$ without any index shall denote derivability that is not clearly specified, for example when the context is applicable to all of $\vdash_m$, $\vdash_i$, $\vdash_{ec}$ and $\vdash_{tc}$.

1.5 Sequent Calculi

We will have to deal with several sequent calculi. There are a few things they all have in common.

A *sequent* is a pair $(\Gamma; \Delta)$ of finite multisets Γ and Δ of formulae, where the first component Γ is called *antecedent* and the second component Δ is called *succedent*. If the succedent is a singleton, the sequent is called *singular*, but notice that this is **not** the same as saying that it only contains one formula, $(\Gamma; \{A, A\})$ is not singular, while $(\Gamma; \{A\})$ is. While there are several notations for separators of antecedent and succedent, we will use the $\Rightarrow$ character. Other notations that can be found in some of the references use $\neg\!\!\rightarrow$, $\bot\!\!\rightarrow$, $\vdash$ and $\rightarrow$, exchanging $\rightarrow$ and $\supset$ in formulae. Thus, sequents are usually denoted by $\Gamma \Rightarrow \Delta$.

If we talk about positivity or negativity of formulae regarding a sequent $\Gamma \Rightarrow \Delta$, we shall consider every formula in Γ as negative and every formula in Δ as positive, with the appropriate consequences for their subformulae. Thus, we consider $\Rightarrow$ similar to an implication.

A common (informal) intuition is that the antecedent is interpreted conjunctively, while the succedent is interpreted disjunctively.

The multisets are usually denoted by capital Greek letters which differ from capital Latin letters, mostly Γ and Δ; formulae are considered as singletons; a comma is used as union operator.

To denote that a sequent was derived by a special calculus, we write $\vdash \Gamma \Rightarrow \Delta$, where $\vdash$ may be replaced by something more specific. We usually give it an index that points out the calculus it was derived from, but we may leave that index out if it is clear from the context, or if we give properties that are applicable to multiple entailment relations.

The rules are given in patterns of the form

$$\frac{\mathfrak{M}_1 \ldots \mathfrak{M}_n}{\mathfrak{N}}\,\mathfrak{a}$$

where $\mathfrak{a}$ denotes an identifier of the rule and may be left out as well, $\mathfrak{M}_1, \ldots, \mathfrak{M}_n$ are sequents that have already been derived, and $\mathfrak{N}$ is the sequent that can be derived from them. Rules with $n = 0$, that is, rules of the form

$$\frac{}{\mathfrak{N}}\,\mathfrak{a}$$

will be called *(logical) axioms*. $\mathfrak{N}$ will be called the *conclusion* of a rule, the $\mathfrak{M}_i$ and additional constraints for that rule will be called the *premises*. The $\mathfrak{M}_i$ will be called the *sequent premises*.

We will sometimes be sloppy and write additional premises for a rule to apply next to the $\mathfrak{M}_i$ when there is no danger of confusion, for example

$$\frac{A \text{ atomic}}{A, \Gamma \Rightarrow \Delta, A}$$

Usually, the $\mathfrak{M}_i$ are patterns of the form $\Gamma, A_1, \ldots, A_n \Rightarrow \Delta, B_1, \ldots, B_m$ or similar, where the Γ and Δ denote parts that do not influence the behavior of the given rule (except for variable conditions they must satisfy), while the explicitly given formulae $A_1, \ldots, A_n, B_1, \ldots, B_m$ are relevant. These parts that mostly do not influence the behavior are called the *context* of the rule, the explicitly stated formulae are called differently throughout different calculi, as there might be a finer distinction between them, in general we will therefore refer to them as formulae outside the context or *explicit formulae* of that rule. We may add $\vdash$-characters to the patterns.

If we want to make statements about the height of the proofs of sequents, we may add a superscript to the entailment relation, telling about the height. $\vdash^n \Gamma \Rightarrow \Delta$ shall mean that $\Gamma \Rightarrow \Delta$ is derivable in a proof of height $\leq n$. Therefore, $\vdash^n \Gamma \Rightarrow \Delta$ implies $\vdash^m \Gamma \Rightarrow \Delta$ for all $m > n$, and for axiomatic rules we have

$$\frac{}{\vdash^0 \Gamma \Rightarrow \Delta}$$

for rules with one premise we have

$$\frac{\vdash^n \Gamma \Rightarrow \Delta}{\vdash^{n+1} \Gamma' \Rightarrow \Delta'}$$

and for rules with two premises we have

$$\frac{\vdash^m \Gamma' \Rightarrow \Delta' \qquad \vdash^n \Gamma'' \Rightarrow \Delta''}{\vdash^{1+max(m,n)} \Gamma \Rightarrow \Delta}$$

1.6 The Calculus of Natural Deduction

The only calculus we will use which is usually not denoted as a sequent calculus is the *calculus of natural deduction*, as described in (11) (however, of course, it can as well be formulated as a sequent calculus, and its main differences to the other sequent calculi is that it mainly operates on the succedent, and formulae can also be decomposed). It has a notation of proofs as trees and a notation of proofs as lambda terms. As we will use both, the following table gives both notations.

$u : A$	u^A
$\dfrac{\begin{array}{c}[u : A] \\ \mid M \\ B\end{array}}{A \to B} \to_+$	$(\lambda u^A M^B)^{A \to B}$

$$\dfrac{\overset{|M}{A\to B} \qquad \overset{|N}{A}}{B}\to_- \qquad\qquad (M^{A\to B}N^{A})^{B}$$

$$\dfrac{\overset{|M}{A} \qquad x\notin FV(M\backslash\{A\})}{\forall_x A}\forall_+ \qquad\qquad (\lambda x M^{A})^{\forall_x A}$$

$$\dfrac{\overset{|M}{\forall_x A} \qquad r}{A[x:=r]}\forall_- \qquad\qquad (M^{\forall_x A}r)^{A[x:=r]}$$

$$\dfrac{\overset{|M}{A_i} \qquad i\in\{0,1\}}{A_0\vee A_1}\vee_+ \qquad\qquad (v_{+,i,A_{1-i}}M^{A_i})^{A_0\vee A_1}$$

$$\dfrac{\overset{|M}{A\vee B} \qquad \overset{[u:A]}{\overset{|N}{C}} \qquad \overset{[v:B]}{\overset{|P}{C}}}{C}\vee_- \qquad\qquad (M^{A\vee B}(u^{A}.N^{C})(v^{B}.P^{C}))^{C}$$

$$\dfrac{\overset{|M}{A} \qquad \overset{|N}{B}}{A\wedge B}\wedge_+ \qquad\qquad \left\langle M^{A},N^{B}\right\rangle^{A\wedge B}$$

$$\dfrac{\overset{|M}{A_1\wedge A_2} \qquad \overset{[u:A_1][v:A_2]}{\overset{|N}{B}}}{B}\wedge_- \qquad\qquad (M^{A_1\wedge A_2}(u^{A_1},v^{A_2}.N^{B}))^{B}$$

$$\dfrac{t \qquad \overset{|M}{A[x:=t]}}{\exists_x A}\exists_+ \qquad\qquad (\exists_{+,x}tM^{A[x:=t]})^{\exists_x A}$$

$$\dfrac{\overset{|M}{\exists_x A} \qquad \overset{[u:A]}{\overset{|N}{B}} \qquad \text{"var cond"}}{B}\exists_- \qquad\qquad (M^{\exists_x A}(u^{A},x.N^{B}))^{B}$$

The left premises of $\to_-$, $\vee_-$ and $\exists_-$ are called *major premises* of the applied rule, the right premises are called *minor premises*. Notice that $\forall_+$ and $\exists_-$ have variable conditions; "var cond" of $\exists_-$ means that x must not occur freely in B, and not freely in any free assumption of N except A.

Definition We define the function *FA* of *free assumption variables* over proofs by

- $FA(u^A) := \{u\}$
- $FA((\lambda u^A M^B)^{A\to B}) := FA(M)\backslash\{u\}$
- $FA((M^{A\to B}N^A)^B) := FA(M) \cup FA(N)$
- $FA((\lambda x M^A)^{\forall_x A}) := FA(M)$
- $FA((M^{\forall_x A}r)^{A[x:=r]}) := FA(M)$
- $FA((\nu_{+,i,A_{1-i}} M^{A_i})^{A_0 \vee A_1}) := FA(M)$
- $FA((M^{A\vee B}(u^A.N^C)(\nu^B.P^C))^C) := FA(M) \cup (FA(N)\backslash\{u\}) \cup (FA(P)\backslash\{\nu\})$
- $FA(\langle M^A, N^B\rangle^{A\wedge B}) := FA(M) \cup FA(N)$
- $FA((M^{A_1\wedge A_2}(u^{A_1}, \nu^{A_2}.N^B))^B) := FA(M) \cup (FA(N)\backslash\{u,\nu\})$
- $FA((\exists_{+,x} t M^{A[x:=t]})^{\exists_x A}) := FA(M)$
- $FA((M^{\exists_x A}(u^A, x.N^B))^B) := FA(M) \cup (FA(N)\backslash\{u\})$

Definition We say that A is *derivable in minimal logic* from a set or multiset of formulae Γ, writing $\Gamma\vdash_m A$, if there is a proof in natural deduction with end formula A and which has only free assumptions that are also in Γ. We say it is *derivable in intuitionistic logic* or *intuitionistically derivable*, writing $\Gamma\vdash_i A$, if $EFQ \cup \Gamma\vdash_m A$, where $EFQ := \{\forall_{\vec{x}}.\bot\to P\vec{x} | P \text{ relation symbol}\}$. We say it is *derivable in (traditional) classical logic* or *(traditional-) classically derivable*, writing $\Gamma\vdash_{tc} A$, if $STAB \cup \Gamma\vdash_m A$, where $STAB := \{\forall_{\vec{x}}.\neg\neg A\to A | A \text{ formula}, \vec{x} \supseteq FV(A)\}$. We say it is *derivable in extended classical logic* or *extended classically derivable* writing $\Gamma\vdash_{ec} A$, if $ASTAB \cup \Gamma\vdash_m A$, where $ASTAB := \{\forall_{\vec{x}}.\neg\neg P\vec{x}\to P\vec{x} | P \text{ relation symbol}\}$.

Furthermore, we write $\Gamma \vdash \Delta$ to denote that $\Gamma \vdash A$ for all $A \in \Delta$, where $\vdash \in \{\vdash_m, \vdash_i, \vdash_{tc}, \vdash_{ec}\}$.

It is known that $EFQ\vdash_m \bot\to A$ for all formulae A, which is why we only put $\bot\to A$ for atomic A into *EFQ*. Unfortunately, the same cannot be done with *STAB* in general, which is why we distinguish traditional and extended classical derivability.

Extended classical logic can be seen as another view on classical logic. In the historic context, classical logic is usually defined according to our definition of traditional classical logic. Therefore we will have a closer look at extended classical logic in the part based on our own work in (10), in Section 4.

Definition By the *sequent* $seq(t^A)$ of a proof t^A we shall mean the pair $\{B|u^B \in FA(t), u \text{ some variable name}\} \Rightarrow A$, denoted as a sequent. Trivially, every proof proves its sequent considered as the appropriate entailment.

We extend the concepts of positivity and negativity of formula occurrences to sequents of proofs in the obvious way with $\Rightarrow$ being considered as implication, and all premises of the sequent being considered negative, and the conclusion being considered positive.

2 Barr's Theorem

We first have a close look at Barr's theorem, which gives one classification of theories for which classical derivability implies intuitionistic derivability, namely the geometric theories.

Definition A formula is called a *geometric formula*, if it does not contain $\to$ and $\forall$. A formula is called a *geometric implication*, if it is of the form $\forall_{\vec{x}}(A \to B)$ where A and B are geometric formulae. A theory Γ is called a *geometric theory*, if it only contains geometric implications.

Theorem 2.1 (Barr's theorem). *Let A be a geometric implication, and Γ be a geometric theory. Then $\Gamma \vdash_{tc} A$ implies $\Gamma \vdash_i A$.*

A lemma which is used in (5) and stated in (8) is

Lemma 2.2. *Every geometric theory is intuitionistically equivalent to a theory where all the axioms have the form*

$$\forall_{\vec{x}}.P_0 \to \exists_{\vec{y}}.P_1 \vee \ldots \vee P_n$$

where the P_i are conjunctions of atomic formulae.

We may additionally require that $\vec{y}$ is non-empty, since we know that $(\exists_y A) \leftrightarrow A$ for $y \notin FV(A)$.

2.1 Palmgren's Proof

We will look at a slight modification of the proof given in (8), which uses the Gödel-Gentzen negative translation and the Dragalin-Friedman translation. We will generalize the Gödel-Gentzen negative translation, avoiding the use of the

Dragalin-Friedman translation. However, we still need ex falso quodlibet, since our generalization of the Gödel-Gentzen negative translation cannot embed classical logic into minimal logic, but only into intuitionistic logic. For the sake of completeness, we will give a definition of the Dragalin-Friedman translation, although we will not use it.

Definition The *Dragalin-Friedman Translation* for a formula is recursively defined as:

- $A^C = A \vee C$ for atomic A
- $(A \circ B)^C = A^C \circ B^C$ for $\circ \in \{\rightarrow, \vee, \wedge\}$
- $(\mathfrak{Q}xA)^C = \mathfrak{Q}xA^C$ for $\mathfrak{Q} \in \{\forall, \exists\}$

2.1.1 The continuation translation

Palmgren uses the Gödel-Gentzen negative translation which gives an embedding of classical logic into minimal logic. We will, however, use a generalization which needs intuitionistic logic.

Definition The *continuation translation* $\cdot_A$ for a formula A is recursively defined as:

- $\perp_A = A$
- $P_A = (P \rightarrow A) \rightarrow A$ for P atomic, $P \neq \perp$
- $(B \vee C)_A = ((B_A \rightarrow A) \wedge (C_A \rightarrow A)) \rightarrow A$
- $(\exists_x B)_A = (\forall_x . B_A \rightarrow A) \rightarrow A$
- $(B \circ C)_A = B_A \circ C_A$ for $\circ \in \{\wedge, \rightarrow\}$
- $(\forall_x B)_A = \forall_x B_A$

This is a generalization of the Gödel-Gentzen negative translation:

Definition The *Gödel-Gentzen negative translation* A^g for formulae A can be defined by $A^g := A_\perp$.

As usual, we set $\Gamma_A := \{B_A | B \in \Gamma\}$. Trivially, B_A can be calculated from B by an algorithm in linear time relative to $\mathtt{len}(B)$.

Lemma 2.3. *Let A not contain free variables that occur bound in C, and C not contain implications or universal quantifiers. Then $C_A \leftrightarrow ((C \rightarrow A) \rightarrow A)$ is derivable in intuitionistic logic. The proof tree t can be generated by an algorithm in linear time relative to* $\mathrm{len}(C)$*, and there is an m such that* $\mathrm{len}(t) \leq m \cdot \mathrm{len}(C)$ *for all C.*

Proof By structural induction. Denote by $M(C)^{C_A \to (C\to A)\to A}$ and $N(C)^{((C\to A)\to A)\to C_A}$ the generated proofs. We will use this notation for recursion.

- $M(\bot) = \lambda_{u^A}\lambda_{v^{\bot\to A}} u$, $N(\bot) = \lambda_{u^{(\bot\to A)\to A}}.uEfq_A$.
- For atomic $C \neq \bot$: $M(C) = N(C) = \lambda_{u^{(C\to A)\to A}} u$.
- $M(\exists_x C) = \lambda_{w^{(\forall x.C_A \to A)\to A}}\lambda_{v^{(\exists_x C)\to A}}.w\lambda_x\lambda_{y^{C_A}}.M(C)y\lambda_{u^C}.v(\exists_{+,x}u)$,
  $N(\exists_x C) = \lambda_{y^{(\exists_x C\to A)\to A}}\lambda_{u^{\forall x.C\to A}}.y\lambda_{w^{\exists_x C}}.w(v^C.uxv)$.
- $M(B\vee D) = \lambda_{x^{(B\vee D)_A}}\lambda_{v^{(B\vee D)\to A}}.$
  $x\left\langle\lambda_{w^{B_A}}.M(B)w\lambda_{u^B}.v(\vee_{+,r,D}u), \lambda_{w^{D_A}}.M(D)w\lambda_{u^D}.v(\vee_{+,l,B}u)\right\rangle$,
  $N(B\vee D) = \lambda_{x^{((B\vee D)\to A)\to A}}\lambda_{u^{(B\to A)\wedge(D\to A)}}.$
  $x\lambda_{w^{B\vee D}}.w(v_1^B.(u(s^{B\to A}, t^{D\to A}.s))v_1, v_2^D.(u(s^{B\to A}, t^{D\to A}.t))v_2)$.
- $M(B\wedge D) = \lambda_{x^{B_A\wedge D_A}}\lambda_{w^{(B\wedge D)\to A}}.$
  $(x(s^{B_A}, t^{D_A}.M(D)t))\lambda_{v^D}.(x(s^{B_A}, t^{D_A}.M(B)s))\lambda_{u^B}.w\langle u, v\rangle$,
  $N(B\wedge D) = \lambda_{w^{((B\wedge D)\to A)\to A}}.$
  $\left\langle N(B)\lambda_{v^{B\to A}}.w\lambda_{u^{B\wedge D}}.v(u(s^B, t^D.s)), N(D)\lambda_{v^{D\to A}}.w\lambda_{u^{B\wedge D}}.v(u(s^B, t^D.t))\right\rangle$.

Trivially the resulting algorithm is linear relative to *len*(C). In every case, the size of the additional tree nodes is bounded, hence, take as m the maximum number of additional tree nodes in a step. □

Remark Only for the case $\bot$ we need intuitionism, which we do not need for the Gödel-Gentzen negative translation in this special case.

Lemma 2.4. *The following formulae are derivable in minimal logic for every* A, B *and* C. *The length of the proof trees and the time consumed to produce them does not depend on the formulae* A, B *and* C.

1. $(((A\to B)\to B)\to B)\to(A\to B)$
2. $(A\to B\to C) \leftrightarrow ((A\wedge B)\to C)$
3. $A\to((A\to C)\to C)$
4. $(((A\to C)\to C)\to A)\to(((B\to C)\to C)\to B)\to(((A\wedge B)\to C)\to C)\to(A\wedge B)$
5. $(((B\to C)\to C)\to B)\to(((A\to B)\to C)\to C)\to A\to B$
6. $\forall_x.(((A\to C)\to C)\to A)\to(((\forall_x A)\to C)\to C)\to A$
7. $(B\to C)\to((B\to A)\to A)\to(C\to A)\to A$

Proof

$$\dfrac{\dfrac{((A\to B)\to B)\to B \qquad \dfrac{\dfrac{[u : A\to B] \qquad [x : A]}{B}}{(A\to B)\to B}\,u}{B}}{A\to B}\,x$$

$$\dfrac{A\land B \qquad \dfrac{\dfrac{A\to B\to C \qquad [u:A]}{B\to C} \qquad [v:B]}{C}}{C}\,u,v$$

$$\dfrac{\dfrac{\dfrac{(A\land B)\to C \qquad \dfrac{[x:A] \qquad [y:B]}{A\land B}}{C}}{B\to C}\,y}{A\to B\to C}\,x$$

$$\dfrac{\dfrac{[u:A\to C] \qquad A}{C}}{(A\to C)\to C}\,u$$

$$\dfrac{\dfrac{((A\to C)\to C)\to A \qquad \dfrac{\dfrac{((A\land B)\to C)\to C \qquad \dfrac{\dfrac{[x:A\to C] \qquad \dfrac{[y:A\land B] \qquad [s:A]}{A}}{C}}{(A\land B)\to C}\,y}{C}}{(A\to C)\to C}\,x}{A} \qquad \begin{matrix}\vdots\\ B\end{matrix}}{A\land B}$$

$$\dfrac{\dfrac{((B\to C)\to C)\to B \qquad \dfrac{\dfrac{((A\to B)\to C)\to C \qquad \dfrac{\dfrac{[x:B\to C] \qquad \dfrac{[y:A\to B] \qquad [w:A]}{B}}{C}}{(A\to B)\to C}\,y}{C}}{(B\to C)\to C}\,x}{B}}{A\to B}\,w$$

$$\dfrac{((A\to C)\to C)\to A \qquad \dfrac{\dfrac{((\forall_x A)\to C)\to C \qquad \dfrac{\dfrac{[v:A\to C] \qquad \dfrac{[u:\forall_x A]}{A}}{C}}{(\forall_x A)\to C}\,u}{C}}{(A\to C)\to C}\,v}{A}$$

$$\dfrac{\dfrac{(B\to A)\to A \qquad \dfrac{\dfrac{[w : C\to A] \qquad \dfrac{B\to C \qquad [x : B]}{C}}{A}}{B\to A}\,x}{A}}{(C\to A)\to A}\,w$$

□

Lemma 2.5. *Let* A *not contain free variables that occur bound in* B*. Then* $(\neg\neg B\to B)_A$ *is derivable in intuitionistic logic. The algorithm generating that proof is linear relative to* $\mathrm{len}(B)$*, the length of the generated proof tree depends linearly on* $\mathrm{len}(B)$*.*

Proof By structural induction. For $B = \bot$, we have $(\neg\neg\bot\to\bot)_A = (((A\to A)\to A)\to A)$ which is derivable by $\lambda_{u^{(A\to A)\to A}} u \lambda_{v^A} v$. For all other atomic B, we have $(\neg\neg B\to B)_A = ((((B\to A)\to A)\to A)\to A)\to((B\to A)\to A)$ which follows by Lemma 2.4/1. For $B = \exists_x C$ we have $(\neg\neg B\to B)_A = ((((\forall_x.C\to A)\to A)\to A)\to A)\to(\forall_x.C\to A)\to A$ which also follows by Lemma 2.4/1, same for $B = C \vee D$, $(\neg\neg B\to B)_A = (((((C\to A)\wedge(D\to A))\to A)\to A)\to A)\to((C\to A)\wedge(D\to A))\to A$. For $B = C \wedge D$ we may assume $(\neg\neg C\to C)_A = ((C_A\to A)\to A)\to C_A$ and $(\neg\neg D\to D)_A = ((D_A\to A)\to A)\to D_A$ by induction and apply Lemma 2.4/4, similarly for $B = \forall_x C$, $(\neg\neg C\to C)_A$ with Lemma 2.4/6. For $B = C\to D$ we may assume $(\neg\neg D\to D)_A$ and apply Lemma 2.4/5.

As we are doing recursion over all connectives, the algorithm to prove this takes at most $\mathrm{len}(B)$ steps. □

Theorem 2.6. *Let* A *be a formula not containing free variables that occur bound in* Γ, B*. Then there is an algorithm transforming a classical proof* t *of* $\Gamma\vdash_{tc}B$ *into an intuitionistic proof* t_m *of* $\Gamma_A\vdash_i B_A$*. Let* Q *be the longest formula occurring in* t*. Then the algorithm takes linear time relative to* $\mathrm{len}(t)\cdot\mathrm{len}(Q)$*, and* $\mathrm{len}(t_m)$ *depends linearly on* $\mathrm{len}(t)\cdot\mathrm{len}(Q)$*.*

Proof Denote by $E(r)$ the converted term r. We give a recursive algorithm for E:

- $E(u^B) = B_A$ if B is not a stability axiom, and otherwise like in Lemma 2.5
- $E((\lambda_{x^B} M^C)^{B\to C}) = (\lambda_{x^{B_A}} E(M)^{C_A})^{B_A\to C_A = (B\to C)_A}$
- $E(\langle M^B, N^C\rangle^{B\wedge C}) = \langle E(M)^{B_A}, E(N)^{C_A}\rangle^{B_A\wedge C_A = (B\wedge C)_A}$
- $E((\lambda_x M^B)^{\forall_x B}) = (\lambda_x E(M)^{B_A})^{\forall_x B_A = (\forall_x B)_A}$
- $E((M^{B\to C} N^B)^C) = (E(M)^{(B\to C)_A = B_A\to C_A} E(N)^{B_A})^{C_A}$
- $E((M^{B\wedge C}(u^B, v^C.N^D))^D) = (E(M)^{(B\wedge C)_A = B_A\wedge C_A}(u^{B_A}, v^{C_A}.E(N)^{D_A}))^{D_A}$
- $E((\vee_{+,r,C} M^B)^{B\vee C}) = (\lambda_{u^{B_A\to A}} \lambda_{v^{C_A\to A}} u E(M)^{B_A})^{(B\vee C)_A}$, similarly for $\vee_{+,l}$

- $\mathrm{E}((M^{\forall_x B} r)^{B[x:=r]}) = (\mathrm{E}(M)^{(\forall_x B)_A = \forall_x B_A} r)^{B_A[x:=r]}$, this is allowed because $A[x := r] = A$, since x occurs in Γ, B.
- $\mathrm{E}((M^{B\vee C}(u^B.N^D, u^C.P^D))^D) = \mathrm{E}(u^{\neg\neg D \to D})\lambda_{w^{D_A \to A}}($ $\mathrm{E}(M)^{((B_A \to A)\wedge(C_A \to A))\to A} \langle \lambda_{u^{B_A}}.w(u.\mathrm{E}(N)^{D_A}), \lambda_{v^{C_A}}.w(v.\mathrm{E}(P)^{D_A})\rangle)$
- $\mathrm{E}(M^{\exists_x B}(u^B.N^C)) = \mathrm{E}(u^{\neg\neg C \to C})\lambda_{w^{C_A \to A}}.\mathrm{E}(M)^{(\forall_x.B_A \to A)\to A}\lambda_x \lambda_{u^{B_A}}.w(u.\mathrm{E}(N)^{C_A})$ where the variable condition for the $\forall_+$ is satisfied since otherwise there would have to be a free assumption other than u in N which contains x freely and then $M^{\exists_x B}(u^B.N^C)$ would not be valid anymore

In every step, the generated tree consists of the recursively calculated subtrees expanded by a constant number of tree nodes or a constant number of applications of Lemma 2.5. From this follows the linearity. □

Lemma 2.7. $\Gamma \vdash_i \Gamma_A$ *for geometric theories* Γ *and* A *not containing variables occurring in* Γ*. Generating a proof of this takes linear time relative to* $\sum_{A\in\Gamma} \mathrm{len}(A)$*.*

Proof Every formula of Γ has the form $\forall \vec{x}.B \to C$ where B and C are geometric formulae. Using Lemma 2.4/7 with Lemma 2.3 we get a proof $t^{(B\to C)\to B_A \to C_A}$. Then we can use $(\lambda_{\vec{x}}.t(u^{\forall \vec{x}.B \to C}\vec{x}))^{(\forall \vec{x}.B\to C)_A}$ to derive the continuation translation. We can do this for all axioms in Γ. As applying Lemma 2.4/7 takes constant time, and applying Lemma 2.3 takes linear time, generating one such proof takes linear time. Hence, proving this takes linear time relative to the sum of the lengths of all formulae. □

2.1.2 Proof of Barr's Theorem

Now we can prove Theorem 2.1: Assume $\Delta \vdash_{tc} A$ with $A = \forall \vec{x}.B \to C$ where Δ, B, C are geometric and C does not contain a free variable occurring bound in Δ, B, let t be the (classical) proof. Then $s := t\vec{x}u^B$ proves $\Delta, B \vdash_{tc} C$, and $\lambda_{\vec{x}}\lambda_B s$ proves $\Delta \vdash_{tc} A$ again, hence, $\Delta \vdash_{tc} A \Leftrightarrow \Delta, B \vdash_{tc} C$, so it suffices to show that $\Delta, B \vdash_i C$. By Theorem 2.6 and Lemma 2.7 we have $\Delta, B \vdash_{tc} C \Rightarrow \Delta_C, B_c \vdash_i C_C \Rightarrow \Delta, B \vdash_i C_C$, and with Lemma 2.3 we get $\Delta, B \vdash_i (C \to C) \to C$ and therefore $\Delta, B \vdash_i C$. As we use a composition of linear algorithms, the algorithm generating this proof is linear relative to $\sum_{X \in \Delta} len(X)$.

2.2 The Proof by Negri, Plato and Strahm

We will look at a modification of the proof given in (5), which is based on a suggestion by Thomas Strahm. It uses the sequent calculi **G3c** and **G3im**, and where the original proof adds additional rules to include the axioms, we will prove that the necessary properties, namely the admissibility of their structural rules and the cut rule, are still valid if one allows additional, "geometric" axioms.

Strahm suggested (for another calculus, however) that one could just add additional axioms to the logical axioms and then prove the admissibility of cut elimination for this new calculus. However, for **G3im** and **G3c**, it turns out that partial cut elimination does not hold with general additional axioms. Still, the full formalism in (5) is not necessary to prove Barr's theorem. We give a similar formalism, using our geometric axioms, and prove the necessary structural properties for a partial cut elimination theorem.

Most of these structural theorems are just straightforward generalizations of the corresponding structural theorems for the calculi without additional axioms.

2.2.1 G3im and G3c

The rules of the calculi **G3im** and **G3c** can be separated into rules for the connectives, structural rules and the cut rule. Their rules handle multiple formulae in the antecedents, which are interpreted conjunctively, and multiple formulae in the succedents, which are interpreted disjunctively.

We write $\vdash_G$ if the rules apply for both **G3im** and **G3c**, $\vdash_{Gc}$ for **G3c** and $\vdash_{Gi}$ for **G3im**. $\vdash_{Gc}$ corresponds to $\vdash_{tc}$ in the sense that if a sequent $\Gamma \Rightarrow A$ can be derived by **G3c**, then $\Gamma \vdash_{tc} A$. $\vdash_{Gi}$ corresponds to $\vdash_i$ in the sense that if a sequent $\Gamma \Rightarrow A$ can be derived by **G3im**, then $\Gamma \vdash_i A$. In 3.2, we prove something similar for Calculus n.

The general rules for the connectives are

$$\frac{P \text{ atomic}}{P, \Gamma \Rightarrow \Delta, P}\ \text{Axiom} \qquad \frac{}{\bot, \Gamma \Rightarrow \Delta}\ \text{L}\bot$$

$$\frac{A, B, \Gamma \Rightarrow \Delta}{A \wedge B, \Gamma \Rightarrow \Delta}\ \text{L}\wedge \qquad \frac{\Gamma \Rightarrow \Delta, A \qquad \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \wedge B}\ \text{R}\wedge$$

$$\frac{A, \Gamma \Rightarrow \Delta \qquad B, \Gamma \Rightarrow \Delta}{A \vee B, \Gamma \Rightarrow \Delta}\ \text{L}\vee \qquad \frac{\Gamma \Rightarrow \Delta, A, B}{\Gamma \Rightarrow \Delta, A \vee B}\ \text{R}\vee$$

$$\frac{A[x := t], \forall_x A, \Gamma \Rightarrow \Delta}{\forall_x A, \Gamma \Rightarrow \Delta}\ \text{L}\forall$$

$$\frac{A[x := y], \Gamma \Rightarrow \Delta}{\exists_x A, \Gamma \Rightarrow \Delta}\ \text{L}\exists \qquad \frac{\Gamma \Rightarrow \Delta, \exists_x A, A[x := t]}{\Gamma \Rightarrow \Delta, \exists_x A}\ \text{R}\exists$$

Additionally, **G3c** has the rules

$$\frac{\Gamma \Rightarrow \Delta, A \qquad B, \Gamma \Rightarrow \Delta}{A \to B, \Gamma \Rightarrow \Delta} \, L{\to_c} \qquad \frac{A, \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \to B} \, R{\to_c}$$

$$\frac{\Gamma \Rightarrow \Delta, A[x := y]}{\Gamma \Rightarrow \Delta, \forall_x A} \, R\forall_c$$

while **G3im** has corresponding rules

$$\frac{A \to B, \Gamma \Rightarrow A \qquad B, \Gamma \Rightarrow \Delta}{A \to B, \Gamma \Rightarrow \Delta} \, L{\to_i} \qquad \frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow \Delta, A \to B} \, R{\to_i}$$

$$\frac{\Gamma \Rightarrow A[x := y]}{\Gamma \Rightarrow \Delta, \forall_x A} \, R\forall_i$$

There are variable conditions on y for $L\exists$ and $R\forall_{c/i}$, namely that y must not occur freely in the conclusions of the rules.

Structural rules for left and right weakening and contraction are

$$\frac{\Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta} \, LW \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A} \, RW$$

$$\frac{A, A, \Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta} \, LC \qquad \frac{\Gamma \Rightarrow \Delta, A, A}{\Gamma \Rightarrow \Delta, A} \, RC$$

The cut rule is

$$\frac{\Gamma \Rightarrow \Delta, A \qquad A, \Gamma' \Rightarrow \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'} \, \mathrm{Cut}$$

and we call A the *cut formula* of that cut.

Definition As we call the rules $L\bot$ and Axiom *logical axioms* we call additional axioms *non-logical axioms*. A *main formula* of a rule is an explicit formula in the conclusion of that rule. The *context preserving cut rule* $\mathrm{Cut}_{\mathrm{CS}}$ is defined by

$$\frac{\Gamma \Rightarrow \Delta, A \qquad A, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta} \, \mathrm{Cut}_{\mathrm{CS}}$$

We also call A the *cut formula* of that cut.

Lemma 2.8. $\vdash_G \Gamma, A \Rightarrow A, \Delta$ *can be derived for every formula* A.

Proof By structural induction on A. For atomic A and $\perp$ we have the axioms. For $A = B \wedge C$, we have

$$\dfrac{\dfrac{A, B, \Gamma \Rightarrow A, \Delta \qquad A, B, \Gamma \Rightarrow B, \Delta}{A, B, \Gamma \Rightarrow A \wedge B, \Delta}}{A \wedge B, \Gamma \Rightarrow A \wedge B, \Delta}$$

Similarly for $A = B \vee C$. For $A = \forall_x B$, let c not occur freely in Γ, Δ, then we have

$$\dfrac{\dfrac{A[x := c], \Gamma \Rightarrow \Delta, A[x := c]}{\forall_x A, \Gamma \Rightarrow \Delta, A[x := c]}}{\forall_x A, \Gamma \Rightarrow \Delta, \forall_x A}$$

Similarly for $A = \exists_x B$. For $A = B \rightarrow C$ we may use

$$\dfrac{\dfrac{A \rightarrow B, \Gamma, A \Rightarrow A \qquad B, \Gamma \Rightarrow B}{A \rightarrow B, \Gamma, A \Rightarrow B}}{A \rightarrow B, \Gamma \Rightarrow \Delta, A \rightarrow B}$$

in **G3im**, similarly for **G3c**. $\square$

Lemma 2.9. *In* **G3c** *with arbitrary non-logical axioms, the cut rule can be replaced by the context preserving cut rule.*

Proof Let a proof tree be given that contains at least one application of the cut rule, and α be a top-most application of it. Then α has the form

$$\dfrac{\begin{matrix}\mathcal{D}\\ \Gamma \Rightarrow \Delta, A\end{matrix} \qquad \begin{matrix}\mathcal{E}\\ \Gamma', A \Rightarrow \Delta'\end{matrix}}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,\text{Cut}$$

Then we can replace α by the following derivation that uses the context preserving cut rule:

$$\dfrac{\dfrac{\dfrac{\begin{matrix}\mathcal{D}\\ \Gamma \Rightarrow \Delta, A\end{matrix}}{\Gamma \Rightarrow \Delta', \Delta, A}\,\text{RW}}{\Gamma', \Gamma \Rightarrow \Delta', \Delta, A}\,\text{LW} \qquad \dfrac{\dfrac{\begin{matrix}\mathcal{E}\\ \Gamma', A \Rightarrow \Delta'\end{matrix}}{\Gamma', A \Rightarrow \Delta', \Delta}\,\text{RW}}{\Gamma', \Gamma, A \Rightarrow \Delta', \Delta}\,\text{LW}}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,\text{Cut}_{\text{cs}}$$

$\square$

Definition A *geometric axiom* is a sequent $P_1, \ldots, P_p \Rightarrow \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m$, where the P_i are atomic, and the M_i are conjunctions of atomic formulae, and $\vec{y}$ is non-empty.

It should be immediately clear that this definition is related to geometric theories. Before we can actually use it, we need to prove some properties of **G3c** that are preserved when allowing geometric axioms.

Denote by **G3cT** the calculus **G3c** with additional geometric axioms (silently ignoring the fact that it depends on the set of additional geometric axioms given, which will always be clear or irrelevant), and by $\vdash_{\mathsf{GcT}}$ the entailment relation of this calculus.

The following is a trivial generalization of the substitution lemma given in (12):

Lemma 2.10. *Assume* $\vdash_{\mathsf{GcT}}{}^{n}\Gamma \Rightarrow \Delta$, *let* x *be free in* Γ, Δ, *such that it can be substituted by* t *without variable collisions, and such that* t *does not contain any free variable used in a rule application of* $\mathsf{L}\exists$ *or* $\mathsf{R}\forall_{i/c}$ *in the proof. Let the geometric axioms be closed under substitution. Then* $\vdash_{\mathsf{GcT}}{}^{n}\Gamma[x := t] \Rightarrow \Delta[x := t]$.

Proof We prove this by induction on n. For $n = 0$, for the logical axioms it is trivial, while for the geometric axioms we just assume it. The induction steps for the rules involving the connectives are trivial as well, as these do not change variable assignments.

For L$\forall$, we have a proof of the form

$$\frac{\vdash_{\mathsf{GcT}}{}^{n}A[y := s], \forall_y A, \Gamma \Rightarrow \Delta}{\vdash_{\mathsf{GcT}}{}^{n+1}\forall_y A, \Gamma \Rightarrow \Delta}$$

For $x \neq y$ the induction step is trivial, so assume that $x = y$, that is, we want to derive $\vdash_{\mathsf{GcT}}{}^{n+1}(\forall_y A)[y := t], \Gamma[y := t] \Rightarrow \Delta[y := t]$ which is $\vdash_{\mathsf{GcT}}{}^{n+1}(\forall_y A), \Gamma[y := t] \Rightarrow \Delta[y := t]$. By induction we can derive $\vdash_{\mathsf{GcT}}{}^{n}A[y := s][y := t], \forall_y A, \Gamma[y := t] \Rightarrow \Delta[y := t]$, and therefore we get

$$\frac{\vdash_{\mathsf{GcT}}{}^{n}A[y := s][y := t], \forall_y A, \Gamma[y := t] \Rightarrow \Delta[y := t]}{\vdash_{\mathsf{GcT}}{}^{n+1}\forall_y A, \Gamma[y := t] \Rightarrow \Delta[y := t]}$$

Similarly for R$\exists$.

For L$\exists$, we have

$$\frac{\vdash_{GcT}{}^{n} A[z := y], \Gamma \Rightarrow \Delta}{\vdash_{GcT}{}^{n+1} \exists_z A, \Gamma \Rightarrow \Delta}$$

and by the variable condition we know that y does not occur freely in Γ, Δ, therefore we may assume $x \neq y$ without loss of generality. By assumption we also know that t does not contain y, and therefore, $\vdash_{GcT}{}^{n} A[z := y][x := t], \Gamma[x := t] \Rightarrow \Delta[x := t]$, which is derivable by induction hypothesis, satisfies the variable condition, and we can derive

$$\frac{\vdash_{G}{}^{n} A[z := y][x := t], \Gamma[x := t] \Rightarrow \Delta[x := t]}{\vdash_{G}{}^{n+1} (\exists_z A)[x := t], \Gamma[x := t] \Rightarrow \Delta[x := t]}$$

Similarly for $R\forall_{i/c}$. □

Lemma 2.11. *In* **G3cT**, *the following properties hold, which can be used to decompose formulae:*

1. *If* $\vdash_{GcT}{}^{n} A \wedge B, \Gamma \Rightarrow \Delta$, *then* $\vdash_{GcT}{}^{n} A, B, \Gamma \Rightarrow \Delta$.
2. *If* $\vdash_{GcT}{}^{n} \Gamma \Rightarrow \Delta, A \vee B$, *then* $\vdash_{GcT}{}^{n} \Gamma \Rightarrow \Delta, A, B$.
3. *If* $\vdash_{GcT}{}^{n} A_0 \vee A_1, \Gamma \Rightarrow \Delta$, *then* $\vdash_{GcT}{}^{n} A_i, \Gamma \Rightarrow \Delta$ *for* $i \in \{0, 1\}$.
4. *If* $\vdash_{GcT}{}^{n} \Gamma \Rightarrow \Delta, A_0 \wedge A_1$, *then* $\vdash_{GcT}{}^{n} \Gamma \Rightarrow \Delta, A_i$ *for* $i \in \{0, 1\}$.
5. *If* $\vdash_{GcT}{}^{n} \Gamma \Rightarrow A \rightarrow B, \Delta$, *then* $\vdash_{GcT}{}^{n} \Gamma, A \Rightarrow B, \Delta$.
6. *If* $\vdash_{GcT}{}^{n} \Gamma, A \rightarrow B \Rightarrow \Delta$, *then* $\vdash_{GcT}{}^{n} \Gamma \Rightarrow \Delta, A$ *and* $\vdash_{GcT}{}^{n} \Gamma, B \Rightarrow \Delta$.
7. *If* $\vdash_{GcT}{}^{n} \Gamma \Rightarrow \Delta, \forall_x A$, *then* $\vdash_{GcT}{}^{n} \Gamma \Rightarrow \Delta, A[x := y]$ *for all* y *such that* $y \notin FV(\Gamma, \Delta, A)$.
8. *If* $\vdash_{GcT}{}^{n} \exists_x A, \Gamma \Rightarrow \Delta$, *then* $\vdash_{GcT}{}^{n} A[x := y], \Gamma \Rightarrow \Delta$ *for* $y \notin FV(\Gamma, \Delta, A)$.

Proof By Lemma 2.9, we may assume that only context preserving cuts are used. This is a generalization of the inversion lemma given in (12), according to which the Theorem can be proved by induction on n, and we only add a special case for $n = 0$: If we have a logical axiom, then we know that the decomposed formula must be in the context of that axiom, since logical axioms only have $\perp$ or atomic formulae outside of their context. Then trivially, the given conclusion in each case is an axiom too, as we only have to change the context of that logical axiom. For geometric axioms, we know that 1,3,6,8 hold, as we only have atomic formulae on the left side of these axioms. As everything must be inside an existential quantifier on the right side, 2,5,7 and 4 hold. Therefore, for $n = 0$, we are done. □

Lemma 2.12. *Assume that all geometric axioms are closed under contraction and substitution. Then*

1. *If* $\vdash_{GcT}{}^{n} \Gamma, A, A \Rightarrow \Delta$, *then* $\vdash_{GcT}{}^{n} \Gamma, A \Rightarrow \Delta$.

2. *If* $\vdash_{GcT}{}^{n}\Gamma \Rightarrow \Delta, A, A$, *then* $\vdash_{GcT}{}^{n}\Gamma \Rightarrow \Delta, A$.

From this Lemma it directly follows that the contraction rules LC and RC are admissible in **G3cT**, except for chains of contractions directly below the geometric axioms. This result, for **G3c** without geometric axioms, is also proved in (12).

Proof We may again assume by 2.9 that only context preserving cuts are used. A proof of the admissibility of contraction is given in (12), and we only add another induction base, namely the geometric axioms, for which we assume the lemma and closedness under substitution, which is needed to apply our version of the substitution lemma, Lemma 2.10. □

We can similarly prove that the weakening rule is admissible as well, except for chains of weakenings directly below (possibly contracted) geometric axioms.

Lemma 2.13. *Let the set of geometric axioms be closed under the application of weakening rules and substitutions. Then every* **G3cT***-proof containing weakening rules can be converted into a proof in which no weakening rules occur.*

Proof Weakenings just extend the context, and we can trivially shift them to the top of the proof in all rules that do not have variable conditions, and as all axioms are closed under weakening, we will gain a shorter proof without the instances of weakening.

For the rule L$\exists$, we would have the situation

$$\dfrac{\dfrac{A[x := y], \Gamma \Rightarrow \Delta}{\exists_x A, \Gamma \Rightarrow \Delta}}{\exists_x A, \Gamma \Rightarrow \Delta, B}$$

If y does not occur freely in B, we can just shift that weakening through the L$\exists$ application. If not, let z be a variable not occurring yet. Then by Lemma 2.10 we can derive $A[x := z], \Gamma \vdash_{GcT} \Delta$ where z is not free in B, and therefore we can convert the whole derivation into

$$\dfrac{\dfrac{A[x := z], \Gamma \Rightarrow \Delta}{A[x := z], \Gamma \Rightarrow \Delta, B}}{\exists_x A, \Gamma \Rightarrow \Delta}$$

and therefore shift this application of weakening up. Similarly for R$\forall_c$. □

Instead of requiring the geometric axioms to be closed under the application of structural rules, we could equivalently prove that structural rules can be omitted except for chains of applications of structural rules directly below them, which is the version we need in what follows.

2.2.2 Partial Cut Elimination

Theorem 2.14 (Partial Cut Elimination). *Let the geometric axioms be closed under substitution. Then we can omit the structural rules, except for chains of structural rules directly below geometric axioms, and the cut rule for cut formulae not being part of a geometric axiom, from* **G3cT**.

Proof By the above lemmata we can assume that the proof does not contain structural rules, except for chains directly below the geometric axioms, and no instance of a cut rule that is not context preserving is present. We define the *adapted height* of a derivation as the height of this derivation, not counting chains of structural rules below geometric axioms. We define the *level* of a cut as the sum of the adapted heights of the proofs of its premises. The *range* of a cut is the range of its cut formula. We show that every cut of which the cut formula does not occur in a geometric axiom can be replaced by a cut with lower range or level. Hence, we will be able to recursively replace all cuts until only cuts with level 0 are left, and then we show that these can be removed, too.

So let a top most instance of a cut be given.

$$\dfrac{\begin{array}{c}\mathcal{D}\\ \Gamma \Rightarrow \Delta, A\end{array} \qquad \begin{array}{c}\mathcal{E}\\ A, \Gamma \Rightarrow \Delta\end{array}}{\Gamma \Rightarrow \Delta}\,\mathrm{Cut_{CS}}$$

If $\Gamma \Rightarrow \Delta, A$ is a geometric axiom, or resulted from a contraction rule applied to one, or from a weakening rule of which A is not the main formula, we are done, as the cut has a cut formula occurring in a geometric axiom.

If $\Gamma \Rightarrow \Delta, A$ resulted from a weakening rule, this must be the last rule of a chain of structural rules applied to a geometric axiom, and we have the situation

$$\dfrac{\begin{array}{c}\mathcal{D}\\ \Gamma \Rightarrow \Delta', A, A\end{array} \qquad \begin{array}{c}\mathcal{E}\\ A, \Gamma \Rightarrow \Delta', A\end{array}}{\Gamma \Rightarrow \Delta', A}\,\mathrm{Cut_{CS}}$$

which can be replaced by

$$\dfrac{\begin{array}{c}\mathcal{D}\\ \Gamma \Rightarrow \Delta', A, A\end{array}}{\Gamma \Rightarrow \Delta', A}\,\mathrm{RC}$$

which is legitimate since it just extends that chain of structural rules.

Assuming $\Gamma \Rightarrow \Delta, A$ is a logical axiom following from the rule Axiom, where A is not main formula, the derivation looks like

$$\dfrac{\text{Axiom}\ \dfrac{}{P, \Gamma' \Rightarrow P, \Delta', A} \qquad \begin{matrix}\mathcal{E} \\ A, P, \Gamma' \Rightarrow P, \Delta'\end{matrix}}{P, \Gamma' \Rightarrow \Delta', P}$$

and we can get $P, \Gamma' \Rightarrow \Delta', P$ directly. If A is main formula of Axiom the derivation looks like

$$\dfrac{\dfrac{}{A, \Gamma' \Rightarrow \Delta', A}\ \text{Axiom} \qquad \begin{matrix}\mathcal{E} \\ A, A, \Gamma' \Rightarrow \Delta'\end{matrix}}{A, \Gamma' \Rightarrow \Delta'}$$

But we can as well apply LC on $A, A, \Gamma' \Rightarrow \Delta'$ to get $A, \Gamma' \Rightarrow \Delta'$, and then remove this application of a structural rule.

Assume $\Gamma \Rightarrow \Delta, A$ follows by L$\perp$. If $A \neq \perp$ then $\perp \in \Gamma$, and hence the result can be obtained from L$\perp$ directly. If $A = \perp$, $\Gamma \Rightarrow \Delta, A$ can also be obtained by Axiom, and therefore follows from the former case.

Similarly for $A, \Gamma \Rightarrow \Delta$ following from Axiom or L$\perp$ where A is not the main formula, and $A, \Gamma \Rightarrow \Delta$ following from Axiom where A is the main formula.

The remaining case for logical axioms is $A, \Gamma \Rightarrow \Delta$ following from L$\perp$, and $A = \perp$. If $\perp$ also is the main formula in $\mathcal{D}$, it can only be of the form Form $\Gamma, \perp \Rightarrow \Delta, \perp$, but then $A, \Gamma \Rightarrow \Delta$ has the form $\perp, \Gamma \Rightarrow \Delta', \perp$, which is an axiom. So let $\perp$ not be main formula in $\mathcal{D}$. If $\mathcal{D}$ has two premises, then the cut application has the form

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0 \\ \Gamma' \Rightarrow \Delta', \perp\end{matrix} \qquad \begin{matrix}\mathcal{D}_1 \\ \Gamma'' \Rightarrow \Delta'', \perp\end{matrix}}{\Gamma \Rightarrow \Delta, \perp}\ R \qquad \perp, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}$$

which can be transformed into

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0 \\ \Gamma' \Rightarrow \Delta', \perp\end{matrix} \qquad \dfrac{}{\perp, \Gamma' \Rightarrow \Delta'}}{\Gamma' \Rightarrow \Delta'} \qquad \dfrac{\begin{matrix}\mathcal{D}_1 \\ \Gamma'' \Rightarrow \Delta'', \perp\end{matrix} \qquad \dfrac{}{\perp, \Gamma'' \Rightarrow \Delta''}}{\Gamma'' \Rightarrow \Delta''}}{\Gamma \Rightarrow \Delta}\ R$$

We can proceed similarly if R has only one premise.

Since we handled with all cases where at least one premise of the cut was followed by an axiom, let us now assume both premises did not follow from axioms.

Assume $\Gamma \Rightarrow \Delta, A$ follows from a rule R in which A is not the main formula. For R with two premises the derivation then has the form

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0 \\ \Gamma' \Rightarrow \Delta', A\end{matrix} \qquad \begin{matrix}\mathcal{D}_1 \\ \Gamma'' \Rightarrow \Delta'', A\end{matrix}}{\Gamma \Rightarrow \Delta, A}\,R \qquad \begin{matrix}\mathcal{D}_2 \\ A, \Gamma \Rightarrow \Delta\end{matrix}}{\Gamma \Rightarrow \Delta}$$

which can be transformed into

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0 \\ \Gamma' \Rightarrow \Delta', A\end{matrix}}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta', A} \qquad \dfrac{\begin{matrix}\mathcal{D}_2 \\ A, \Gamma \Rightarrow \Delta\end{matrix}}{A, \Gamma, \Gamma' \Rightarrow \Delta, \Delta'}}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'(*)}$$

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_1 \\ \Gamma'' \Rightarrow \Delta'', A\end{matrix}}{\Gamma, \Gamma'' \Rightarrow \Delta, \Delta'', A} \qquad \dfrac{\begin{matrix}\mathcal{D}_2 \\ A, \Gamma \Rightarrow \Delta\end{matrix}}{A, \Gamma, \Gamma'' \Rightarrow \Delta, \Delta''}}{\Gamma, \Gamma'' \Rightarrow \Delta, \Delta''(**)}$$

$$\dfrac{(*) \qquad (**)}{\Gamma \Rightarrow \Delta}\,R$$

The structural rules can be replaced again by the above lemma. The levels of the new cuts are therefore strictly smaller than the level of the cut before. Similarly for the case that $A, \Gamma \Rightarrow \Delta$ follows from a rule in which A is not the main formula. Similarly for rules with one premise. The remaining case is that A is main formula in both premises. In that case, we must look at the structure of A. For $A = \bot$, it must be an axiom, since no rule except an axiom can have $\bot$ as a main formula. Assume $A = X \wedge Y$, then we have

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0 \\ \Gamma \Rightarrow \Delta, X\end{matrix} \qquad \begin{matrix}\mathcal{D}_1 \\ \Gamma \Rightarrow \Delta, Y\end{matrix}}{\Gamma \Rightarrow \Delta, X \wedge Y} \qquad \dfrac{\begin{matrix}\mathcal{D}_2 \\ X, Y, \Gamma \Rightarrow \Delta\end{matrix}}{X \wedge Y, \Gamma \Rightarrow \Delta}}{\Gamma \Rightarrow \Delta}$$

It can be transformed into

$$
\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \Rightarrow \Delta, Y\end{array} \qquad \dfrac{\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma \Rightarrow \Delta, X\end{array}}{Y, \Gamma \Rightarrow \Delta, X} \qquad \begin{array}{c}\mathcal{D}_2\\ X, Y, \Gamma \Rightarrow \Delta\end{array}}{Y, \Gamma \Rightarrow \Delta}}{\Gamma \Rightarrow \Delta}
$$

Again the applications of Struct can be removed. The level of neither new cut is greater than the level of the cut we had before, and the ranges of both new cuts are strictly smaller than the range of the cut we had before. Similarly for $A = X \vee Y$. For $A = X \to Y$ we have

$$
\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma, X \Rightarrow Y, \Delta\end{array}}{\Gamma \Rightarrow X \to Y, \Delta} \qquad \dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \Rightarrow X, \Delta\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ \Gamma, Y \Rightarrow \Delta\end{array}}{\Gamma, X \to Y \Rightarrow \Delta}}{\Gamma \Rightarrow \Delta}
$$

It can be transformed into

$$
\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \Rightarrow X, \Delta\end{array} \qquad \dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma, X \Rightarrow Y, \Delta\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ \Gamma, X, Y \Rightarrow \Delta\end{array}}{\Gamma, X \Rightarrow \Delta}}{\Gamma \Rightarrow \Delta}
$$

Again we can remove the structural rules, the levels of the new cuts are not greater, and the ranges of the new cuts are smaller. For $A = \forall_x D$ the derivation has the form

$$
\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma \Rightarrow \Delta, D[x := y]\end{array}}{\Gamma \Rightarrow \Delta, \forall_x D} \qquad \dfrac{\begin{array}{c}\mathcal{D}_1\\ \forall_x D, D[x := t], \Gamma \Rightarrow \Delta\end{array}}{\forall_x D, \Gamma \Rightarrow \Delta}}{\Gamma \Rightarrow \Delta}
$$

Firstly, we can replace y in $\mathcal{D}_0$ by a term t, we call that new derivation $\mathcal{D}_0'$.

$$
\dfrac{\begin{array}{c}\mathcal{D}_0'\\ \Gamma \Rightarrow \Delta, D[x := t]\end{array} \qquad \dfrac{\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma \Rightarrow \Delta, D[x := y]\end{array}}{D[x := t], \Gamma \Rightarrow \Delta, D[x := y]}}{D[x := t], \Gamma \Rightarrow \Delta, \forall_x D} \qquad \begin{array}{c}\mathcal{D}_1\\ \forall_x D, D[x := t], \Gamma \Rightarrow \Delta\end{array}}{D[x := t], \Gamma \Rightarrow \Delta}}{\Gamma \Rightarrow \Delta}
$$

Again the structural rules can be removed, the level of the resulting cut stays the same since the height of the left premise is decreased by one while the height of the right premise is increased by one, and the range of the resulting cut is smaller. The cut on the right side has a smaller level. Similarly for $A = \exists_x D$. □

Corollary 2.15. *In* **G3c** *(without additional non-logical axioms) we can omit the cut rule and the structural rules.*

2.2.3 The Proof

Lemma 2.16. *A sequent of the form* $\Rightarrow \forall_{\vec{x}}.P_0 \rightarrow \exists_{\vec{y}}.P_1 \vee \ldots \vee P_n$ *where* $\vec{y}$ *is not empty, and the* P_i *are conjunctions of atomic formulae is derivable from geometric axioms without using cuts, using the rules of* **G3im** *(or* **G3c***).*

Proof Firstly, we can derive $\Rightarrow \forall_{\vec{x}}.P_0 \rightarrow \exists_{\vec{y}}.P_1 \vee \ldots \vee P_n$ from $\Rightarrow P_0 \rightarrow \exists_{\vec{y}}.P_1 \vee \ldots \vee P_n$ by applying $R\forall_i$ or similarly, $R\forall_c$. There is no danger of violating a variable condition, as only one formula is present. Then by $R\rightarrow_{c/i}$, we can derive this sequent from $P_0 \Rightarrow \exists_{\vec{y}}.P_1 \vee \ldots \vee P_n$. Then $P_0 = P_{00} \wedge \ldots \wedge P_{0p_0}$ can be derived from $P_{00}, \ldots, P_{0p_0} \Rightarrow \exists_{\vec{y}}.P_1 \vee \ldots \vee P_n$ by multiple applications of $L\wedge$, and this is a geometric axiom. □

Now assume we have a proof for $\vdash_{Gc} \Gamma \Rightarrow G$ where Γ is a geometric theory and G is a geometric implication, and all variables occurring freely in some element of Γ, G are parameters. By 2.1, we may assume that all elements of Γ have the form

$$\forall_{\vec{x}}.P_0 \rightarrow \exists_{\vec{y}}.P_1 \vee \ldots \vee P_n$$

and these can be derived from geometric axioms by Lemma 2.16. If $G = \forall_{\vec{x}}.A \rightarrow B$ with geometric formulae A and B, from $\Rightarrow G$ we may derive the sequent $A \Rightarrow B$ in a way similar to the destructuring of the G_i shown below, using the cut rule.

The resulting proof is a proof in **G3cT**, and therefore, partial cut elimination and admissibility of weakening rules except for chains directly below the geometric axioms hold. But in such a proof, the formulae in the premises of rules are subformulae of the formulae in the conclusion or in the geometric axioms, and thus, as the final conclusion of this proof does not contain implications and universal quantifications, none of the formulae above may contain them. But this means that no rule for $\forall$ and $\rightarrow$ can be applied, and hence, the proof only uses rules that **G3c** and **G3im** have in common. From the conclusion, by $R\rightarrow_c$ and $R\forall_i$ we can derive $\Rightarrow G$ again. As all the G_i are closed except for parameters, we may extend all contexts in the antecedents by Γ without the danger of violating a variable condition, gaining a proof of $\Gamma \Rightarrow G$, but losing the property that the axioms are all geometric axioms. However, $\vdash_{Gi} \Gamma \Rightarrow G_i$ is derivable by Lemma 2.8.

Let $G_i = \forall_{\vec{x}}.M_0 \to \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m$ where the M_i are conjunctions of atomic formulae, and $M_0 = A_0 \wedge \ldots \wedge A_a$. From $\vdash_{Gi} \Gamma \Rightarrow G_i$ we can then firstly derive

$$\dfrac{\Gamma \Rightarrow G_i \qquad \dfrac{M_0 \to \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m \Rightarrow M_0 \to \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m}{\forall_{\vec{x}}.M_0 \to \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m \Rightarrow M_0 \to \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m}}{\Gamma \Rightarrow M_0 \to \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m}$$

in **G3im**. We use the cut rule for **G3im** (which is also admissible according to (5)). By

$$\dfrac{M_0, M_0 \to \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m, \Gamma \Rightarrow M_0 \qquad \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m, \Gamma \Rightarrow \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m}{M_0, M_0 \to \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m, \Gamma \Rightarrow \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m}$$

we can decompose the implication. Using

$$\dfrac{\dfrac{A_0, \ldots, A_a \Rightarrow A_0 \qquad \cdots \qquad A_0, \ldots, A_a \Rightarrow A_a}{A_0, \ldots, A_a \Rightarrow A_0 \wedge \ldots \wedge A_a} \qquad A_0 \wedge \ldots \wedge A_a, \Gamma \Rightarrow \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m}{A_0, \ldots, A_a, \Gamma \Rightarrow \exists_{\vec{y}}.M_1 \vee \ldots \vee M_m}$$

we may also decompose the conjunction. Using this, we can derive the geometric axioms with the Γ-extended contexts intuitionistically.

Applying this to all geometric axioms we get an intuitionistic proof for $\vdash_{Gi} \Gamma \Rightarrow G$ (that may not be cut free, though). This proves Theorem 2.1.

3 Orevkov's Theorem

In (7), Orevkov gives a complete classification of singular sequents $\Gamma \Rightarrow A$ for which classical derivability implies intuitionistic derivability in terms of forbidden positive or negative occurrences of connectives. An equivalent result with a slightly different notation is given in (4) by Nadathur.

Glivenko sequent classes are sets of sequents that are classically derivable iff they are intuitionistically derivable. In (7), a full classification of such Glivenko sequent classes with respect to positivity and negativity of connectives is given.

Definition A σ-class is a subset of restrictions $\{\wedge^+, \wedge^-, \vee^+, \vee^-, \to^+, \to^-, \neg^+, \neg^-, \forall^+, \forall^-, \exists^+, \exists^-\}$. A sequent belongs to a σ-class if it does **not** contain any of its $+$-signed connectives positively, and **not** contain any of its $-$-signed connectives negatively. A σ-class is called *complete Glivenko class* if every classically derivable singular sequent belonging to this class is also intuitionistically derivable, while for every class containing fewer restrictions, this is not the case.

Theorem 3.1 (Orevkov). *The following list is a complete list of σ-classes which are complete Glivenko classes:*

$$\{\to^+, \neg^+, \forall^+\}, \{\to^+, \neg^+, \vee^-\}, \{\to^+, \neg^+, \forall^-\}, \{\to^-, \neg^-, \vee^+, \exists^+\}, \{\to^-, \neg^-, \vee^+, \to^+, \forall^+\},$$
$$\{\to^-, \neg^-, \vee^+, \to^+, \vee^-\}, \{\to^-, \neg^-, \vee^+, \to^+, \forall^-\}$$

In the following chapters, we will concentrate on $\{\to^+, \neg^+, \vee^-\}$.

3.1 Nadathur's Proof

In (4), the author allows the symbols $\top$, $\bot$, $\wedge$, $\vee$, $\to$, $\exists$, $\forall$, and defines $\neg A := A \to \bot$. He uses the following calculus, which we will refer to as *Calculus* n. We will use $\vdash_n$ as its derivability relation.

There are axioms

$$\frac{}{\Gamma \Rightarrow \top} \qquad\qquad \frac{A \text{ atomic}}{A, \Gamma \Rightarrow \Delta, A}$$

There are contraction rules

$$\frac{\Gamma, B, B \Rightarrow \Delta}{\Gamma, B \Rightarrow \Delta}\text{contr-L} \qquad\qquad \frac{\Gamma \Rightarrow \Delta, B, B}{\Gamma \Rightarrow \Delta, B}\text{contr-R}$$

There are operational rules

$$\frac{\Gamma \Rightarrow \Delta, \bot}{\Gamma \Rightarrow \Delta, D}\bot\text{-R}$$

$$\frac{B_i, \Gamma \Rightarrow \Delta \qquad i \in \{1,2\}}{B_1 \wedge B_2, \Gamma \Rightarrow \Delta}\wedge\text{-L} \qquad\qquad \frac{\Gamma \Rightarrow \Delta, B \qquad \Gamma \Rightarrow \Delta, D}{\Gamma \Rightarrow \Delta, B \wedge D}\wedge\text{-R}$$

$$\frac{B, \Gamma \Rightarrow \Delta \qquad D, \Gamma \Rightarrow \Delta}{B \vee D, \Gamma \Rightarrow \Delta}\vee\text{-L} \qquad\qquad \frac{\Gamma \Rightarrow \Delta, B_i \qquad i \in \{1,2\}}{\Gamma \Rightarrow \Delta, B_1 \vee B_2}\vee\text{-R}$$

$$\frac{\Gamma \Rightarrow \Delta, B \qquad D, \Gamma \Rightarrow \Theta}{B \to D, \Gamma \Rightarrow \Delta, \Theta}\to\text{-L} \qquad\qquad \frac{B, \Gamma \Rightarrow \Delta, D}{\Gamma \Rightarrow \Delta, B \to D}\to\text{-R}$$

In the following operational rules, t denotes a term, and c denotes a constant that does not occur in the formulae of the conclusion.

$$\frac{B[x := t], \Gamma \Rightarrow \Delta}{\forall_x B, \Gamma \Rightarrow \Delta}\ \forall\text{-L} \qquad \frac{\Gamma \Rightarrow \Delta, B[x := c]}{\Gamma \Rightarrow \Delta, \forall_x B}\ \forall\text{-R}$$

$$\frac{B[x := c], \Gamma \Rightarrow \Delta}{\exists_x B, \Gamma \Rightarrow \Delta}\ \exists\text{-L} \qquad \frac{\Gamma \Rightarrow \Delta, B[x := t]}{\Gamma \Rightarrow \Delta, \exists_x B}\ \exists\text{-R}$$

Definition A proof in Calculus n shall be called a C-*proof*. A proof in the Calculus n where every occurring succedent contains exactly one formula shall be called an I-*proof*. C-proofs are classical proofs, in the sense that there is a C-proof for $\vdash_n \Gamma \Rightarrow A$ iff we have $\Gamma \vdash_{tc} A$. I-proofs are intuitionistic proofs, in the sense that iff there is an I-proof for $\vdash_n \Gamma \Rightarrow A$ then we have $\Gamma \vdash_i A$. (This will be proved in 3.2).

Notice that we did not require the sequents to be singular for I-proofs.

Firstly, a substitution lemma similar to Lemma 2.10 holds, from which we can conclude the admissibility of a weakening rule, as we did for **G3c**:

Lemma 3.2. *Assume* $\vdash_n{}^n \Gamma \Rightarrow \Delta$*, let* x *be free in* Γ, Δ*, such that it can be substituted by* t *without variable collisions, and such that* t *does not contain any free variable used in a rule application of* $\exists$*-L or* $\forall$*-R in the proof. Then* $\vdash_n{}^n \Gamma[x := t] \Rightarrow \Delta[x := t]$. *If the first proof is an* I*-proof, then the resulting proof is an* I*-proof as well.*

Lemma 3.3. *If* $\vdash_n{}^n \Gamma \Rightarrow \Delta$*, then* $\vdash_n{}^n A, \Gamma \Rightarrow \Delta$*. If the first proof is an* I*-proof, then the resulting proof is an* I*-proof as well.*

Proof By induction on n. The case $n = 0$ is trivial, all axioms have a context in the antecedent. For the induction step, we assume that for all proofs of height $\leq n$ we have already proved the property, and distinguish according to the last rule applied.

Since the rules without variable conditions all have an extensible context in the antecedent, we can by induction hypothesis extend the contexts of the premises, and derive the weakened sequence. For example, for $\rightarrow$-L, we can convert

$$\frac{\vdash_n{}^n \Gamma \Rightarrow \Delta, B \qquad \vdash_n{}^n D, \Gamma \Rightarrow \Theta}{\vdash_n{}^{n+1} B \rightarrow D, \Gamma \Rightarrow \Delta, \Theta}$$

into

$$\frac{\vdash_n{}^n A, \Gamma \Rightarrow \Delta, B \qquad \vdash_n{}^n D, A, \Gamma \Rightarrow \Theta}{\vdash_n{}^{n+1} B \rightarrow D, A, \Gamma \Rightarrow \Delta, \Theta}$$

For the rules with variable conditions, namely $\exists$-L and $\forall$-R the added formula A may violate the variable condition. Assume we have

$$\frac{\vdash_n{}^n B[x := c], \Gamma \Rightarrow \Delta}{\vdash_n{}^{n+1} \exists_x B, \Gamma \Rightarrow \Delta}$$

If c does not occur in A we can just extend the premise. If c occurs in A, the variable condition would not be satisfied. Let d be a new variable that does not occur yet. By the variable condition, we know that c does not occur in $\forall_x B, \Gamma, \Delta$, therefore $B[x := d] = B[x := c][c := d]$, $\Gamma[c := d] = \Gamma$, $\Delta[c := d] = \Delta$, and therefore by Lemma 3.2 we can derive $\vdash_n{}^n B[x := d], \Gamma \Rightarrow \Delta$. Then by induction, we can also derive $\vdash_n{}^n B[x := d], A, \Gamma \Rightarrow \Delta$, and since A does not contain d, the variable condition holds and we get

$$\frac{\vdash_n{}^n B[x := d], A, \Gamma \Rightarrow \Delta}{\vdash_n{}^{n+1} \exists_x B, A, \Gamma \Rightarrow \Delta}$$

Similarly for $\forall$-R. □

We will focus on the part of Nadathur's proof that is the analogon to Orevkov's σ-class $\{\to^+, \neg^+, \vee^-\}$.

Theorem 3.4. *Let* Δ *be nonempty, and* $\Gamma \Rightarrow \Delta$ *have a* C*-proof in which no instance of* $\to$*-R or* $\vee$*-L is used. Then* $\Gamma \Rightarrow G$ *has an* I*-proof for some* $G \in \Gamma$*.*

Proof We use induction on the heights of C-proofs. For height 0, only axioms can be used. For

$$\overline{A, \Gamma' \Rightarrow \Delta', A}$$

we can give the I-proof

$$\overline{A \Rightarrow A}$$

and

$$\overline{\Gamma \Rightarrow \top}$$

is already an I-proof. For heights $n > 1$ we assume the theorem holds for all proofs of height $m < n$, consider the last rule used in the proof, and prove that it holds for the whole proof if it holds for the proofs of the sequent premises of this last rule. For the contraction rules, $\forall$-L, $\exists$-L, $\vee$-R and $\wedge$-L this is trivial. For $\perp$-R, we have an I-proof for $\Gamma \Rightarrow A$ for some $A \in \Delta', \perp$. If $A = \perp$, then we can use

$$\frac{\begin{array}{c}|\\ \Gamma \Rightarrow \perp\end{array}}{\Gamma \Rightarrow D}\ \perp\text{-R}$$

for every $D \in \Delta$. If $A \in \Delta'$, then already $A \in \Delta$. For $\rightarrow$-L, we assume the induction hypothesis for proofs of $\Gamma' \Rightarrow \Delta', B$ and $D, \Gamma' \Rightarrow \Theta$. If there is an I-proof of $\Gamma' \Rightarrow A$ for some $A \in \Delta'$, we can weaken the antecedent by $B \rightarrow D$ and get an I-proof for $B \rightarrow D, \Gamma' \Rightarrow A$. Otherwise, there must be an I-proof of $\Gamma' \Rightarrow B$, and an I-proof of $D, \Gamma \Rightarrow A$ for some $A \in \Theta$. Then by

$$\dfrac{\begin{array}{c}|\\ \Gamma' \Rightarrow B\end{array} \qquad \begin{array}{c}|\\ D, \Gamma' \Rightarrow A\end{array}}{B \rightarrow D, \Gamma' \Rightarrow A} \rightarrow\text{-L}$$

the theorem holds. For $\wedge$-R, we have the premises $\Gamma \Rightarrow \Delta', B$ and $\Gamma \Rightarrow \Delta', D$. If we have an I-proof for $\Gamma \Rightarrow A$ for some $A \in \Delta'$, then we are already done. Otherwise, by induction hypothesis, we have I-proofs for $\Gamma \Rightarrow B$ and $\Gamma \Rightarrow D$, and therefore by

$$\dfrac{\begin{array}{c}|\\ \Gamma \Rightarrow B\end{array} \qquad \begin{array}{c}|\\ \Gamma \Rightarrow D\end{array}}{\Gamma \Rightarrow B \wedge D} \wedge\text{-R}$$

For $\forall$-R, we have an I-proof for $\Gamma \Rightarrow X$ for some $X \in \Delta', B[x := c]$. For $X \in \Delta'$ we are done, for $X = B[x := c]$ we can simply apply $\forall$-R on $\Gamma \Rightarrow X$. Similarly for $\exists$-R. □

The given algorithm transforming a C-proof in an I-proof is linear in the number of nodes of the proof tree, if implemented in the right way; the only case where this is non-trivial is when we have to extend the contexts in the $\rightarrow$-R case. On the one hand, with Lemma 3.3 we could add a left-weakening rule

$$\dfrac{\Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}$$

On the other hand, multisets can be efficiently implemented using binary trees, and with some pointer magic, extending all contexts may only involve the change of a pointer; however, this requires that the proof is already given in that form. If none of these actions is taken, then the algorithm remains quadratic, at least.

Notice that Calculus n is cut free by default, though, according to (4), cuts are admissible.

Corollary 3.5. *If a singular sequent $\Gamma \Rightarrow G$ not containing positive occurrences of $\rightarrow$ and not containing negative occurrences of $\vee$ has a C-proof, then it has an I-proof.*

Proof This follows directly from the above theorem: In no rule in Calculus n, a connective can be removed, and the usage of $\rightarrow$-R and $\vee$-L yields a positive $\rightarrow$ and a negative $\vee$. Thus, in a proof of our $\Gamma \Rightarrow G$ these rules cannot occur. □

This is equivalent to the case $\{\rightarrow^+, \neg^+, \vee^-\}$ of sufficiency of Theorem 3.1: Positive occurrences of $\rightarrow$ and negative occurrences of $\vee$ are forbidden. A positive occurrence of $\neg$ in terms of (4) yields a positive occurrence of an implication, as $\neg A$ is an abbreviation of $A \rightarrow \bot$.

For the necessity, see Section 5.

3.2 The Equivalence of Calculus n and Natural Deduction

For the conversion of proofs in Calculus n into proofs in natural deduction, we assume all sequences are finite, and regard sequents $\Gamma \Rightarrow \Delta$ as the universal closure of $(\bigwedge \Gamma) \rightarrow \bigvee \Delta$, that is, if $\Gamma = A_1, \ldots, A_n$ and $\Delta = B_1, \ldots, B_m$ and the sequent $\Gamma \Rightarrow \Delta$ is derivable in Calculus n, we show that $\forall_{\vec{x}}.A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m$ is derivable in natural deduction, where $\vec{x}$ contains all free variables of $A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m$ and as usual, the empty disjunction becomes $\bot$ and the empty conjunction becomes $\top$. As we add $\top$ to the list of the axioms, we have $\bot \vee D \leftrightarrow D$ and $\top \wedge B \leftrightarrow B$ which keeps everything consistent. (If we do not want to add $\top$ as an axiom, we may as well define it by $\neg\bot$, which is derivable.)

With this, of course, $\Gamma \Rightarrow$ and $\Rightarrow \Delta$ have the same embedding as $\Gamma \Rightarrow \bot$ and $\top \Rightarrow \Delta$, but there is no need of injectivity.

Associativity and commutativity of $\wedge$ and $\vee$ are trivially derivable in natural deduction. Therefore, we silently assume that a multiset Γ, A has been translated into $G \circ A$, where $\circ \in \{\vee, \wedge\}$ and G is the conjunction or disjunction of the formulae in Γ for explicit formulae A.

For the axiomatic rules, we have $\lambda_{\vec{x}} \lambda_{y^G} t^\top$ and $\lambda_{\vec{x}} \lambda_{y^{A \wedge G}}.y(u^A, v^G. \vee_+ u)$. contr-L and contr-R are derivable via

$$\dfrac{\dfrac{\dfrac{\dfrac{\forall_{\vec{x}}.G \wedge B \wedge B \rightarrow D}{G \wedge B \wedge B \rightarrow D} \qquad \dfrac{[g : G \wedge B] \qquad \dfrac{[g : G \wedge B] \qquad [B]}{B}}{G \wedge B \wedge B}}{D}}{G \wedge B \rightarrow D}\, g}{\forall_{\vec{x}}.G \wedge B \rightarrow D}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall_{\vec{x}}.G \rightarrow D \vee B \vee B}{G \rightarrow D \vee B \vee B} \qquad [g : G]}{D \vee B \vee B} \qquad [D \vee B] \qquad \dfrac{[B]}{D \vee B}}{D \vee B}}{G \rightarrow D \vee B}\, g}{\forall_{\vec{x}}.G \rightarrow D \vee B}$$

$\bot$-R is derivable by

$$
\begin{prooftree}
\AxiomC{$\forall_{\vec{x}}.G\to D\lor\bot$}
\UnaryInfC{$G\to D\lor\bot$}
\AxiomC{$[g:G]$}
\BinaryInfC{$D\lor\bot$}
\AxiomC{$[D]$}
\UnaryInfC{$D\lor E$}
\AxiomC{$\bot\to D$}
\AxiomC{$[\bot]$}
\BinaryInfC{D}
\UnaryInfC{$D\lor E$}
\TrinaryInfC{$D\lor E$}
\RightLabel{g}
\UnaryInfC{$G\to D\lor E$}
\UnaryInfC{$\forall_{\vec{x}}.G\to D\lor E$}
\end{prooftree}
$$

$\land$-L and $\land$-R are derivable by

$$
\begin{prooftree}
\AxiomC{$\forall_{\vec{x}}.B_i\land G\to D$}
\UnaryInfC{$B_i\land G\to D$}
\AxiomC{$[w:B_1\land B_2]$}
\AxiomC{$[B_i]$}
\BinaryInfC{B_i}
\AxiomC{$[g:G]$}
\BinaryInfC{$B_i\land G$}
\BinaryInfC{D}
\RightLabel{g}
\UnaryInfC{$G\to D$}
\RightLabel{w}
\UnaryInfC{$B_1\land B_2\to G\to D$}
\RightLabel{Lemma 2.4/2}
\UnaryInfC{$(B_1\land B_2)\land G\to D$}
\UnaryInfC{$\forall_{\vec{x}}.(B_1\land B_2)\land G\to D$}
\end{prooftree}
$$

$$
\begin{prooftree}
\AxiomC{$\forall_{\vec{x}}.G\to D\lor B$}
\UnaryInfC{$G\to D\lor B$}
\AxiomC{$g:G$}
\BinaryInfC{$D\lor B$}
\AxiomC{$[D]$}
\noLine
\UnaryInfC{$\vdots$}
\noLine
\UnaryInfC{$D\lor(B\land E)$}
\AxiomC{$\forall_{\vec{x}}.G\to D\lor E$}
\UnaryInfC{$G\to D\lor E$}
\AxiomC{$g:G$}
\BinaryInfC{$D\lor E$}
\AxiomC{$[D]$}
\UnaryInfC{$D\lor(B\land E)$}
\AxiomC{$[B]\ [E]$}
\UnaryInfC{$B\land E$}
\UnaryInfC{$D\lor(B\land E)$}
\TrinaryInfC{$D\lor(B\land E)$}
\TrinaryInfC{$D\lor(B\land E)$}
\RightLabel{g}
\UnaryInfC{$G\to D\lor(B\land E)$}
\UnaryInfC{$\forall_{\vec{x}}.G\to D\lor(B\land E)$}
\end{prooftree}
$$

$\lor$-L and $\lor$-R are derivable by

$$
\begin{prooftree}
\AxiomC{$u:B\lor E$}
\AxiomC{$\forall_{\vec{x}}.B\land G\to D$}
\UnaryInfC{$B\land G\to D$}
\AxiomC{$[B]$}
\AxiomC{$[g:G]$}
\BinaryInfC{$B\land G$}
\BinaryInfC{D}
\AxiomC{$\forall_{\vec{x}}.E\land G\to D$}
\UnaryInfC{$E\land G\to D$}
\AxiomC{$[E]$}
\AxiomC{$[g:G]$}
\BinaryInfC{$E\land G$}
\BinaryInfC{D}
\TrinaryInfC{D}
\RightLabel{g}
\UnaryInfC{$G\to D$}
\RightLabel{u}
\UnaryInfC{$(B\lor E)\to G\to D$}
\RightLabel{Lemma 2.4/2}
\UnaryInfC{$(B\lor E)\land G\to D$}
\UnaryInfC{$\forall_{\vec{x}}.(B\lor E)\land G\to D$}
\end{prooftree}
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall_{\vec{x}}.G \to D \lor B_i}{G \to D \lor B_i} \qquad [g : G]}{D \lor B_i} \qquad \dfrac{[D]}{D \lor (B_1 \lor B_2)} \qquad \dfrac{\dfrac{[B_i]}{B_1 \lor B_2}}{D \lor (B_1 \lor B_2)}}{D \lor (B_1 \lor B_2)}}{G \to D \lor (B_1 \lor B_2)}\, g}{\forall_{\vec{x}}.G \to D \lor (B_1 \lor B_2)}
$$

For $\forall$-L and $\to$-L we have

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall_{\vec{y}}.B[x := t] \land G \to D}{B[x := t] \land G \to D} \qquad \dfrac{\dfrac{[b : \forall_x B]}{B[x := t]} \qquad [g : G]}{B[x := t] \land G}}{D}\, g}{G \to D}\, b}{(\forall_x B) \to G \to D}\, \text{Lemma 2.4/2}}{(\forall_x B \land G) \to D}}{\forall_{\vec{y}}.(\forall_x B) \land G \to D}
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall_{\vec{x}}.G \to D \lor B}{G \to D \lor B} \qquad [g : G]}{D \lor B} \qquad \dfrac{[D]}{D \lor Q} \qquad \dfrac{\dfrac{\dfrac{\forall_{\vec{x}}.E \land G \to Q}{E \land G \to Q} \qquad \dfrac{\dfrac{[i : B \to E] \qquad [B]}{E} \qquad [g : G]}{E \land G}}{Q}}{D \lor Q}}{D \lor Q}\, g}{G \to D \lor Q}\, i}{(B \to E) \to G \to D \lor Q}\, \text{Lemma 2.4/2}}{(B \to E) \land G \to D \lor Q}}{\forall_{\vec{x}}.(B \to E) \land G \to D \lor Q}
$$

For $\exists$-L and $\exists$-R we have

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[b : \exists_x B] \qquad \dfrac{[B(= B[x := c][c := x])]}{\exists_c B[x := c]}}{\exists_c B[x := c]} \qquad \dfrac{\dfrac{\forall_{\vec{y}}.B[x := c] \land G \to D}{B[x := c] \land G \to D} \qquad \dfrac{[B[x := c]] \qquad [g : G]}{B[x := c] \land G}}{D}}{D}}{G \to D}\, g}{(\exists_x B) \to G \to D}\, b}{(\exists_x B) \land G \to D}\ \text{Lemma 2.4/2}}{\forall_{\vec{y}}.(\exists_x B) \land G \to D}
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall_{\vec{y}}.G \to D \lor B[x := t]}{G \to D \lor B[x := t]} \qquad [g : G]}{D \lor B[x := t]} \qquad \dfrac{[D]}{D \lor \exists_x B} \qquad \dfrac{\dfrac{[B[x := t]]}{\exists_x B}}{D \lor \exists_x B}}{D \lor \exists_x B}}{G \to D \lor \exists_x B}\, g}{\forall_{\vec{y}}.G \to D \lor \exists_x B}
$$

For the singular versions of $\to$-R and $\forall$-R we have

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall_{\vec{x}}.B \land G \to E}{B \land G \to E} \qquad \dfrac{[b : B] \qquad [g : G]}{B \land G}}{E}}{B \to E}\, b}{G \to (B \to E)}\, g}{\forall_{\vec{x}}.G \to (B \to E)}
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall_{\vec{y}}.G \to B[x := c]}{G \to B[x := c]} \qquad [g : G]}{B[x := c]}}{\forall_c B[x := c]}\,(*)}{B[x := c][c := x](= B)}}{\forall_x B}\,(**)}{G \to \forall_x B}\, g}{\forall_{\vec{y}}.G \to \forall_x B}
$$

where in (*) we use the fact that c is a constant that does not occur in G so the variable condition is satisfied, and in (**) we use that if x occurs in G then $x \in \vec{y}$.

For the multi-succedent versions of $\to$-R and $\forall$-R we need traditional classical logic, which complies with the definitions of I-proofs and C-proofs.

We will use the *tertium non datur* $A \vee \neg A$ which is derivable in traditional classical logic for all formulae A by

$$
\begin{prooftree}
\AxiomC{$[u : \neg(A \vee \neg A)]$}
\AxiomC{$[u : \neg(A \vee \neg A)]$}
\AxiomC{$[v : A]$}
\UnaryInfC{$A \vee \neg A$}
\BinaryInfC{$\bot$}
\RightLabel{v}
\UnaryInfC{$\neg A$}
\UnaryInfC{$A \vee \neg A$}
\BinaryInfC{$\bot$}
\RightLabel{u}
\UnaryInfC{$\neg\neg(A \vee \neg A)$}
\RightLabel{stab}
\UnaryInfC{$A \vee \neg A$}
\end{prooftree}
$$

Now the proof for multi-succedent $\to$-R goes

$$
\begin{prooftree}
\AxiomC{$\forall_{\vec{x}}.B \wedge G \to D \vee E$}
\UnaryInfC{$B \wedge G \to D \vee E$}
\AxiomC{$b : B$}
\AxiomC{$g : G$}
\BinaryInfC{$B \wedge G$}
\BinaryInfC{$D \vee E$}
\AxiomC{$[D]$}
\UnaryInfC{$D \vee (B \to E)$}
\AxiomC{$[E]$}
\RightLabel{$[k : B]$}
\UnaryInfC{$B \to E$}
\UnaryInfC{$D \vee (B \to E)$}
\TrinaryInfC{$D \vee (B \to E)$}
\noLine
\UnaryInfC{(*)}
\end{prooftree}
$$

$$
\begin{prooftree}
\AxiomC{$|$}
\noLine
\UnaryInfC{$\bot \to E$}
\AxiomC{$n : \neg B$}
\AxiomC{$[b : B]$}
\BinaryInfC{$\bot$}
\BinaryInfC{E}
\RightLabel{b}
\UnaryInfC{$B \to E$}
\UnaryInfC{$D \vee (B \to E)$}
\noLine
\UnaryInfC{(**)}
\end{prooftree}
$$

$$
\begin{prooftree}
\AxiomC{$|$}
\noLine
\UnaryInfC{$B \vee \neg B$}
\AxiomC{(*)}
\AxiomC{(**)}
\RightLabel{b, n}
\TrinaryInfC{$D \vee (B \to E)$}
\RightLabel{g}
\UnaryInfC{$G \to D \vee (B \to E)$}
\UnaryInfC{$\forall_{\vec{x}}.G \to D \vee (B \to E)$}
\end{prooftree}
$$

And the proof for multi-succedent $\forall$-R goes

$$
\begin{prooftree}
\AxiomC{$\forall_{\vec{y}}.G\to D\vee B[x:=c]$}
\UnaryInfC{$G\to D\vee B[x:=c]$}
\AxiomC{$g:G$}
\BinaryInfC{$D\vee B[x:=c]$}
\AxiomC{$\mid$}
\noLine
\UnaryInfC{$\bot\to B[x:=c]$}
\AxiomC{$n:\neg D$}
\AxiomC{$[D]$}
\BinaryInfC{$\bot$}
\BinaryInfC{$B[x:=c]$}
\AxiomC{$[B[x:=c]]$}
\TrinaryInfC{$B[x:=c]$}
\RightLabel{(*)}
\UnaryInfC{$\forall_c B[x:=c]$}
\UnaryInfC{$B(=B[x:=c][c:=x])$}
\RightLabel{(**)}
\UnaryInfC{$\forall_x B$}
\UnaryInfC{$D\vee\forall_x B$}
\noLine
\UnaryInfC{(***)}
\end{prooftree}
$$

$$
\begin{prooftree}
\AxiomC{$\mid$}
\noLine
\UnaryInfC{$D\vee\neg D$}
\AxiomC{$[d:D]$}
\UnaryInfC{$D\vee\forall_x B$}
\AxiomC{(***)}
\RightLabel{d, n}
\TrinaryInfC{$D\vee\forall_x B$}
\RightLabel{g}
\UnaryInfC{$G\to D\vee\forall_x B$}
\UnaryInfC{$\forall_{\vec{y}}.G\to D\vee\forall_x B$}
\end{prooftree}
$$

Where (*) and (**) are like in the singular variant.

Notice that only $\bot$-R and the multi-succedent versions of $\forall$-R and $\to$-R use ex falso quodlibet, thus the following lemma holds.

Lemma 3.6. *An* I*-proof in Calculus* n *without applications of* $\bot$*-R can be converted into a minimal proof in natural deduction.*

We gave an embedding of the Calculus n into natural deduction. As for every rule there can be given a simple template to translate it into natural deduction, if we ignore the necessity of ordering disjunctions and conjunctions, this would be possible in linear time relative to the number of nodes of the proof, so in an actual implementation, it is desirable to minimize the parts where conjunctions and disjunctions are reorganized. Reorganizing conjunctions works in linear time: Disassembling the conjunctions takes one step for every $\wedge$, and then putting them together again in another direction as well. For example, to convert $A\wedge(B\wedge(C\wedge D))$ into $((A\wedge B)\wedge C)\wedge D$, we need the derivation

$$
\dfrac{A\wedge(B\wedge(C\wedge D)) \qquad \dfrac{[p : B\wedge(C\wedge D)] \qquad \dfrac{[q : C\wedge D] \qquad \dfrac{\dfrac{\dfrac{[a : A] \qquad [b : B]}{A\wedge B} \qquad [c : C]}{(A\wedge B)\wedge C} \qquad [d : D]}{((A\wedge B)\wedge C)\wedge D}}{((A\wedge B)\wedge C)\wedge D}\; c, d}{((A\wedge B)\wedge C)\wedge D}\; b, q}{((A\wedge B)\wedge C)\wedge D}\; a, p
$$

Ordering disjunctions takes at most quadratic time relative to the length of the formula: Disassembling it forks once for every $\wedge$ or $\vee$, and in every fork, we need to introduce (in the worst case) every $\wedge$ or $\vee$ again.

For the other direction we show that if t is an intuitionistic (traditional-classical) proof and $seq(t) = \Gamma \Rightarrow A$, then there is an I-proof (C-proof) in Calculus n deriving $\Gamma \Rightarrow A$.

We need two more rules which can be shown to be valid in Calculus n. Firstly, a left weakening rule for singular sequents, which is valid according to Lemma 3.3.

$$
\dfrac{\Gamma \Rightarrow A}{B, \Gamma \Rightarrow A}\, LW_n
$$

Secondly, the axiomatic rule deriving $A, \Gamma \Rightarrow \Delta, A$ requires A to be atomic, for the singular calculus, it becomes $A, \Gamma \Rightarrow A$, and we show that in the singular calculus we can derive $B, \Gamma \Rightarrow B$ for arbitrary B by structural induction (where c shall be a constant not occurring in Γ in every step):

$$
\dfrac{\dfrac{B, \Gamma \Rightarrow B}{B\wedge C, \Gamma \Rightarrow B} \qquad \dfrac{C, \Gamma \Rightarrow C}{B\wedge C, \Gamma \Rightarrow C}}{B\wedge C, \Gamma \Rightarrow B\wedge C}
\qquad
\dfrac{\dfrac{B, \Gamma \Rightarrow B}{B, \Gamma \Rightarrow B\vee C} \qquad \dfrac{C, \Gamma \Rightarrow C}{C, \Gamma \Rightarrow B\vee C}}{B\vee C, \Gamma \Rightarrow B\vee C}
$$

$$
\dfrac{\dfrac{B, \Gamma \Rightarrow B \qquad C, B, \Gamma \Rightarrow C}{B\rightarrow C, B, \Gamma \Rightarrow C}}{B\rightarrow C, \Gamma \Rightarrow B\rightarrow C}
$$

$$
\dfrac{\dfrac{B[x := c], \Gamma \Rightarrow B[x := c]}{\forall_x B, \Gamma \Rightarrow B[x := c]}}{\forall_x B, \Gamma \Rightarrow \forall_x B}
\qquad
\dfrac{\dfrac{B[x := c], \Gamma \Rightarrow B[x := c]}{B[x := c], \Gamma \Rightarrow \exists_x B}}{\exists_x B, \Gamma \Rightarrow \exists_x B}
$$

By (4) we know that a cut rule holds:

$$\frac{\Gamma_1 \Rightarrow A \qquad A, \Gamma_2 \Rightarrow B}{\Gamma_1, \Gamma_2 \Rightarrow B}$$

We limit ourselves to singular sequents, except for showing stability, as stability - of course - cannot be shown intuitionistically. Stability can be derived by

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg\neg A, A \Rightarrow A, \bot}{\neg\neg A \Rightarrow A, \neg A} \qquad \dfrac{\bot, \neg\neg A \Rightarrow \bot}{\bot, \neg\neg A \Rightarrow A}}{\neg\neg A, \neg\neg A \Rightarrow A, A}}{\neg\neg A \Rightarrow A, A}}{\neg\neg A \Rightarrow A}}{\Rightarrow \neg\neg A \to A}$$

which actually requires an instance of the multi-succedent implication introduction. For ex falso quodlibet, we have the derivation

$$\dfrac{\dfrac{\bot \Rightarrow \bot}{\bot \Rightarrow A}}{\Rightarrow \bot \to A}$$

The introduction rules for $\forall$, $\exists$, $\wedge$, $\vee$, $\to$ in natural deduction directly correspond to the rules for the right side in Calculus $\mathfrak{n}$, except that we may not have equal contexts on both premises, which can be changed by applying instances of the left weakening rule. For the elimination rules we need a cut rule

$$\frac{\Gamma_1 \Rightarrow \Delta_1, B \qquad B, \Gamma_2 \Rightarrow \Delta_2}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2}$$

$$\frac{\Gamma_1 \Rightarrow B \qquad B, \Gamma_2 \Rightarrow C}{\Gamma_1, \Gamma_2 \Rightarrow C}$$

which is admissible according to (4). For $\to_{-}$ we have the sequents $\Gamma'' \Rightarrow A \to B$ and $\Gamma' \Rightarrow A$ and derive $\Gamma'', \Gamma' \Rightarrow B$. So we assume that we already can derive sequents $\Gamma'' \Rightarrow A \to B$ and $\Gamma' \Rightarrow A$, and weaken both sequents to $\Gamma \Rightarrow A$ and $\Gamma \Rightarrow A \to B$ with $\Gamma = \Gamma', \Gamma''$. Then we can derive $\Gamma \Rightarrow B$ by

$$\dfrac{\Gamma \Rightarrow A \to B \qquad \dfrac{\Gamma \Rightarrow A \qquad \dfrac{\Gamma, A \Rightarrow A \qquad \Gamma, B \Rightarrow B}{\Gamma, A, A \to B \Rightarrow B}}{\Gamma, A \to B \Rightarrow B}}{\Gamma \Rightarrow B}$$

We proceed similarly for $\vee_-$

$$\dfrac{\Gamma \Rightarrow A \vee B \qquad \dfrac{\Gamma, A \Rightarrow C \qquad \Gamma, B \Rightarrow C}{\Gamma, A \vee B \Rightarrow C}}{\Gamma \Rightarrow C}$$

and for $\wedge_-$

$$\dfrac{\Gamma \Rightarrow A \wedge B \qquad \dfrac{\dfrac{\dfrac{\Gamma, A, B \Rightarrow C}{\Gamma, A \wedge B, B \Rightarrow C}}{\Gamma, A \wedge B, A \wedge B \Rightarrow C}}{\Gamma, A \wedge B \Rightarrow C}}{\Gamma \Rightarrow C}$$

and for $\forall_-$

$$\dfrac{\Gamma \Rightarrow \forall_x A \qquad \dfrac{\Gamma, A[x := t] \Rightarrow A[x := t]}{\Gamma, \forall_x A \Rightarrow A[x := t]}}{\Gamma \Rightarrow A[x := t]}$$

For $\exists_-$, we must take care of the variable condition: We have a derivation with the sequent $\Gamma \Rightarrow \exists_x A$, and a derivation $\Gamma', A \Rightarrow C$ where x must not occur freely in C and Γ', therefore we may have to replace formulae by α-equivalent ones, that is, by formulae in which bound variables are renamed. By

$$\dfrac{\dfrac{A[x := c] \Rightarrow A[x := c]}{\forall_x A \Rightarrow A[x := c]}}{\forall_x A \Rightarrow \forall_y . A[x := y]}$$

and

$$\dfrac{\dfrac{A[x := c] \Rightarrow A[x := c]}{A[x := c] \Rightarrow \exists_y . A[x := y]}}{\exists_x A \Rightarrow \exists_y . A[x := y]}$$

we know that this is allowed, and therefore, for $\exists$-L it is sufficient to have the variable not occurring freely in the conclusion, and we may use

$$\dfrac{\Gamma \Rightarrow \exists_x A \qquad \dfrac{\Gamma', A \Rightarrow C}{\Gamma', \exists_x A \Rightarrow C}}{\Gamma, \Gamma' \Rightarrow C}$$

This way we have an embedding of natural deduction into Calculus n, and vice versa. We can therefore use some structural properties of Calculus n to prove properties of proofs in natural deduction:

Lemma 3.7. *If $\Gamma \vdash_i A$ such that the sequent $\Gamma \Rightarrow A$ does not contain $\perp$ negatively, then $\Gamma \vdash_m A$.*

Proof By Lemma 3.6, it suffices to show that no instance of $\perp$-R can be used in an I-proof of $\vdash_n \Gamma \Rightarrow A$.

Firstly we notice that all the rules preserve signa of the (sub)formulae in the sequents. The signa of the formulae in the contexts stay the same. Introducing a connective or quantifier other than $\rightarrow$ does not change the signa of the explicit formulae either, nor does contraction. $\rightarrow$-R and $\rightarrow$-L take a positive and a negative formula, and create an implication out of them. Still, the signa of the explicit formulae of these rules are preserved.

The rules that may introduce a subformula $\perp$ in the succedent while $\perp$ not being a subformula in a succedent of the premises are $\vee$-R, $\rightarrow$-R and the axiom deriving $\Gamma, \perp \Rightarrow \perp$. However, after an application of $\vee$-R or $\rightarrow$-R, $\perp$ will be a proper subformula, and since all rules only build up formulae, there cannot be an application of $\perp$-R, which clearly requires $\perp$ not to be a proper subformula. Thus, if we have an application of $\perp$-R, this $\perp$ must have been introduced by an axiom deriving $\Gamma, \perp \Rightarrow \perp$.

But that sequent contains a negative occurrence of $\perp$, and as $\perp$-R only removes a positive $\perp$, after all further rules a negative occurrence of $\perp$ will remain, possibly being a subformula of some larger formula, as all rules only build up (it may be main formula of a contraction rule, but the resulting sequent still contains a negative occurrence of $\perp$).

Therefore, with an application of $\perp$-R, there must be a negative occurrence of $\perp$ in the resulting sequent, which we assumed not to have.

Thus, we will in fact get a minimal proof. $\square$

3.3 A Sketch of Orevkov's Proof

The proof given by Orevkov in (7) is extremely hard to read, relevant parts are just referenced to other papers, and the major part is rather technical, and since we are using a different notation, outside our scope. For historical reasons, we sketch the proof for the sequent class $\{\rightarrow^+, \neg^+, \vee^-\}$, as this will be the case we will concentrate on. We will additionally give some essential concepts from (2) and (3) which are only cited in (7). However, we are trying to give an impression on how this proof essentially works, especially since it seems to be the first proof of that kind and is very similar to the other proofs we have seen. We will adapt the rules given in (7) to our notation. In particular, (7) considers rules of a calculus as templates, and what we actually call a rule (the template with explicit values) is considered as a realization of that rule. We will however use our usual notation, differing from (7), and this should be kept in mind when comparing this text to (7).

3.3.1 The Calculi

The formalism used in (7) does not have a falsum $\perp$, but therefore has negation $\neg$ as an elementary connective. By a we shall mean a constant in what follows, and by t an arbitrary term. The classical calculus C^+, for which we use the entailment relation $\vdash_{C^+}$, has the rules

$$\frac{}{A \Rightarrow A}\ \text{Axiom}$$

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}\ [\mathrm{Y}\vdash_{C^+}] \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A}\ [\vdash_{C^+}\mathrm{Y}]$$

$$\frac{A, \Gamma' \Rightarrow \Delta', B, \Delta''}{\Gamma' \Rightarrow \Delta', (A \rightarrow B), \Delta''}\ [\vdash_{C^+}\rightarrow] \qquad \frac{\Gamma', \Gamma'' \Rightarrow \Delta', A \qquad \Gamma', B, \Gamma'' \Rightarrow \Delta'}{\Gamma', (A \rightarrow B), \Gamma'' \Rightarrow \Delta'}\ [\rightarrow\vdash_{C^+}]$$

$$\frac{\Gamma' \Rightarrow \Delta', A, \Delta'' \qquad \Gamma' \Rightarrow \Delta', B, \Delta''}{\Gamma' \Rightarrow \Delta', (A \wedge B), \Delta''}\ [\vdash_{C^+}\wedge] \qquad \frac{\Gamma', A, B, \Gamma'' \Rightarrow \Delta'}{\Gamma', (A \wedge B), \Gamma'' \Rightarrow \Delta'}\ [\wedge\vdash_{C^+}]$$

$$\frac{\Gamma' \Rightarrow \Delta', A, B, \Delta''}{\Gamma' \Rightarrow \Delta'', (A \vee B), \Delta''}\ [\vdash_{C^+}\vee] \qquad \frac{\Gamma', A, \Gamma'' \Rightarrow \Delta' \qquad \Gamma', B, \Gamma'' \Rightarrow \Delta'}{\Gamma', (A \vee B), \Gamma'' \Rightarrow \Delta'}\ [\vee\vdash_{C^+}]$$

$$\frac{A, \Gamma' \Rightarrow \Delta', \Delta''}{\Gamma' \Rightarrow \Delta', \neg A, \Delta''}\,[\vdash_{C^+}\neg] \qquad \frac{\Gamma', \Gamma'' \Rightarrow \Delta', A}{\Gamma', \neg A, \Gamma'' \Rightarrow \Delta'}\,[\neg\vdash_{C^+}]$$

$$\frac{\Gamma' \Rightarrow A[x := a]}{\Gamma' \Rightarrow \Delta', \forall_x A, \Delta''}\,[\vdash_{C^+}\forall] \qquad \frac{\Gamma', A[x := t], \forall_x A, \Gamma'' \Rightarrow \Delta'}{\Gamma', \forall_x A, \Gamma'' \Rightarrow \Delta'}\,[\forall\vdash_{C^+}]$$

$$\frac{\Gamma' \Rightarrow \Delta', A[x := t], \exists_x A, \Delta''}{\Gamma' \Rightarrow \Delta', \exists_x A, \Delta''}\,[\vdash_{C^+}\exists] \qquad \frac{\Gamma', A[x := a], \Gamma'' \Rightarrow \Delta'}{\Gamma', \exists_x A, \Gamma'' \Rightarrow \Delta'}\,[\exists\vdash_{C^+}]$$

where in $\vdash_{C^+}\forall$ and $\exists\vdash_{C^+}$ a is free for x in A, and does not occur freely in Γ', Γ'', Δ, $\exists_x A$ and $\forall_x A$, and where in $\forall\vdash_{C^+}$ and $\vdash_{C^+}\exists$, t is free for x in A.

We call $[Y\vdash_{C^+}]$ and $[\vdash_{C^+}Y]$ thinning rules, and every other rule except Axiom we call logical rules.

The intuitionistic (multi-succedent) calculus K_m^+, for which we use the entailment relation $\vdash_{K_m^+}$, has the rules

$$\frac{}{A \Rightarrow A}\,\text{Axiom}$$

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}\,[Y\vdash_{K_m^+}] \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A}\,[\vdash_{K_m^+}Y]$$

$$\frac{A, \Gamma' \Rightarrow B}{\Gamma' \Rightarrow \Delta', A \to B, \Delta''}\,[\vdash_{K_m^+}\to] \qquad \frac{\Gamma', A \to B, \Gamma'' \Rightarrow \Delta' \qquad \Gamma', B, \Gamma'' \Rightarrow \Delta'}{\Gamma', A \to B, \Gamma'', \Delta'}\,[\to\vdash_{K_m^+}]$$

$$\frac{\Gamma' \Rightarrow \Delta', A, \Delta'' \qquad \Gamma' \Rightarrow \Delta', B, \Delta''}{\Gamma' \Rightarrow \Delta', A \wedge B, \Delta''}\,[\vdash_{K_m^+}\wedge] \qquad \frac{\Gamma', A, B, \Gamma'' \Rightarrow \Delta'}{\Gamma', A \wedge B, \Gamma'' \Rightarrow \Delta'}\,[\wedge\vdash_{K_m^+}]$$

$$\frac{\Gamma' \Rightarrow \Delta', A, B, \Delta''}{\Gamma' \Rightarrow \Delta', A \vee B, \Delta''}\,[\vdash_{K_m^+}\vee] \qquad \frac{\Gamma', A, \Gamma'' \Rightarrow \Delta' \qquad \Gamma', B, \Gamma'' \Rightarrow \Delta'}{\Gamma', A \vee B, \Gamma'' \Rightarrow \Delta'}\,[\vee\vdash_{K_m^+}]$$

$$\frac{A, \Gamma' \Rightarrow}{\Gamma' \Rightarrow \Delta', \neg A, \Delta''}\,[\vdash_{K_m^+}\neg] \qquad \frac{\Gamma', \neg A, \Gamma'' \Rightarrow \Delta', A}{\Gamma', \neg A, \Gamma'' \Rightarrow \Delta'}\,[\neg\vdash_{K_m^+}]$$

$$\frac{\Gamma' \Rightarrow A[x := a]}{\Gamma' \Rightarrow \Delta', \forall_x A, \Delta''}\,[\vdash_{K_m^+}\forall] \qquad \frac{\Gamma', A[x := t], \forall_x A, \Gamma'' \Rightarrow \Delta'}{\Gamma', \forall_x A, \Gamma'' \Rightarrow \Delta'}\,[\forall\vdash_{K_m^+}]$$

$$\frac{\Gamma' \Rightarrow \Delta', A[x := t], \exists_x A, \Delta''}{\Gamma' \Rightarrow \Delta', \exists_x A, \Delta''}\,[\vdash_{K_m^+}\exists] \qquad \frac{\Gamma', A[x := a], \Gamma'' \Rightarrow \Delta'}{\Gamma', \exists_x A, \Gamma'' \Rightarrow \Delta'}\,[\exists\vdash_{K_m^+}]$$

The same variable conditions as for C^+ hold, and they are separated into logical or thinning rules in the same way.

In an introductory rule for a connective or quantifier $\circ$, introducing $A \circ B$ or $\circ A$, we say that the explicit occurrence of $A \circ B$ or $\circ A$ in the conclusion is the *principal occurrence* of that rule. The explicit occurrences of A and B will be called *side occurrences*. The explicit occurrence of $A \circ B$ or $\circ A$ in the premises will be called *quasi-principal occurrences*.

The formula $A \circ B$ or $\circ A$ will be called *principal formula*, and the formulae A and B *side formulae*. The formulae in the context will be called *parametric formulae*, and their occurrences *parametric occurrences*.

Lemma 3.8. *If the end sequent of a C^+-deduction D does not contain occurrences of the type $\circ^+$ (of the type $\circ^-$), then D does not contain applications of the rule $\vdash_{C^+}\circ$ (of the rule $\circ\vdash_{C^+}$).*

Proof This is trivial. If a rule $\vdash_{C^+}\circ$ ($\circ\vdash_{C^+}$) is applied, then the resulting sequent will contain $\circ$ positively (negatively). As no rule replaces any subformula, and as no rule changes any signum of a connective, this occurrence will remain until the end sequent is reached. $\square$

3.3.2 Pruned Deductions

Proofs may contain steps that are not necessary. The following definitions address this problem.

Definition A deduction D has the *pure variable property*, if no variable occurs both bound and free in D, and for each application of $\vdash_{C^+/K^+_m}\forall$ and $\exists\vdash_{C^+/K^+_m}$, using the namings in the above definitions, a only occurs freely in sequents above the conclusion of that application. (If $x \notin FV(A)$, a should be chosen so, too - that is, in the worst case, uniquely chosen for this application).

Definition A formula occurrence o of a formula A is called *traceable to the axioms* if one of the following conditions holds:

- o is the explicit occurrence of A in the rule Axiom.
- o occurs in the context of the conclusion of a rule, and there is a formula occurrence of A in the corresponding context of one of the sequent premises of this rule that is itself traceable to the axioms.
- o is the conclusion of a rule for a connective $\circ \in \{\rightarrow, \vee, \wedge\}$, such that $A = B \circ C$, and the explicit occurrences of B and C in the premises of this rule are traceable to the axioms.

- o is the conclusion of a rule for $\neg$, such that $A = \neg B$, and the explicit occurrence of B in the premise of this rule is traceable to the axioms.
- o is the explicit formula $\mathfrak{Q}_x A$ of a rule for a quanifier, and every explicit occurrence of A (which is possibly changed by a substitution) or every explicit occurrence of $\mathfrak{Q}_x A$ in the premises is traceable to the axioms.

Definition A deduction D in the calculus C^+ or in the calculus K_m^+ is *pruned* if it has the pure variable property and satisfies:

1. All occurrences of formulae in any sequent occurring in D as the conclusion of an application of a logical rule are traceable to the axioms.
2. All side formulae of occurrences of applications of the rules $\vdash_{C^+/K_m^+}\wedge$, $\vee\vdash_{C^+/K_m^+}$, $\to\vdash_{C^+/K_m^+}$, $\forall\vdash_{C^+/K_m^+}$, $\vdash_{C^+/K_m^+}\exists$ and $\neg\vdash_{C^+/K_m^+}$ are traceable to the axioms.

Theorem 3.9. *Every C^+ derivation has a pruned form.*

Lemma 3.10. *Let D be a pruned C^+-deduction without applications of $\vdash_{C^+}\to$, $\vdash_{C^+}\neg$ and $\vee\vdash_{C^+}$, then all occurrences of formulae except maybe one are introduced by thinnings into the succedent of sequents occurring in D.*

We will not prove Theorem 3.9 and Lemma 3.10 here. They are proved in (7).

Corollary 3.11. *Any pruned C^+-deduction not containing applications of $\vdash_{C^+}\to$, $\vdash_{C^+}\neg$ and $\vee\vdash_{C^+}$ can be transformed into a deduction in the calculus K_m^+ with the same end sequent.*

Proof We have to look at the remaining rules. Axiom and the thinning rules in both calculi are equal, so are the rules for $\wedge$, $\forall$, $\exists$ and so is $\vdash_{C^+/K_m^+}\vee$. $\to\vdash_{K_m^+}$ and $\to\vdash_{C^+}$ differ, as $\to\vdash_{K_m^+}$ requires a quasi-principal occurrence of $A\to B$ while $\to\vdash_{C^+}$ does not. On the other hand, by Lemma 3.10, in $\to\vdash_{C^+}$, at least one of A or B must have been introduced by a thinning into the succedent, and this would contradict the prunedness. In $\neg\vdash_{C^+}$, if A was introduced by a thinning into the succedent, this would contradict the prunedness. If A was introduced by an axiom $A \Rightarrow A$, we can safely introduce the $\neg A$ in the antecedent by a thinning rule, and then relpace $\neg\vdash_{C^+}$ by $\neg\vdash_{K_m^+}$, and then push this thinning rule to the bottom as far as possible to regain a pruned deduction. Doing this, we get a derivation in K_m^+. $\square$

The sufficiency of $\{\to^+, \neg^+, \vee^-\}$ follows directly from this. For the necessity, see 5.

Notice that the calculi given here are cut free.

4 Removing stability axioms

Again we focus on the complete Glivenko class $\{\to^+, \neg^+, \vee^-\}$. In the notation of the calculus of natural deduction, $\neg A$ is expressed by $A \to \bot$. Furthermore, we usually do not talk about sequents, but of proofs with free assumptions. So, deriving $\Gamma \vdash A$ means that we have a proof of A in natural deduction with all free assumptions of this proof being in Γ.

4.1 Concepts on Natural Deduction

The following two theorems are well-known for the calculus of natural deduction, and proofs are given in (11).

Theorem 4.1 (Normalization)**.** *For every proof M^A there is a proof N^A with the same or fewer free assumptions, in which no introduced connective is eliminated again. We call such a proof normal or in normal form, and the process of transforming a proof into its normal form we call normalization.*

Remark In terms of the lambda calculus, this corresponds to the β normal Form.

Theorem 4.2 (Subformula Property)**.** *Let t be a proof of $\Gamma \vdash_m A$ in normal form. Then every formula occurring in t is a subformula of one of the formulae in $\Gamma \cup \{A\}$.*

Theorem 4.3 (Long Normal Form)**.** *For every proof M^A there is a proof N^A with the same or fewer free assumptions, which is normal, and in which every implication and universal quantification that is not introduced is eliminated. We say such a proof has long normal form.*

This is obvious, as we may expand every problematic u in a normal term to $\lambda_x.ux$ (this is called η expansion).

Notice that another usual definition of the long normal form requires all connectives that are not introduced to be eliminated. However, one may lose uniqueness of the normal form with this definition, and in every case we will deal with, we neither need its properties, nor do we need the uniqueness, so it makes no difference which definition we use.

The following definitions look complicated at first sight, but in fact, they are the canonical definitions for their purpose. By the *discharge* of a formula occurrence we shall mean the binding of a free assumption in which the formula occurs, through $\vee_-$, $\wedge_-$ or $\exists_-$, which usually makes this formula occurrence disappear.

Definition A *track* of a derivation M is a sequence $(A_i)_{0\le i\le n}$ of formula occurrences such that

- A_0 is a top formula occurrence in M that is not discharged by an instance of $\lor_-$, $\land_-$ or $\exists_-$.
- For $i < n$, A_i is not the minor premise of an instance of $\to_-$, and exactly one of the following conditions holds:
  - A_i is not the major premise of an instance of $\lor_-$, $\land_-$ or $\exists_-$, and A_{i+1} is directly below A_i.
  - A_i is the major premise of an instace of $\lor_-$, $\land_-$ or $\exists_-$, and A_{i+1} is an assumption discharged by this instance.
- Exactly one of the following conditions hold for A_n:
  - A_n is the minor premise of an instance of $\to_-$
  - A_n is the conclusion of M
  - A_n is the major premise of an instance of $\land_-$, $\lor_-$ or $\exists_-$ and there are no assumptions discharged by this instance

Definition A *segment* (of length n) in a derivation M is a sequence $(A_i)_{1\leq i\leq n}$ of occurrences of a formula A, such that

- for $1 \leq i < n$, A_i is a minor premise of an application of $\lor_-$, $\land_-$ or $\exists_-$ with conclusion A_{i+1}.
- A_n is not a minor premise of $\lor_-$, $\land_-$ or $\exists_-$.
- A_1 is not the conclusion of $\lor_-$, $\land_-$ or $\exists_-$.

Notice that by this definition, a formula occurrence which is neither a minor premise nor the conclusion of an application of $\lor_-$, $\land_-$ or $\exists_-$, also forms a segment of length 1.

The following Lemma will be useful in the further proofs, and is crucial, as we will show in a counterexample in Section 5.

Lemma 4.4. *Let* $\Gamma\vdash_{ec}A$ *where* A *is* $\to$*-free and* Γ *contains* $\to$ *only positively. Then in a proof of this in long normal form no bound variables occur that contain negative implications.*

Proof Assume a bound variable $u : B$, where B contains negative implications. By Theorem 4.2, B is subformula of a stability axiom, which means that $B = \neg\neg C$ or $B = \neg\neg C\to C$ for some atomic C.

In case $B = \neg\neg C$, from Theorem 4.2 clearly follows that it cannot be bound by anything different from $\to_+$ with conclusion C. However, this formula will then be a subformula of all further derived formulae, as all elimination rules would require a formula with a negative occurrence of $\neg\neg C\to C$, which is forbidden according to Theorem 4.2.

In case $B = \neg\neg C\to C$, every way of binding it would require a formula with B as subformula, which is forbidden due to Theorem 4.2. $\square$

4.2 Extended Classical Proofs

It is known that we cannot derive all of *STAB* from *ASTAB*, but by Lemma 2.4/4, 5, 6 and 7 we know that we can derive stability at least for all formulae not containing $\exists$ and $\vee$.

Trivially, extended classical provability follows from traditional-classical provability. Because of

$$\frac{\neg\neg A\to A \qquad \dfrac{\bot}{\neg\neg A}\,[u:\neg A]}{A}$$

we know that $ASTAB\vdash_m EFQ$ and therefore intuitionistic derivability implies extended classical derivability.

The main difference now lies in formulae containing $\exists$ and $\vee$: We can in general not derive $A\vee B\leftrightarrow\neg(\neg A\wedge\neg B)$ and $\exists_x A\leftrightarrow\neg\forall_x\neg A$, as we can do in traditional classical logic.

We therefore define the weak existence quantifier $\tilde{\exists}$ and the weak disjunction $\tilde{\vee}$ by $\tilde{\exists}_x A := \neg\forall_x\neg A$ and $A\tilde{\vee}B := \neg(\neg A\wedge\neg B)$.

We get the introduction rules $A\to A\tilde{\vee}B$, $B\to A\tilde{\vee}B$, $\forall_x.A\to\tilde{\exists}_x A$ by

$$\dfrac{\dfrac{A \qquad \dfrac{[\neg A\wedge\neg B] \qquad [\neg A]}{\neg A}}{\bot}}{A\tilde{\vee}B} \qquad\qquad \dfrac{\dfrac{B \qquad \dfrac{[\neg A\wedge\neg B] \qquad [\neg B]}{\neg B}}{\bot}}{A\tilde{\vee}B}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\forall_x\neg A]}{\neg A} \qquad [A]}{\bot}}{\tilde{\exists}_x A}}{A\to\tilde{\exists}_x A}}{\forall_x.A\to\tilde{\exists}_x A}$$

and the elimination rules $A\tilde{\vee}B\to(A\to C)\to(B\to C)\to C$ and $\forall_x.\tilde{\exists}_x A\to(\forall_x.A\to C)\to C$ we get from

$$\dfrac{\neg\neg C\to C \qquad \dfrac{\dfrac{A\tilde{\vee}B \qquad \dfrac{\begin{matrix}[u:\neg C] \quad A\to C\\ \vdots \\ \neg A\end{matrix} \qquad \begin{matrix}[u:\neg C]\quad B\to C\\ \vdots\\ \neg B\end{matrix}}{\neg A\wedge\neg B}}{\bot}}{\neg\neg C}\,u}{C}$$

$$\dfrac{\neg\neg C\to C \qquad \dfrac{\dfrac{\tilde{\exists}_x A \qquad \dfrac{\begin{matrix}\dfrac{\forall_x.A\to C}{A\to C} \quad [\neg C]\\ \vdots\\ \neg A\end{matrix}}{\forall_x.\neg A}}{\bot}}{\neg\neg C}}{C}$$

Using these abbrevitations, we may consider extended classical logic as a fragment of minimal logic.

4.3 Removing Stability Axioms in Extended Classical Proofs

For extended classical logic, we get a similar theorem to the case of the σ-class $\{\to^+, \vee^-, \neg^-\}$ of Orevkov's theorem:

Theorem 4.5. *Assume* $\Gamma \vdash_{ec} A$, *where the sequent* $\Gamma \Rightarrow A$

1. *does not contain positive occurrences of* $\to$.
2. *does not contain negative occurrences of* $\bot$ *and* $\vee$.

Then there is a quadratic algorithm transforming a proof t *of* $\Gamma \vdash_{ec} A$ *in long normal form into a minimal one.*

Proof As positive implications are forbidden, A must not contain implications at all, and as all negative implications in Γ are positive in $\Gamma \Rightarrow A$, there must not be a negative implication in Γ. $\bot$ must not occur positively in Γ. Furthermore, negative strong disjunctions are forbidden in $\Gamma \Rightarrow A$, that is, in Γ there must not be positive disjunctions.

If a stability axiom $\forall_{\vec{x}}.\neg\neg P\vec{x}\to P\vec{x}$ is used, since we are in long normal form, it must be eliminated by a sequence of $\forall$ eliminations, followed by an $\to_-$ application with

$\neg\neg P\vec{r}$, deriving $P\vec{r}$. Assume it was a top most application of stability. As $\neg\neg P\vec{r}$ contains a negative implication, by Lemma 4.4 it cannot be an assumption, and must hence have been derived by $\to_+$, which must then have been in the scope of a derivation of $\perp$ from $\neg P\vec{r}$. Thus, every top most stability axiom occurs in a context

$$
\dfrac{\begin{array}{c} Stab_{P,\vec{x}} \\ \mid \\ \neg\neg P\vec{r}\to P\vec{r} \end{array} \qquad \begin{array}{c} \mid M \\ \dfrac{\perp}{\neg\neg P\vec{r}}\to_+ u \\ \mid S \\ \neg\neg P\vec{r} \end{array}}{\begin{array}{c} P\vec{r} \\ \mid K \\ A \end{array}}
$$

where S is a segment. S can only fork if it contains an instance of $\vee_-$ as this is the only rule with two minor premises, but this is clearly impossible since Γ must not contain positive occurrences of $\vee$. Thus, the instance $\to_+ u$ is unique.

Now assume M had no occurrence of u. Assume B was a top node of a main track in M. Then B must contain a strictly positive occurrence of $\perp$. On the other hand, B can never be bound: $\vee_-$ or $\exists_-$ requires a formula with a strictly positive occurrence of $\perp$ which cannot be derived if there are no such occurrences in at least one of the assumptions, while $\to_+$ would contradict normality, as the implication would have to be eliminated again. Thus, B would have to be in Γ. Contradiction.

So we still have our top most application of stability, and we know that M must contain a free occurrence of $u : \neg P\vec{r}$, which must then be eliminated with a derivation N of $P\vec{r}$. We therefore have the form

$$
\dfrac{\begin{array}{c} Stab_{P,\vec{x}} \\ \mid \\ \neg\neg P\vec{r}\to P\vec{r} \end{array} \qquad \begin{array}{c} \dfrac{[u : \neg P\vec{r}] \qquad \begin{array}{c} \mid N \\ P\vec{r} \end{array}}{\perp} \\ \mid M' \\ \dfrac{\perp}{\neg\neg P\vec{r}}\to_+ u \\ \mid S \\ \neg\neg P\vec{r} \end{array}}{\begin{array}{c} P\vec{r} \\ \mid K \\ A \end{array}}
$$

where u should not be free in N; such an N trivially exists, we can for example choose a top most free occurrence of u, then it must be eliminated by such an N.

First notice that no variable that actually occurs in N can be bound by $\to_+$ inside M' and S: Any application of $\to_+$ to bind a variable must be in M', since S is a segment. The introduced $\to$ must disappear again, which can only be achieved by $\to_-$ with a major premise that has a negative implication due to normality.

With an additional occurrence of $P\vec{r}$ or $\neg P\vec{r}$ this is impossible since $P\vec{r}$ is not an implication. With another additional assumption variable with a negative implication, there would have to be another $\to_-$ in the proof which either violates the normality, or leads to the necessity of yet another negative implication, and so on, and can therefore not be possible.

Thus, this is only possible with a stability axiom, but we chose a top most application of stability.

Furthermore, there can trivially be no $\vee_-$ in M' or S since that would require a positive occurrence of $\vee$. Thus, all variables are bound by instances of $\wedge_-$ and $\exists_-$.

Now let $E_1(m_1, s_1), \dots, E_k(m_k, s_k)$ be the instances of $\wedge_-$ and $\exists_-$ in M' and S below N in the right order, where E_1 should be bottommost. Then $E_1(m_1, E_2(m_2, \dots E_k(m_k, N)))^{P\vec{r}}$ has at most the same free variables as $\mathtt{Stab}_{P,\vec{x}}\vec{r}(S(M(uN)))$, since every variable that was bound before is still bound. Furthermore, since we kept the order, $E_1(m_1, E_2(m_2, \dots E_k(m_k, N)))^{P\vec{r}}$ remains normal. Hence, we can replace $\mathtt{Stab}_{P,\vec{x}}\vec{r}(S(M(uN)))^{P\vec{r}}$ by $E_1(m_1, E_2(m_2, \dots E_k(m_k, N)))^{P\vec{r}}$.

Repeating this will give us a proof in minimal logic.

We now want to create an algorithm doing what is described in this proof. In this algorithm, we will iterate through the proof tree in depth-first way.

1. If $t = u$ is an assumption or axiom, return u.
2. If $t = \vee_+ a$, $t = \exists_+ a$ or $t = (\lambda_x a)^{\forall_x A}$ recur with a obtaining b, and return $\vee_+ b$, $\exists_+ b$ or $\lambda_x b$.
3. If $t = \langle a, b \rangle$, recur with a and b obtaining c and d and return $\langle c, d \rangle$.
4. If $t = a(b^{\exists_x A}.d)$ or $t = a(b^A, c^B.d)$ recur with a and d obtaining e and f, and return $e(f)$.
5. If $t = a^{\forall_x A} r$ recur with a obtaining b and return br.
6. If $t = a^{A \to B} b$ where $A \neq \neg\neg B$, recur with a and b obtaining c and d and return cd.
7. If $t = a^{\neg\neg B \to B} b^{\neg\neg B}$ recur with b obtaining c (to make sure we are at a top-most stability application). We now know that c is of the form $S(\lambda_u M(uN))$, where S

is a segment, thus, we can easily skip all $\wedge_-$ and $\exists_-$ instances on c to obtain $\lambda_u M(uN)$ (notice that we have not calculated N yet).

- If $u \notin FA(M(uN))$ we are not on a proper instance, thus, return αc (this actually cannot happen).
- Otherwise, iterate through $M(uN)$ to find uN: Check for applications of u with an argument not containing an application of u to get a top-most such application.

We now know S, M and N. Iterate through S and M and collect all the elimination axioms, concatenate them obtaining M'. Then return $S(M(N))$.

Except for 7, the algorithm just goes through the proof tree and is hence linear. In 7 we have to iterate through the minor premise a few times: Once to skip S, then to check whether $u \in FA(M(uN))$, then to find the proper uN, and then to collect the elimination axioms from S and M. All of this can be done in linear time.

Thus, the algorithm is quadratic relative to the length of the proof tree: for every tree node we have to iterate less than the whole tree a few more times. □

The statement of this theorem, though not the algorithm, can be concluded from Corollary 3.5 and Lemma 3.7:

From Corollary 3.5 immediately follows

Lemma 4.6. *Assume $\Gamma \vdash_{ec} A$, where the sequent $\Gamma \Rightarrow A$*

1. *does not contain positive occurrences of $\to$.*
2. *does not contain negative occurrences of $\vee$.*

Then $\Gamma \vdash_i A$.

Proof This clearly follows from Corollary 3.5: From $\Gamma \vdash_{ec} A$ trivially follows $\Gamma \vdash_{tc} A$, as in the latter case we may use more axioms, and from $\Gamma \vdash_{tc} A$, the rest follows by Corollary 3.5. □

Still, we only get an intuitionistic proof from Lemma 4.6, that is, a proof that may contain ex falso quodlibet. On the other hand, Theorem 4.5 also forbids positive occurrences of $\perp$ in Γ, so by Lemma 3.7, this proof will be minimal. We proved Theorem 4.5 using Corollary 3.5 and Lemma 3.7.

5 Examples and Limitations

We want to show that the case $\{\to^+, \neg^+, \vee^-\}$ of Theorem 3.1 and Theorem 4.5 are optimal, in the sense that if we drop any of their premises, the theorems do not hold anymore. We therefore give counterexamples satisfying all but one of these premises, for which the theorems do not hold.

Let us go through the premises of the case $\{\to^+, \neg^+, \lor^-\}$ of Theorem 3.1 firstly.

No positive negation: Dropping $\neg^+$, the sequent $\vdash_{tc} A \lor \neg A$, which proves tertium non datur, lies in our class. We already gave a classical proof in long normal form in Section 3.2. It is intuitionistically equivalent to stability, as the following proof shows, and can therefore not be derivable intuitionistically.

$$\dfrac{\dfrac{A \lor \neg A \qquad [A] \qquad \dfrac{\bot \to A \qquad \dfrac{[u : \neg\neg A] \qquad [\neg A]}{\bot}}{A}}{A}}{\neg\neg A \to A}\, u$$

No positive implication: Similarly, $\vdash_{tc} A \lor (A \to B)$ proves that the $\to^+$ cannot be dropped.

No negative disjunction: As an example that we cannot omit $\lor^-$, in (7) the counterexample $\forall_x \forall_y . R(a, x) \lor R(b, y) \vdash_{tc} \exists_y \forall_x R(y, x)$, is given, which contains negative disjunctions. A classical proof is:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{u : \neg\exists_y \forall_x R(y,x) \qquad \dfrac{[g : \forall_x R(y,x)]}{\exists_y \forall_x R(y,x)}}{\bot}}{\neg \forall_x R(y,x)}\, g \qquad \dfrac{\dfrac{\dfrac{\dfrac{[e : \neg\exists_x \neg R(y,x)] \qquad \dfrac{\dfrac{\dfrac{[f : \neg R(y,x)] \; [f' : R(y,x)]}{\bot}}{\neg R(y,x)}\, f'}{\exists_x \neg R(y,x)}}{\bot}}{\neg\neg R(y,x)}\, f}{R(y,x)}\, \text{stab (**)}}{\forall_x R(y,x)}}{\bot}}{\neg\neg \exists_x \neg R(y,x)}\, e}{\exists_x \neg R(y,x)}\, \text{stab (***)}}{\forall_y \exists_x \neg R(y,x)}$$

(*)

\begin{prooftree}
\AxiomC{(*)}
\noLine
\UnaryInfC{$\forall_y\exists_x\neg R(y,x)$}
\UnaryInfC{$\exists_x\neg R(b,x)$}
\AxiomC{(*)}
\UnaryInfC{$\forall_y\exists_x\neg R(y,x)$}
\UnaryInfC{$\exists_x\neg R(a,x)$}
\AxiomC{$\forall_x\forall_y.R(a,x)\lor R(b,y)$}
\UnaryInfC{$R(a,x)\lor R(b,x)$}
\AxiomC{$\vdots$}
\noLine
\UnaryInfC{$\bot$}
\AxiomC{$[R(b,x)]$}
\AxiomC{$[\neg R(b,x)]$}
\BinaryInfC{$\bot$}
\TrinaryInfC{$\bot$}
\BinaryInfC{$\bot$}
\BinaryInfC{$\bot$}
\RightLabel{u}
\UnaryInfC{$\neg\neg\exists_y\forall_x R(y,x)$}
\RightLabel{stab}
\UnaryInfC{$\exists_y\forall_x R(y,x)$}
\end{prooftree}

We have shown that $\{\to^+, \lor^-, \neg^+\}$ is indeed necessary for Orevkov's theorem.

Now let us go through the premises of Theorem 4.5.

No positive implication: The previous counterexample for the necessity of $\to^+$ in Orevkov's theorem is not applicable here, as it needs traditional classical logic. However, the following example can also be used as a counterexample for the previous case. It is a (partial) proof of the *Peirce formula*, which is known not to be intuitionistically derivable. We have in fact $(P\to Q)\to P\vdash_{ec} P$, which is provable by

\begin{prooftree}
\AxiomC{$[u:\neg P]$}
\AxiomC{$(P\to Q)\to P$}
\AxiomC{$[u:\neg P]$}
\AxiomC{$[v:P]$}
\BinaryInfC{$\bot$}
\RightLabel{$[w:\neg Q]$}
\UnaryInfC{$\neg\neg Q$}
\RightLabel{stab}
\UnaryInfC{Q}
\RightLabel{v}
\UnaryInfC{$P\to Q$}
\BinaryInfC{P}
\BinaryInfC{$\bot$}
\RightLabel{u}
\UnaryInfC{$\neg\neg P$}
\RightLabel{stab}
\UnaryInfC{P}
\end{prooftree}

Removing the top most instance of stability can be done as stated above and gives us

\begin{prooftree}
\AxiomC{$[u:\neg P]$}
\AxiomC{$(P\to Q)\to P$}
\AxiomC{$\bot\to Q$}
\AxiomC{$[u:\neg P]$}
\AxiomC{$[v:P]$}
\BinaryInfC{$\bot$}
\BinaryInfC{Q}
\RightLabel{v}
\UnaryInfC{$P\to Q$}
\BinaryInfC{P}
\BinaryInfC{$\bot$}
\RightLabel{u}
\UnaryInfC{$\neg\neg P$}
\RightLabel{stab}
\UnaryInfC{P}
\end{prooftree}

The lower stability however cannot be removed this way: v is bound by an implication introduction, and we cannot bind it in another way.

No negative falsum: Theorem 4.5 does not allow negative occurrences of $\bot$, and there is a trivial example showing that this is necessary: $\neg P, P \vdash_{ec} Q$, which is provable by

$$\dfrac{\dfrac{\dfrac{\neg P \qquad P}{\bot}}{\neg\neg Q}\,[w:\neg Q]}{Q}\,\text{stab}$$

We know that since it implies ex falso quodlibet, it cannot be derivable in minimal logic. Trying to apply the algorithm here fails because w does not occur at all, and thus no derivation of Q, but there is an abstraction of it. (This is the nasty side case in Section 4.3 where we proved that M' actually has to contain u freely.)

No negative disjunction: Theorem 4.5 does not allow negative disjunctions. Assuming we drop this premise, then let Q be atomic and not contain x freely, and consider $(\forall_x Px) \to Q, \forall_x (Px \vee Q) \vdash_{ec} Q$, an example given in (10). It is provable by

$$\dfrac{\dfrac{[u:\neg Q] \qquad \dfrac{(\forall_x Px)\to Q \qquad \dfrac{\dfrac{\dfrac{\forall_x .Px \vee Q}{Px \vee Q} \qquad [w':Px] \qquad \dfrac{\dfrac{\dfrac{[u:\neg Q] \qquad [w:Q]}{\bot}}{\neg\neg Px}\,[\neg Px]}{Px}\,\text{stab}}{Px}\,w',w}{\forall_x Px}}{Q}}{\dfrac{\bot}{\neg\neg Q}\,u}}{Q}\,\text{stab}$$

but it is known not to be provable in minimal logic. Trying to apply the algorithm firstly removes the upper application of stability, and replaces it by $w : Q$. But then we are in the situation of

$$\dfrac{Px \vee Q \qquad [w':Px] \qquad [w:Q]}{Px}$$

which is not a correct proof anymore.

We have shown that all premises of Theorem 4.5 are indeed necessary. Now, one could think about whether the given algorithm also works for traditional classical logic. To show that it does not in general, we give an additional example:

Extended classical logic is necessary: The sequent $\forall_x R(a,x) \vdash_{tc} \exists_y \forall_x R(y,x)$, which is similar to our $\vee^-$-example for Theorem 3.1, shows that the algorithm in Section 4.3 fails if the given proof is in traditional classical logic rather than extended classical logic. It is trivial to prove this sequent, even in minimal logic. However, the following proof uses general stability.

$$
\begin{prooftree}
\AxiomC{$[u : \neg\exists_y \forall_x R(y,x)]$}
\AxiomC{$[g : \forall_x R(a,x)]$}
\UnaryInfC{$R(a,x)$}
\UnaryInfC{$\forall_x R(a,x)$}
\UnaryInfC{$\exists_y \forall_x R(y,x)$}
\BinaryInfC{$\bot$}
\RightLabel{g}
\UnaryInfC{$\neg\forall_x R(a,x)$}
\AxiomC{$[e : \neg\exists_x \neg R(a,x)]$}
\AxiomC{$[f : \neg R(a,x)]$}
\AxiomC{$[R(a,x)]$}
\BinaryInfC{$\bot$}
\UnaryInfC{$\neg R(a,x)$}
\UnaryInfC{$\exists_x \neg R(a,x)$}
\BinaryInfC{$\bot$}
\RightLabel{f}
\UnaryInfC{$\neg\neg R(a,x)$}
\RightLabel{stab(**)}
\UnaryInfC{$R(a,x)$}
\UnaryInfC{$\forall_x R(a,x)$}
\BinaryInfC{$\bot$}
\RightLabel{e}
\UnaryInfC{$\neg\neg\exists_x \neg R(a,x)$}
\RightLabel{stab(***)}
\UnaryInfC{$\exists_x \neg R(a,x)$}
\AxiomC{$\forall_y R(a,y)$}
\UnaryInfC{$R(a,x)$}
\AxiomC{$[h : \neg R(a,x)]$}
\BinaryInfC{$\bot$}
\RightLabel{h}
\BinaryInfC{$\bot$}
\RightLabel{u}
\UnaryInfC{$\neg\neg\exists_y \forall_x R(y,x)$}
\RightLabel{stab}
\UnaryInfC{$\exists_y \forall_x R(y,x)$}
\end{prooftree}
$$

The algorithm from Section 4.3 will fail here when eliminating the stability marked with (**), because Lemma 4.4 does not hold in this case: The assumption e contains a negative implication, but it disappears in the scope of the stability application marked with (***). Therefore, it is necessary for this algorithm that only atomic stability is used.

6 Possible Applications

In (1) an example, a weakening of a theorem by Heitmann, is given. Some effort is taken to bring this theorem into a form that looks a little more like formulae to which Barr's theorem can be applied. However, with Orevkov's theorem, except for "cosmetical" reasons, there is no need to do that.

The common notion of a ring is well known, and we only give a more model theoretic definition of it, as it fits better into our setting. (Actually, we define a commutative ring with unity.)

Definition A *ring* is a model R with equality $= \subset R \times R$, two operations $+ : R \times R \rightarrow R$, $\cdot : R \times R \rightarrow R$ and two distinguished elements $0, 1 \in R$, of the following theory:

- $\forall_x \forall_y . x + y = y + x$, $\forall_x \forall_y . x \cdot y = y \cdot x$.
- $\forall_x . x + 0 = x$, $\forall_x . x \cdot 1 = x$.
- $\forall_x \forall_y \forall_z . x + (y + z) = (x + y) + z$, $\forall_x \forall_y \forall_z . x \cdot (y \cdot z) = (x \cdot y) \cdot z$.
- $\forall_x \forall_y \forall_z . x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

We fix the convention that $\cdot$ binds stronger than $+$, and that we may write xy instead of $x \cdot y$.

Notice that all the axioms may be contained in the antecedent of a sequent belonging to the complete Glivenko class $\{\rightarrow^+, \neg^+, \vee^-\}$.

A common problem when applying formal methods to algebra is the fact that, maybe due to the reason that algebra was mainly developed before the rise of symbolic logic, it makes excessive use of higher-order objects, such as ideals, and in addition to that, often in a most-possible unconstructive way. Therefore, to make it accessible for constructive mathematics, a lot of effort was put into avoiding these higher order objects.

In our case, we will not define ring ideals as a subsets of a ring, but as unary relations on rings, which satisfy certain conditions, and instead of defining quotients of rings and ideals as sets of equivalence classes, we define them by the same model as the original ring except for another interpretation of the equality relation, which can be expressed in the original ring.

Definition An *ideal* of a ring R is an unary relation $\mathcal{A} \subseteq R$ such that

- $\forall_b \forall_r . \mathcal{A} r \rightarrow \mathcal{A}(br)$.
- $\forall_b . \forall_c . \mathcal{A} b \rightarrow \mathcal{A} c \rightarrow \mathcal{A}(b + c)$.

For a given ideal $\mathcal{A}$, the *quotient* $R/\mathcal{A}$, defined by changing the interpretation of $b = c$ to the relation $= /\mathcal{A}$ we get from the interpretation of $\exists_a . \mathcal{A} a \wedge b = c + a$, models the ring axioms.

Classically, the Jacobson radical can be defined as the intersection of all maximal ideals, where a maximal ideal is an ideal which is not a proper subset of a non trivial ideal. However, this is not useful here, as it involves higher order objects.

Definition Let R be a ring. Then we define that an element $b \in R$ is *invertible*, writing $b \in R^*$, and that an element $a \in R$ is *in the Jacobson radical* J_R, writing $a \in J$, and that an element z is in the *boundary* of a, writing $z \in J_a$, by

- $b \in R^* :\Leftrightarrow \exists_x . bx = 1$.
- $a \in J_R :\Leftrightarrow \forall_x . (1 - ax) \in R^*$.
- $z \in J_a :\Leftrightarrow \exists_x \exists_y . z = xa + y \wedge ax \in J_R$.

Notice that all these relations may belong to the antecedent of a sequent belonging to the complete Glivenko class $\{\rightarrow^+, \neg^+, \vee^-\}$.

Definition We inductively define the relation $\mathit{Heit}(R) < n$, saying the *Heitmann dimension* of R is smaller than n, by:

- $\mathit{Heit}(R) < 0 :\Leftrightarrow 1 = 0$.
- $\mathit{Heit}(R) < n + 1 :\Leftrightarrow \forall_a . \mathit{Heit}(R/J_a) < n$.

The problem with this definition is that it is not intrinsic. For $n = 0$, we have $\mathit{Heit}(R) < 1 :\Leftrightarrow \forall_a . \mathit{Heit}(R/J_a) < 0$, which means $\forall_a . (\mathit{Heit}(R/J_a) \models 1 = 0)$ and thus, $\forall_a . \exists_z . z \in J_a \wedge 1 = z$, which is equivalent to $\forall_a . 1 \in J_a$. Unwinding this definition, we get $\forall a . \exists_x \exists_y . 1 = xa + y \wedge ax \in J_R$, which becomes $\forall_a . \exists_x \exists_y . 1 = xa + y \wedge \forall_z . (1 - zax) \in R^*$, which then becomes $\forall_a . \exists_x \exists_y . 1 = xa + y \wedge \forall_z . \exists_b . b(1 - zax) = 1$.

For $n = 1$ we have $\mathit{Heit}(R) < 2 :\Leftrightarrow \forall a . \mathit{Heit}(R/J_a) < 1$ which we can express via $\forall a' . ((R/J_{a'}) \models \forall_a . \exists_x \exists_y . 1 = xa + y \wedge \forall_z . \exists_b . b(1 - zax) = 1)$ which can be unwound into $\forall a' . \forall_a . \exists_x \exists_y . (\exists_t \in J_{a'} . 1 = t + xa + y) \wedge \forall_z . \exists_b . (\exists_t \in J_{a'} . b(1 - zax) = t + 1)$, which then can be unwound again into something that is obviously $\rightarrow$ and $\vee$ free.

In general, one sees that proceeding to arbitrary n cannot produce a formula that contains $\rightarrow$ or $\vee$, hence, sequents containing these formulae in their antecedent cannot be outside of $\{\rightarrow^+, \neg^+, \vee^-\}$ because of these formulae.

Definition An $n \times n$-matrix over a ring R can be defined as a function $\{0, \ldots, n-1\} \rightarrow \{0, \ldots, n-1\} \rightarrow R$. The set of all such matrices shall be denoted by $R^{n \times n}$. For every $n \times n$-matrix a, we can define the x, y-submatrix function by $\tilde{a} := \lambda x \lambda y \lambda p \lambda q . a(\mathit{if}(p < x)(p)(1 + p))(\mathit{if}(q < y)(q)(1 + q))$, and therefore the set of $n - 1$-submatrices by $\{a\}_{n-1} := \{\tilde{a}xy | 0 \leq x, y \leq n\}$, and the set of m-submatrices inductively as the set of submatrices of the $(m + 1)$-submatrices. The *determinant* $[a]$ of a matrix a can be defined as usual, for example recursively through the Laplace expansion $[i] = i$ for $i \in R = R^{1 \times 1}$, $[a] = \sum\limits_{0 \leq i, j < n} (aij) \cdot (-1)^{i+j} \cdot [\tilde{a}ij]$ for $(n \times n)$-matrices, $n > 1$.

For a given matrix $F \in R^{k \times k}$ and a given natural $n < k$, let $\{a_1, \ldots, a_m\} = \{[x] | x \in \{F\}_n\}$, and define the ideal $\Delta_n(F)$ by

- $\Delta_n(F)r :\Leftrightarrow \exists_{b_1} \dots \exists b_m.r = \sum_{1\leq i\leq m} b_i a_i$.

Obviously, $\Delta_n(F)r$ is expressible in first order without $\vee$ and $\rightarrow$.

Theorem 6.1. $\text{Heit}(R) < n \wedge \forall_r \Delta_n(F)r \rightarrow \exists_{X,Y}.1 = XFY$, *where* X *is a row vector,* Y *is a column vector, and* XFY *is defined in the usual way.*

This theorem can be classically proved. On the other hand, $\mathit{Heit}(R) < n \wedge \forall_r \Delta_n(F)r$ is expressible without $\rightarrow$ and $\wedge$, and the same holds for $\exists_{X,Y}.1 = XFY$, obviously, since this can be expressed as the existence of elements of R satisfying the conjunction of some equations. Denote by $\mathfrak{R}$ the ring axioms, then
$\mathfrak{R}, \mathit{Heit}(R) < n, \forall r.\Delta_n(F)r \Rightarrow \exists_{X,Y}.1 = XFY$ is therefore in the complete Glivenko class $\{\rightarrow^+, \neg^+, \vee^-\}$, and we can apply Orevkov's theorem.

References

[1] T. Coquand and H. Lombardi. A logical approach to abstract algebra. *Math. Struct. in Comp. Science*, 16:885–900, 2006.

[2] S. C. Kleene. *Introduction to metamathematics*. Van Nostrand, Princeton, N. J., 1952.

[3] S. C. Kleene. "Permutability of inferences in Gentzen's calculi LK and LJ." Two papers on the predicate calculus. *Memoirs of the American Mathematical Society*, 10:1–26, 1952.

[4] G. Nadathur. Correspondences between classical, intuitionistic and uniform provability. *Theoretical Computer Science*, 232(1-2):273–298, 2000.

[5] S. Negri. Contraction-free sequent calculi for geometric theories with an application to barr's theorem. *Arch. Math. Log.*, 42(4):389–401, 2003.

[6] V. P. Orevkov, editor. *Logical and logico-mathematical calculi. I*. English translation, The calculi of symbolic logic. I, Proceedings of the Steklov Institute of Mathematics, vol. 98 (1971). Trudy Matematicheskogo Instituta imeni V.A. Steklova, 1968.

[7] V. P. Orevkov. On Glivenko sequent classes. In *Logical and logico-mathematical calculi* (6), pages 131–154 (Russian), 147–173 (English).

[8] E. Palmgren. An intuitionistic axiomatisation of real closed fields. *Mathematical Logic Quarterly*, 48(2):297–299, 2002.

[9] H. Schwichtenberg. Minimal from classical proofs. In E. Börger, G. Jäger, H. Kleine-Büning, and M. Richter, editors, *Computer Science Logic*, volume 262 of *LNCS*, pages 326–328. Springer-Verlag, Berlin, Heidelberg, New York, 1992.

[10] H. Schwichtenberg and C. Senjak. Minimal from classical proofs. submitted, 2011.

[11] H. Schwichtenberg and S. S. Wainer. Proofs and computations. to appear 2011.

[12] A. S. Troelstra and H. Schwichtenberg. *Basic Proof Theory*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1996.