MARTIN FISCHER

Sequent Calculi for the Propositional Logic of HYPE

Abstract. In this paper we discuss sequent calculi for the propositional fragment of the logic of HYPE. The logic of HYPE was recently suggested by Leitgeb (*Journal of Philosophical Logic* 48:305–405, 2019) as a logic for hyperintensional contexts. On the one hand we introduce a simple **G1**-system employing rules of contraposition. On the other hand we present a **G3**-system with an admissible rule of contraposition. Both systems are equivalent as well as sound and complete proof-system of HYPE. In order to provide a cut-elimination procedure, we expand the calculus by connections as introduced in Kashima and Shimura (*Mathematical Logic Quarterly* 40:153–172, 1994).

Keywords: Conditionals, Negation, Partial, Cut-elimination, Connections.

1. Introduction

Leitgeb [4] proposed a logic for hyperintensional contexts that is called HYPE.[1] In his paper Leitgeb worked with an axiomatic calculus. In this paper, we are going to develop sequent calculi for the propositional fragment of HYPE.

The direct motivation for this paper is the presentation of sequent systems that have some desirable proof-theoretic features. Most prominently, it would be desirable to have a calculus that allows for cut-elimination and with it to be able to establish results, such as the conservativity over intuitionistic logic, in a proof-theoretic way. An indirect motivation is that these calculi are able to function as the basis for further proof-theoretic investigations of theories formulated in the logic of HYPE, such as theories of truth.

Before going into the details of the presentation, we informally describe the logic of HYPE. The logic of HYPE can be understood as an extension of intuitionistic logic, IL, by an additional negation or as an extension of First-Degree-Entailment, FDE, by an intuitionistic conditional. In

---

[1] According to Odintsov and Wansing [7] the work of Moisil [5] constitutes a predecessor of the propositional fragment of HYPE.

both cases it is a restriction of classical logic. If we look at the first way of characterizing HYPE the analogy to intuitionistic logic is given mainly by the primitive conditional of HYPE that behaves very similar to the intuitionistic conditional. From this perspective the innovation lies in adding a negation that does not behave intuitionistically, but rather close to the behaviour of an FDE negation. We have the de Morgan laws and double negation elimination, but not the full classical rules. This is in line with a semantic interpretation of formulas by truth-value gaps or gluts. In the semantics therefore we consider states that can have a partial interpretation or an overdetermined interpretation. If we look at it from the second perspective, we have a nontrivial extension of FDE by a suitable conditional, i.e. an intuitionistic conditional that allows for modus ponens and the deduction theorem. The conditional can therefore be interpreted as an object linguistic correspondent of the sequent arrow, that can additionally be iterated.

In the following we address several questions in turn. In the second section we introduce the propositional fragment of the logic of HYPE. We start by repeating the axiomatic calculus as presented in Leitgeb [4] and simplified in Speranski [9]. The main focus however, will be on sequent calculi. We start by looking at $\mathbf{G1h_p}$, a system employing rules of contraposition. The main advantage of this system is its simplicity. Then we move to a system in which not only the structural rules of weakening and contraction are admissible, but also the contraposition rule. This system will be called $\mathbf{G3h_p}$. The most innovative aspect is the treatment of negated conditionals. We establish the connection of both systems to the axiomatic calculus, and with it the soundness and completeness of $\mathbf{G1h_p}$ and $\mathbf{G3h_p}$.

The third section focusses on cut-elimination. We first present problems for the standard strategy of $\mathbf{G3}$-systems in the case of our introduction rules for the conditional and present a counterexample for cut-elimination in $\mathbf{G3h_p}$. Then we consider an extension $\mathbf{cGh_p^+}$ of the system $\mathbf{G1h_p}$ by connections and more liberal rules. This is based on the solution for cut-elimination in the case of constant domains in IL as suggested by Kashima and Shimura in [3]. We establish a cut-elimination theorem for $\mathbf{cGh_p^+}$.

In Section 4 we establish the equivalence of $\mathbf{cGh_p^+}$ with our previous systems.

In the last section we show how the cut-elimination can be used to establish the conservativity of HYPE over IL as well as FDE.

2. The Propositional Logic of HYPE

We consider a propositional language based on the following primitive logical signs $\neg, \vee, \rightarrow, \perp$.[2]

Since we have the usual de Morgan laws we can make use of the following abbreviations: $\top$ is short for $\neg\perp$, $A \wedge B$ is short for $\neg(\neg A \vee \neg B)$ and $A \leftrightarrow B$ for $(A \rightarrow B) \wedge (B \rightarrow A)$.[3] The notion of a well-formed formula is standard.

Before we present the sequent systems we repeat the axiomatic system of HYPE. We make use of the propositional part $\mathbf{N_p^\bullet}$ of the system $\mathsf{QN^\bullet}$ due to Speranski in [9], which is a simplification of Leitgeb's system in [4] on p.336f. Speranski [9] shows in detail how to derive in $\mathbf{N_p^\bullet}$ the missing principles stated in Leitgeb [4]. The propositional fragment $\mathbf{N_p^\bullet}$ includes the intuitionistic axioms of **Hi**

$$A \rightarrow (B \rightarrow A) \qquad (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C)) \tag{1}$$

$$A \wedge B \rightarrow A \qquad A \wedge B \rightarrow B \tag{2}$$

$$A \rightarrow A \vee B \qquad B \rightarrow A \vee B \tag{3}$$

$$A \rightarrow (B \rightarrow A \wedge B) \qquad (A \rightarrow C) \rightarrow ((B \rightarrow C) \rightarrow (A \vee B \rightarrow C)) \tag{4}$$

$$\perp \rightarrow A \tag{5}$$

and the axioms for double negation:

$$A \rightarrow \neg\neg A \qquad\qquad \neg\neg A \rightarrow A \tag{6}$$

The system $\mathbf{N_p^\bullet}$ is closed under the rule of conditional contraposition:

$$\frac{\vdash_{\mathbf{N_p^\bullet}} A \rightarrow B}{\vdash_{\mathbf{N_p^\bullet}} \neg B \rightarrow \neg A}$$

and under Modus ponens: if $\Gamma \vdash_{\mathbf{N_p^\bullet}} A$ and $\Gamma \vdash_{\mathbf{N_p^\bullet}} A \rightarrow B$, then $\Gamma \vdash_{\mathbf{N_p^\bullet}} B$.

$\mathbf{N_p^\bullet}$ is a neater presentation of HYPE where a few redundant principles are dropped. The consequences of the two systems are identical.

---

[2]Here we deviate slightly from Leitgeb [4]. Officially in Leitgeb $\perp$ is defined as $\neg\top$ and the intutionistic negation is then recovered as $A \rightarrow \perp$.

[3]Although we do neither have $A \rightarrow B$ iff $\neg A \vee B$ nor $\neg A$ iff $A \rightarrow \perp$, it is possible to introduce the abbreviations $A \supset B$ for $\neg A \vee B$, a classical conditional, and $\sim A$ for $A \rightarrow \perp$, an intuitionistic negation.

2.1. G1h$_\mathbf{p}$

We present multiconclusion systems based on a multiconclusion calculus for intuitionistic logic.[4] Sequents are understood as (ordered pairs of) multisets. The system **G1h$_\mathbf{p}$** consists of the following initial sequents and rules, where $\neg\Gamma$ is short for the (multi-)set of all the negations of formulas in Γ:

$$(\texttt{ID}_p) \qquad A \Rightarrow A \qquad\qquad (\texttt{L}\bot) \qquad \bot \Rightarrow$$

$$(\texttt{Cut}) \ \frac{\Gamma \Rightarrow \Delta, A \qquad A, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}$$

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

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

$$(\texttt{L}\vee) \ \frac{A, \Gamma \Rightarrow \Delta \qquad B, \Gamma \Rightarrow \Delta}{A \vee B, \Gamma \Rightarrow \Delta} \qquad (\texttt{R}\vee) \ \frac{\Gamma \Rightarrow A, B, \Delta}{\Gamma \Rightarrow A \vee B, \Delta}$$

$$(\texttt{L}\rightarrow) \ \frac{\Gamma \Rightarrow \Delta, A \qquad B, \Gamma \Rightarrow \Delta}{A \rightarrow B, \Gamma \Rightarrow \Delta} \qquad (\texttt{R}\rightarrow) \ \frac{\Gamma, A \Rightarrow B}{\Gamma \Rightarrow A \rightarrow B}$$

$$(\texttt{ConCp}) \ \frac{\Gamma \Rightarrow \neg\Delta}{\Delta \Rightarrow \neg\Gamma} \qquad\qquad (\texttt{ClCp}) \ \frac{\neg\Gamma \Rightarrow \Delta}{\neg\Delta \Rightarrow \Gamma}$$

In the presence of weakening and contraction the choice of a context-sharing version of (\texttt{Cut}) is not of greater significance. The next lemma collects some basic facts about **G1h$_\mathbf{p}$**. They mostly concern the admissibility of some basic inferences in **G1h$_\mathbf{p}$**.

LEMMA 1.

1. *The sequents* $\Rightarrow \top$, $A \Rightarrow \neg\neg A$, $\neg\neg A \Rightarrow A$, *are derivable in* **G1h$_\mathbf{p}$**.

2. *The rule of contraposition*

$$(\texttt{Cp}) \ \frac{\Gamma \Rightarrow \Delta}{\neg\Delta \Rightarrow \neg\Gamma}$$

   is admissible in **G1h$_\mathbf{p}$**.

---

[4]This goes back to Maehara's version used in Takeuti [10] p. 52f and Dragalin's system used in Negri and Plato [6] p. 108f.

3. *The following rules are admissible in* **G1h$_\mathbf{p}$**:

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

4. *Intersubstitutivity: If* $C \Rightarrow C'$ *and* $C' \Rightarrow C$, *as well as* $\Gamma \Rightarrow \Delta$ *are derivable in* **G1h$_\mathbf{p}$**, *then* $\Gamma(C'/C) \Rightarrow \Delta(C'/C)$ *is derivable, where* $\Gamma(C'/C) \Rightarrow \Delta(C'/C)$ *is obtained by replacing all occurrences of* C *in* Γ *and* Δ *by* C'.

PROOF. For 1: $\Rightarrow \top$ is derived by (L$\bot$), (ConCp) and the definition of $\top$. From $\neg A \Rightarrow \neg A$ the sequent $A \Rightarrow \neg\neg A$ is derived by (ConCp) and $\neg\neg A \Rightarrow A$ by (ClCp).

For 2: By a series of cuts we get from $\Gamma \Rightarrow \Delta$ and $A \Rightarrow \neg\neg A$ for all $A \in \Delta$ to $\Gamma \Rightarrow \neg\neg\Delta$. Using (ConCP) gives us $\neg\Delta \Rightarrow \neg\Gamma$.

For 3: Starting with $A, B, \Gamma \Rightarrow \Delta$, we use (Cp) and then (R$\vee$). Then (Cp) again and a series of cuts to eliminate the double negations. The argument for the rule (R$\wedge$) is analogous.

For 4: By induction on the length of the derivation. If C is replaced in the context then there is nothing special to consider. Also within the principal formulas the replacement is unproblematic. ∎

The next observation is that HYPE is an extension of intuitionistic logic.

LEMMA 2. *All the axioms* (1)–(5) *of the axiomatic system* **Hi** *of propositional intuitionistic logic are derivable in* **G1h$_\mathbf{p}$**.

Combining Lemma 1 and Lemma 2 establishes that the axioms of **N$_\mathbf{p}^\bullet$** are derivable in **G1h$_\mathbf{p}$**. That it is also a subsystem will be established in Lemma 11 in combination with Theorem 1. The first-order extension QN$^\bullet$ of **N$_\mathbf{p}^\bullet$** is discussed in Speranski [9]. QN$^\bullet$ is an extension of Došen's logic **N** by the principles of double negation $\neg\neg A \to A$ and $A \to \neg\neg A$. Speranski [9] establishes the equivalence of QN$^\bullet$ and Leitgeb's first-order axiomatic system of HYPE as presented in [4].[5]

2.2. G3h$_\mathbf{p}$

We are now going to present a system in which the structural rules of weakening and contraction are absorbed.

In order to achieve this we will employ an alternative inductive definition of well-formed formula, similar to a Tait language. We start with the literals

[5]Corollary 7.4. p. 22 in [9].

and then close it under the application of $\vee, \rightarrow$ as well as $\neg\vee, \neg\rightarrow, \neg\neg$. Given a set Φ of propositional variables $\mathsf{LIT}_\Phi := \Phi \cup \{\neg p \,|\, p \in \Phi\}$. We use v as a metavariable for literals and $\bar{v}$ as the contrapositive of v, i.e. if v is a propositional variable p, then $\bar{v}$ is $\neg p$, and if v is a negated propositional variable $\neg p$, then $\bar{v}$ is p.

The system $\mathbf{G3h_p}$ consists of the following initial sequents and rules that are based on the positive build-up of formulas:

For v a literal:

$$(\text{ID}) \qquad v, \Gamma \Rightarrow \Delta, v$$

$$(\text{L}\bot) \qquad \bot, \Gamma \Rightarrow \Delta \qquad\qquad\qquad (\text{R}\neg\bot) \qquad \Gamma \Rightarrow \Delta, \neg\bot$$

$$(\text{Cut}) \ \frac{\Gamma \Rightarrow \Delta, A \qquad A, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}$$

PROPOSITIONAL

$$(\text{L}\neg\neg) \ \frac{A, \Gamma \Rightarrow \Delta}{\neg\neg A, \Gamma \Rightarrow \Delta} \qquad\qquad\qquad (\text{R}\ \neg\neg) \ \frac{\Gamma \Rightarrow \Delta, A}{\Gamma \Rightarrow \Delta, \neg\neg A}$$

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

$$(\text{L}\rightarrow) \ \frac{A \rightarrow B, \Gamma \Rightarrow \Delta, A \qquad B, \Gamma \Rightarrow \Delta}{A \rightarrow B, \Gamma \Rightarrow \Delta} \qquad (\text{R}\rightarrow) \ \frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow \Delta, A \rightarrow B}$$

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

$$(\text{L}\neg\rightarrow) \ \frac{\neg B \Rightarrow \Delta, \neg A}{\neg(A \rightarrow B), \Gamma \Rightarrow \Delta} \qquad (\text{R}\neg\rightarrow) \ \frac{\Gamma \Rightarrow \Delta, \neg B \qquad \neg A, \Gamma \Rightarrow \Delta, \neg(A \rightarrow B)}{\Gamma \Rightarrow \Delta, \neg(A \rightarrow B)}$$

We have the following admissible rules for the defined symbols:

LEMMA 3. *The following rules are admissible in* $\mathbf{G3h_p}$*:*

$$(\text{R}\top) \qquad \Gamma \Rightarrow \Delta, \top \qquad\qquad\qquad (\text{L}\neg\top) \qquad \neg\top, \Gamma \Rightarrow \Delta$$

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

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

The most obvious differences to other calculi are the rules for the negated conditional. The main reason for their introduction is the aim of having the admissibility of contraposition.

It is easy to see that our system $\mathbf{G3h_p}$ is an extension of FDE in the version $\mathbf{LE_{fde2}}$ as presented in Anderson and Belnap [1] §17 (without their

$\rightarrow$). The main difference are the additional rules for $\rightarrow$ and $\neg \rightarrow$. Without the rules our system basically is $\mathbf{LE_{fde2}}$, neglecting the minor difference of sequences versus multisets. Without the cut rule the conservativity is obvious. We will establish the conservativity including the cut rule in Lemma 22.

2.3. Properties of G3h$_p$

Before we state some properties of the calculus we clarify the terminology:

The (positive) depth $|A|$ of a formula A is defined as:

$$|A| = 0 \text{ for } A \text{ a literal, } \bot \text{ or } \neg\bot; \qquad |\neg\neg A| = |A| + 1;$$
$$|A \circ B| = \mathsf{max}(|A|, |B|) + 1; \qquad |\neg(A \circ B)| = \mathsf{max}(|\neg A|, |\neg B|) + 1;$$

where $\circ$ is either $\vee$ or $\rightarrow$.

We also adjust the notion of a subformula accordingly:

DEFINITION 1.

(i) A is a subformula of A;

(ii) if $B \circ C$ is a subformula of A, then so are B, C;

(iii) if $\neg(B \circ C)$ is a subformula of A, then so are $\neg B, \neg C$;

(iv) if $\neg\neg B$ is a subformula of A, then so is B.

The depth $|\mathcal{D}|$ of a derivation $\mathcal{D}$ is the maximum length of a path in the prooftree, i.e. $|\mathcal{D}| := \mathsf{sup}_{i<n}(|(\mathcal{D}_i)| + 1)$ (where $\mathcal{D}_0, \ldots, \mathcal{D}_{n-1}$ are the immediate subderivations of $\mathcal{D}$).

We use $\vdash^n \Gamma \Rightarrow \Delta$ as short for: there exists a derivation $\mathcal{D}$ of $\Gamma \Rightarrow \Delta$ with $|\mathcal{D}| \leq n$.

We have the depth preserving admissibility of Weakening.

LEMMA 4. *(Weakening depth-preserving)* If $\vdash^n \Gamma \Rightarrow \Delta$, then $\vdash^n \Gamma, \Gamma' \Rightarrow \Delta, \Delta'$.

LEMMA 5. $A, \Gamma \Rightarrow \Delta, A$ for all A.

LEMMA 6. *(Inversion)*

(i) If $\vdash^n A \vee B, \Gamma \Rightarrow \Delta$, then $\vdash^n A, \Gamma \Rightarrow \Delta$ and $\vdash^n B, \Gamma \Rightarrow \Delta$;

(ii) if $\vdash^n \Gamma \Rightarrow \Delta, A \vee B$, then $\vdash^n \Gamma \Rightarrow \Delta, A, B$;

(iii) if $\vdash^n A \rightarrow B, \Gamma \Rightarrow \Delta$, then $\vdash^n B, \Gamma \Rightarrow \Delta$;

(iv) if $\vdash^n \Gamma \Rightarrow \Delta, \neg\neg A$, then $\vdash^n \Gamma \Rightarrow \Delta, A$;

(v) if $\vdash^n \neg\neg A, \Gamma \Rightarrow \Delta$, then $\vdash^n A, \Gamma \Rightarrow \Delta$;

(vi) if $\vdash^n \neg(A \vee B), \Gamma \Rightarrow \Delta$, then $\vdash^n \neg A, \neg B, \Gamma \Rightarrow \Delta$;

(vii) if $\vdash^n \Gamma \Rightarrow \Delta, \neg(A \vee B)$, then $\vdash^n \Gamma \Rightarrow \Delta, \neg A$ and $\vdash_n \Gamma \Rightarrow \Delta, \neg B$;

(viii) if $\vdash^n \Gamma \Rightarrow \Delta, \neg(A \rightarrow B)$, then $\vdash^n \Gamma \Rightarrow \Delta, \neg B$.

Remark: The rules $(\mathtt{R} \rightarrow)$ and $(\mathtt{L}\neg \rightarrow)$ are not invertible in contrast to the single conclusion calculus **G3i**.[6] Our system is therefore closer to the calculus **G3im** as presented in Negri and von Plato [6]. In the problematic case $A \rightarrow B$ is non-principal and the last inference is an application of $(\mathtt{L}\neg\rightarrow)$:

$$\frac{\neg D \Rightarrow \neg C, A \rightarrow B}{\neg(C \rightarrow D) \Rightarrow A \rightarrow B}$$

We cannot use the IH on the upper sequent, because it contains $\neg C$ in the context, which is not allowed. Even if we could employ the IH to the upper sequent we would get $\neg D, A \Rightarrow \neg C, B$ and then we cannot use $(\mathtt{L}\neg\rightarrow)$, because of the additional A in the context.

In our $\mathbf{G1_h}$ system as well as Leitgeb's axiomatic system there are (restricted) rules of contraposition. In our $\mathbf{G3_h}$ system the rule of contraposition is admissible. Although the admissibility is not height preserving, it is interesting to note that cut is not involved in the admissibility.

LEMMA 7. *[Admissibility of contraposition] If $\vdash^n \Gamma \Rightarrow \Delta$, then $\vdash \neg\Delta \Rightarrow \neg\Gamma$.*

PROOF. By induction on n. For $n = 0$ we have the case $(\mathtt{ID})$, which is obvious as it is stated for literals. The case $(\mathtt{L}\bot)$ is handled by $(\mathtt{R}\neg\bot)$. The case $(\mathtt{R}\neg\bot)$ by $(\mathtt{L}\bot)$ and $(\mathtt{L}\neg\neg)$.

For $n \neq 0$ we distinguish the positive and the negative cases:

In the positive cases we use the IH on the premises and then the corresponding negative rule, for example if the last inference is an application of $(\mathtt{L} \rightarrow)$, then by the IH applied to the upper sequents we have $\neg\Delta \Rightarrow \neg\Gamma, \neg B$ as well as $\neg A, \neg\Delta \Rightarrow \neg\Gamma, \neg(A \rightarrow B)$. Then we can apply $(\mathtt{R}\neg\rightarrow)$ to get the desired.

In the negative cases we additionally make use of the $\neg\neg$ rules, for example if the last inference was $(\mathtt{L}\neg\rightarrow)$, then by IH we have $\neg\neg A, \neg\Delta \Rightarrow \neg\neg B$. By inversion for $\neg\neg$ we have $A, \neg\Delta \Rightarrow B$. Using $(\mathtt{R} \rightarrow)$ we get $\neg\Delta \Rightarrow \neg\Gamma, A \rightarrow B$ and $(\mathtt{R}\neg\neg)$ gives the desired. ∎

LEMMA 8. *(Admissibility of contraction)*

(i) If $\vdash^n \Gamma \Rightarrow \Delta, A, A$, then $\vdash^n \Gamma \Rightarrow \Delta, A$;

---

[6]Compare Troelstra and Schwichtenberg [11], p. 79, Proposition 3.5.4. v.

(ii) if $\vdash^n A, A, \Gamma \Rightarrow \Delta$*, then* $\vdash^n A, \Gamma \Rightarrow \Delta$*.*

PROOF. By induction on n. If $n = 0$, then for any initial sequent $\Gamma \Rightarrow \Delta, A, A$, the sequent $\Gamma \Rightarrow \Delta, A$ is also an initial sequent.

For $n \neq 0$ we distinguish the cases where A is non-principal and cases where it is principal.

If A is non-principal, then we argue as usual by the IH, except for the cases with restricted contexts in the upper sequents, i.e. $(R \rightarrow)$ and $(L \neg \rightarrow)$. In these cases we get the desired directly by using the relevant rule.

If A is principal in an inference for which we have inversion, for example if the last inference is $(R \neg \rightarrow)$: $\vdash^n \Gamma \Rightarrow \Delta, \neg(A \rightarrow B), \neg B$ and $\vdash^n \neg A, \Gamma \Rightarrow \Delta, \neg(A \rightarrow B), \neg(A \rightarrow B)$. By inversion $\vdash^n \Gamma \Rightarrow \Delta, \neg B, \neg B$ and by IH $\vdash^n \Gamma \Rightarrow \Delta, \neg B$ and $\vdash^n \neg A, \Gamma \Rightarrow \Delta, \neg(A \rightarrow B)$. Using $(R \neg \rightarrow)$ we get the desired.

If A is principal in an inference for which we do not have inversion, such as $(R \rightarrow)$ and $(L \neg \rightarrow)$ we argue directly by a suitable application of the rule itself to the upper sequent. ∎

2.4. The Equivalence of $\mathbf{G1h_p}$ and $\mathbf{G3h_p}$

LEMMA 9. *The following rules are admissible in* $\mathbf{G1h_p}$*:*
$(L \neg \neg)$*,* $(R \neg \neg)$*,* $(L \neg \vee)$*,* $(R \neg \vee)$*,* $(L \neg \rightarrow)$ *and* $(R \neg \rightarrow)$*.*

PROOF. In all cases we argue by contraposition and double negation elimination. We only treat one exemplary case $(L \neg \rightarrow)$: Assume that $\neg B \Rightarrow \Delta, \neg A$ is derivable. Then by admissible (Cp) $\neg \neg A, \neg \Delta \Rightarrow \neg \neg B$. By double negation elimination and suitable cuts we get $A, \neg \Delta \Rightarrow B$. By $(R \rightarrow)$ we have $\neg \Delta \Rightarrow A \rightarrow B$. Again using (Cp), double negation elimination and weakening we get the desired $\neg(A \rightarrow B) \Rightarrow \Delta$. ∎

LEMMA 10. *The following rules are admissible in* $\mathbf{G3h_p}$*:* $(\mathtt{ConCp})$ *and* $(\mathtt{ClCp})$*.*

PROOF. Again we use admissible contraposition and inversion for $\neg \neg$. For example for $(\mathtt{ConCp})$ we argue by assuming that $\Gamma \Rightarrow \neg \Delta$ is derivable. Admissible contraposition gives $\neg \neg \Delta \Rightarrow \neg \Gamma$. We can eliminate the double negations due to inversion for $\neg \neg$. ∎

Combining these lemmata we get:

THEOREM 1. $\mathbf{G1h_p}$ *and* $\mathbf{G3h_p}$ *are equivalent, i.e.* $\Gamma \Rightarrow \Delta$ *is derivable in* $\mathbf{G1h_p}$ *iff* $\Gamma \Rightarrow \Delta$ *is derivable in* $\mathbf{G3h_p}$*.*

2.5.   Equivalence with the Axiomatic Calculus of HYPE

LEMMA 11. *For all finite Γ, if $\Gamma \vdash_{\mathbf{N_p^\bullet}} A$, then $\Gamma \Rightarrow A$ is derivable in $\mathbf{G3h_p}$.*

PROOF. All the axioms are derivable in $\mathbf{G3h_p}$. Conditional contraposition is admissible in $\mathbf{G3h_p}$, because if $\Rightarrow A \to B$ is derivable, then we can cut on the derivable sequent $A \to B, A \Rightarrow B$ and therefore $A \Rightarrow B$ is also derivable. Using admissible contraposition we also have $\neg B \Rightarrow \neg A$ and by $(\mathtt{R} \to)$ we have $\Rightarrow \neg B \to \neg A$.   ∎

As usual $\Gamma \Rightarrow \Delta$ corresponds to $\bigwedge \Gamma \to \bigvee \Delta$. If Γ is empty $\bigwedge \Gamma$ is $\neg\bot$ and if Δ is empty, then $\bigvee \Delta$ is $\bot$. Then we have:

LEMMA 12. *If $\Gamma \Rightarrow \Delta$ is derivable in $\mathbf{G3h_p}$, then $\vdash_{\mathbf{N_p^\bullet}} \bigwedge \Gamma \to \bigvee \Delta$.*

PROOF. For the initial sequents we know that $\vdash_{\mathbf{N_p^\bullet}} A \to A$. We also have the structural property that if $\Gamma \vdash A$ and $\Gamma \subseteq \Gamma'$, then $\Gamma' \vdash A$. With the deduction theorem and the derivable import export law, we get the contexts in the antecedent and with the disjunction axioms we get them in the succedents. For $(\mathtt{L}\bot)$ we distinguish the case Δ is empty and nonempty. In the former case we have $\bot \to \bot$ as an axiom and then argue as before. If Δ is nonempty we choose one of the formulas A in Δ for our instance of $\bot \to A$. $(\mathtt{R}\neg\bot)$ is then treated similarly with the addition of contraposition. The admissibility of the propositional rules is simple. We only show $(\mathtt{L}\neg\to)$. By IH $\vdash_{\mathbf{N_p^\bullet}} \neg B \to \bigvee \Delta \vee \neg A$, then $\neg B \vdash_{\mathbf{N_p^\bullet}} \bigvee \Delta \vee \neg A$ and $\neg B \vdash_{\mathbf{N_p^\bullet}} \neg(\neg \bigvee \Delta \wedge A)$. Then $\vdash_{\mathbf{N_p^\bullet}} \neg B \to \neg(\neg \bigvee \Delta \wedge A)$ and by contraposition $\vdash_{\mathbf{N_p^\bullet}} \neg \bigvee \Delta \wedge A \to B$. By import export $\neg \bigvee \Delta \vdash_{\mathbf{N_p^\bullet}} A \to B$ giving $\vdash_{\mathbf{N_p^\bullet}} \neg \bigvee \Delta \to (A \to B)$ and contraposing and double negation elimination gives $\vdash_{\mathbf{N_p^\bullet}} \neg(A \to B) \to \bigvee \Delta$.
                                                                                 ∎

With these equivalence results we can infer the soundness and completeness of our systems $\mathbf{G3h_p}$ and also $\mathbf{G1h_p}$ from the soundness and completeness theorems of Leitgeb [4] and Speranski [9]. Since we do not repeat the semantics here, we just state it as an observation.[7]

OBSERVATION 1. *The proof systems $\mathbf{G3h_p}$ and $\mathbf{G1h_p}$ are sound and complete with respect to the semantics of* HYPE.

---

[7]For a presentation of the semantics of HYPE consult Leitgeb [4] and for a simplification Speranski [9].

3.   Cut Elimination

3.1.   A Counterexample for Cut-Elimination in G3h$_\mathbf{p}$

One of the central questions concerning sequent systems is that of cut-eliminability. For a positive answer we could either try to follow Gentzen's strategy of introducing a more general rule of cut called mix or we could try to make use of the admissibility of weakening and contraction. Ono [8] used the first strategy to establish cut-elimination for several non-classical systems and Negri and van Plato [6] used the second strategy. However, in the course of establishing the cut-elimination for our systems **G1h$_\mathbf{p}$**, **G3h$_\mathbf{p}$**, we run into a problem, due to the additional rules of contraposition or the negated conditional. In the case of contraposition rules an inductive argument is problematic, because the logical complexity of formulas in the upper sequents of contraposition is not less than the ones in the lower sequent. Moreover, we do not have a single principal formula.

In the **G3h$_\mathbf{p}$** case with a negated conditional everything seems to be prepared. We have propositional rules of the correct form, we have admissibility of weakening and contraction. However, we encounter a problem due to the additional rules for the negated conditional.

Consider a cut in which the cut-formula is non-principal in the right upper sequent, which is inferred by a rule with restricted contexts, such as $(\texttt{R} \to)$. In this case the usual cut-elimination procedure does not work. If we try to push the cut upwards on the right, then we introduce a context that blocks the applicability of $(\texttt{R} \to)$. A similar problem already occurs in the case of multi-conclusion systems of intuitionistic logic. Although there are ways to solve the problem in the case of intuitionistic logic, these strategies are not applicable in the additional cases due to the negated rules. In the **G3h$_\mathbf{p}$** case, we have no inversion for $(\texttt{R} \to)$ and $(\texttt{L}\neg \to)$ and therefore the strategy employed in Negri and van Plato [6] does not work. Also Ono's strategy in [8] is not applicable in our case.

The problem bears some analogy to the case of constant domains in intuitionistic logic, where we have a well-known counterexample to cut-elimination due to the asymmetry of the conditional introduction with restricted contexts and universal quantifier introduction without a restriction. It is also possible to construct a counterexample to cut elimination for **G3h$_\mathbf{p}$**. Consider the following sequent with the propositional variables P, Q, R:

$$\neg(R \to \neg(P \to Q)), P \Rightarrow Q \tag{†}$$

654 M. Fischer

A proof search shows that there is no derivation of (†) in **G3h$_\mathbf{p}$** without cut. An application of (L¬→) would require that there is no formula in the context, but $\neg\neg(P \to Q) \Rightarrow Q, \neg R$ is not derivable. On the other hand the following are derivable:

$$\neg(R \to \neg(P \to Q)), P \Rightarrow P \to Q \tag{7}$$

$$P \to Q, \neg(R \to \neg(P \to Q)), P \Rightarrow Q \tag{8}$$

Although (†) is not cut-free derivable in **G3h$_\mathbf{p}$** it would be, if we could use a negated conditional introduction on the sequent $\neg\neg(P \to Q), P \Rightarrow Q, \neg R$, which is easily seen to be derivable. Interesting is that the occurrence of $\neg R$ in the succedent is introduced by weakening and therefore independent of the occurrence of P in the antecedent.

Kashima and Shimura [3] solve the problem in the constant domain case by introducing connections and allowing for additional instances of conditional introduction to establish a cut-elimination procedure. The main idea is to strengthen the introduction rules for the conditional to allow for contexts that are not connected to the antecedent. The connections enable an explicit version of the relevant independence. Therefore we will follow Kashima and Shimura's strategy in the next subsection.

3.2. Connections

The main innovation of Kashima's calculus is that we additionally keep track of connections between formula occurrences at sequents throughout proofs. These connections were first introduced by Kashima in [2] to solve the problems for the constant domain variation of intuitionistic logic and refined in Kashima and Shimura [3].[8]

The connections provide an extra layer of information regarding the sequents within a derivation. As a step towards our goal we have to clearly distinguish different formula occurrences of one and the same formula within a sequent. In order to keep track of occurrences we use labellings. A labelling of the formula occurrences in a sequent by natural numbers is a one-to-one function from a finite set of natural numbers to the set of formulas in the sequent. We only require that the labelling function is one-to-one and not that the ordering is preserved. So for example for the sequent $A, B, B \Rightarrow A, C$ we can use the labelling $A^1, B^2, B^3 \Rightarrow A^4, C^5$, but also $A^7, B^5, B^3 \Rightarrow A^1, C^{16}$.

---

[8]Kashima and Shimura [3] work with sequences instead of multisets.

We introduce a binary relation $\sim$ of connection between formula occurrences within a sequent.[9] When $A^i \sim B^j$ we say that the formula occurrence A^i is connected to the formula occurrence B^j. The main idea is that there is a dependence relation between the two formula occurrences in an initial sequent of the form $A \Rightarrow A$. Formulas introduced by weakening on the other hand lack this dependence relation. These dependencies are preserved along derivations and the connections are a bookkeeping device. Since the formula occurrences in the lower sequent of an inference in most cases do have a unique set of ancestors, the connections are inherited from its ancestors.

In the following we will explicitly add the information of connections of a formula occurrence. We will accomplish this by listing the labels of all the connected formula occurrences within brackets behind the formula occurrence. This will also help to see the locality of the inferences.

Let us illustrate this with an example. Consider our previous sequent now with some additional information $A^1[], B^2[3,5], \Rightarrow B^3[2], C^4[], C^5[2]$. Then at this sequent A^1 has no connections, but B^2 is connected to B^3 and C^5, i.e. $B^2 \sim B^3$ and $B^2 \sim C^5$ and C^4 has no connections.[10]

In order to simplify the presentation of the calculus we will employ some conventions:

We will use sequents of the form $A^i[1_i, \ldots, n_i], \Gamma \Rightarrow \Delta, B^j[1_j, \ldots, m_j]$.

With $[]^i$ we refer to the finite set of labels of formula occurrences connected to the formula occurrence A^i, i.e. $[1_i, \ldots, n_i]$. Since formula occurrences and labels are uniquely identifiable we will switch back and forth between the two, so that for example $[]^{A^i}$ will be used for the same purpose. These sets of labels $[]^i$ keeping track of the connections can also be empty.

In a sequent we usually do not display the sets of connections of the contexts, but we assume that every formula occurrence $\gamma^k \in \Gamma$ has a set of connections $[]^{\gamma^k}$. In general we will use the corresponding small letter γ to refer to a(n arbitrary) formula of Γ. In order to keep track of the connections to side formula occurrences i we make use of the successor labels i', i.e. if i is the label of a side formula occurrence in one of the premises of a rule, then i' is the label of the occurrence of the successor formula in the lower sequent. To avoid overlaps in cases like $(R \rightarrow')$ and $(L\neg\rightarrow')$ as well as $(R \rightarrow^+)$ and $(L\neg\rightarrow^+)$, we assume that the labelling of the lower sequent is chosen, such

---

[9]Not to be confused with the intuitionistic negation.

[10]The brackets convention is not the most parsimonious version as we can recover the connections of the formula occurrences in the succedent from the connections in the antecedent.

that no successor label of non-principal formula is identical to a label of an active formula.

In the case of non-principal formulas in the context the connections are inherited directly, i.e. $[1', \ldots, n']^{i'}$ at the lower sequent is inherited from $[1, \ldots n]^{i}$ at an upper sequent in all the cases except for the rules $(\mathtt{L} \rightarrow)$, $(\mathtt{R}\neg\rightarrow)$ and $(\mathtt{Cut})$, which will be explicitly stated.

The expression $[1, \ldots, n \setminus i]$ refers to the set of labels $1, \ldots, n$ without i.

For inferences with two premises we use the shorthand 'lus' for left upper sequent and 'rus' for right upper sequent.

The Systems $\mathbf{cGh_p}$ and $\mathbf{cGh_p^+}$. Due to a technical consideration it is convenient to work with a sequent system in which there are no empty cedents. In order to guarantee this we slightly deviate from the previous systems. One modification concerns the initial sequents of $\bot$ and $\neg\bot$.[11] Another modification concerns the conditional rules with restricted contexts. Here we introduce an arbitrary new formula C. The additional arbitrary formula C in the conditional rules is introduced for technical reasons similar to the occurrences of A in the initial sequents for $\bot$ and $\top$. We call the occurrences of A^i in the initial sequents for $\bot$ and $\top$ as well as the occurrences of C^l in the $\rightarrow$-rules auxiliary occurrences.

The system $\mathbf{cGh_p}$ consists of the following rules:

Initial sequents

For all formulas A:

$$\begin{array}{ll}
(\mathtt{ID}) & A^i[j] \Rightarrow A^j[i] \\
(\mathtt{L}\bot') & \bot^i[j] \Rightarrow A^j[i] \\
(\mathtt{R}\neg\bot') & A^i[j] \Rightarrow (\neg\bot)^j[i]
\end{array}$$

Structural rules

$$(\mathtt{LW}) \ \frac{\Gamma \Rightarrow \Delta}{A^i, \Gamma \Rightarrow \Delta} \quad (\mathtt{RW}) \ \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A^i}$$

$$(\mathtt{LC}) \ \frac{A^i[1_i, \ldots, n_i], A^j[1_j, \ldots, m_j], \Gamma \Rightarrow \Delta}{A^k[1'_i, \ldots, n'_i, 1'_j, \ldots, m'_j], \Gamma \Rightarrow \Delta} \quad (\mathtt{RC}) \ \frac{\Gamma \Rightarrow \Delta, A^i[1_i, \ldots, n_i], A^j[1_j, \ldots, m_j]}{\Gamma \Rightarrow \Delta, A^k[1'_i, \ldots, n'_i, 1'_j, \ldots, m'_j]}$$

Propositional rules

---

[11]This modification is already present in Kashima [2] and Kashima and Shimura [3].

$$(\text{L}\neg\neg)\ \frac{A^i[1_i,\ldots,n_i],\Gamma\Rightarrow\Delta}{(\neg\neg A)^i[1'_i,\ldots,n'_i],\Gamma\Rightarrow\Delta}\qquad(\text{R }\neg\neg)\ \frac{\Gamma\Rightarrow\Delta,A^i[1_i,\ldots,n_i]}{\Gamma\Rightarrow\Delta,(\neg\neg A)^i[1'_i\ldots,n'_i]}$$

$$(\text{L}\vee)\ \frac{A^i[1_i,\ldots,n_i],\Gamma\Rightarrow\Delta\qquad B^j[1_j,\ldots,m_j],\Pi\Rightarrow\Lambda}{(A\vee B)^k[1'_i,\ldots,n'_i,1'_j,\ldots,m'_j],\Gamma,\Pi\Rightarrow\Delta,\Lambda}$$

$$(\text{R}\vee)\ \frac{\Gamma\Rightarrow\Delta,A^i[1_i,\ldots,n_i]}{\Gamma\Rightarrow\Delta,(A\vee B)^i[1'_i,\ldots,n'_i]}\qquad(\text{R}\vee)\ \frac{\Gamma\Rightarrow\Delta,A^i[1_i,\ldots,n_i]}{\Gamma\Rightarrow\Delta,(B\vee A)^i[1'_i,\ldots,n'_i]}$$

$$(\text{L}\neg\vee)\ \frac{(\neg A)^i[1_i,\ldots,n_i],\Gamma\Rightarrow\Delta}{(\neg(A\vee B))^i[1'_i,\ldots,n'_i],\Gamma\Rightarrow\Delta}\qquad(\text{L}\neg\vee)\ \frac{(\neg A)^i[1_i,\ldots,n_i],\Gamma\Rightarrow\Delta}{(\neg(B\vee A))^i[1'_i,\ldots,n'_i],\Gamma\Rightarrow\Delta}$$

$$(\text{R}\neg\vee)\ \frac{\Gamma\Rightarrow\Delta,(\neg A)^i[1_i,\ldots,n_i]\qquad\Pi\Rightarrow\Lambda,(\neg B)^j[1_j,\ldots,m_j]}{\Gamma,\Pi\Rightarrow\Delta,\Lambda,(\neg(A\vee B))^k[1'_i,\ldots,n'_i,1'_j,\ldots,m'_j]}$$

$$(\text{R}\rightarrow')\ \frac{A^i[1_i,\ldots,n_i],\Gamma\Rightarrow B^j[1_j,\ldots,m_j]}{C^l[k],\Gamma\Rightarrow(A\rightarrow B)^k[l,1'_j,\ldots,m'_j\setminus i]}$$

$$(\text{L}\neg\rightarrow')\ \frac{(\neg B)^i[1_i,\ldots,n_i]\Rightarrow\Delta,(\neg A)^j[1_j,\ldots,m_j]}{(\neg(A\rightarrow B))^k[l,1'_i,\ldots,n'_i\setminus j]\Rightarrow\Delta,C^l[k]}$$

$$(\text{L}\rightarrow)\ \frac{\Gamma\Rightarrow\Delta,A^i[1_i,\ldots,n_i]\qquad B^j[1_j,\ldots,m_j],\Pi\Rightarrow\Lambda}{(A\rightarrow B)^k[1'_j,\ldots,m'_j],\Gamma,\Pi\Rightarrow\Delta,\Lambda}$$

With the sets of connections at the lower sequent defined as:

$$n'\in[]^k\ \text{iff}\ \ n'\in\Lambda\text{ and }n\in[]^j\text{ at rus;}$$

$$n'\in[]^{\gamma^{l'}}\ \text{iff}\ \ \begin{cases}n'\in\Delta\ \&\ n\in[]^{\gamma^l}\text{ at lus, or}\\ n'\in\Lambda\ \&\ i\in[]^{\gamma^l}\text{ at lus }\ \&\ n\in[]^j\text{ at rus;}\end{cases}$$

$$n'\in[]^{\pi^{l'}}\ \text{iff}\ \ n'\text{ in }\Lambda\text{ and }n\in[]^{\pi_l}\text{ at rus.}$$

$$(\text{R}\neg\rightarrow) \; \frac{\Gamma \Rightarrow \Delta, (\neg B)^i[1_i, \ldots, n_i] \qquad (\neg A)^j[1_j, \ldots, m_j], \Pi \Rightarrow \Lambda}{\Gamma, \Pi \Rightarrow \Delta, \Lambda, (\neg(A \rightarrow B))^k[1'_j, \ldots, m'_j]}$$

With the sets of connections at the lower sequent defined as:

$$
\begin{aligned}
n' \in []^{\delta^{l'}} \;\; &\text{iff} \quad n' \text{ in } \Gamma \text{ and } n \in []^{\delta^l} \text{ at lus.} \\
n' \in []^{\lambda^{l'}} \;\; &\text{iff} \quad
\begin{cases}
n' \in \Pi \;\&\; n \in []^{\lambda^l} \text{ at rus, or} \\
n' \in \Gamma \;\&\; n \in []^i \text{ at lus} \quad \&\; j \in []^{\lambda^l} \text{ at rus;}
\end{cases} \\
n' \in []^k \;\; &\text{iff} \quad n' \in \Gamma \text{ and } n \in []^i \text{ at lus;}
\end{aligned}
$$

$$(\text{Cut}) \; \frac{\Gamma \Rightarrow \Delta, C^i[1_i, \ldots, n_i] \qquad C^j[1_j, \ldots, m_j], \Pi \Rightarrow \Lambda}{\Gamma, \Pi \Rightarrow \Delta, \Lambda}$$

With the sets of connections at the lower sequent defined as:

$$
\begin{aligned}
n' \in []^{\gamma^{l'}} \;\; &\text{iff} \quad
\begin{cases}
n' \in \Delta \;\&\; n \in []^{\gamma^l} \text{ at lus, or} \\
n' \in \Lambda \;\&\; i \in []^{\gamma^l} \text{ at lus} \quad \&\; n \in []^j \text{ at rus;}
\end{cases} \\
n' \in []^{\pi^{l'}} \;\; &\text{iff} \quad n' \text{ in } \Lambda \text{ and } n \in []^{\pi^l} \text{ at rus.}
\end{aligned}
$$

The system $\mathbf{cGh_p^+}$ is gained from $\mathbf{cGh_p}$ by replacing the rules with restricted contexts $(\text{R}\rightarrow')$ and $(\text{L}\neg\rightarrow')$ by their less restrictive counterparts $(\text{R}\rightarrow^+)$ and $(\text{L}\neg\rightarrow^+)$:

$$(\text{R}\rightarrow^+) \; \frac{A^i[1_i, \ldots, n_i], \Gamma \Rightarrow \Delta, B^j[1_j, \ldots, m_j]}{C^l[k], \Gamma \Rightarrow \Delta, (A \rightarrow B)^k[l, 1'_j, \ldots, m'_j \setminus i]} \; \text{ for all } \delta \in \Delta, \, \delta \notin []^i$$

$$(\text{L}\neg\rightarrow^+) \; \frac{(\neg B)^i[1_i, \ldots, n_i], \Gamma \Rightarrow \Delta, (\neg A)^j[1_j, \ldots, m_j]}{(\neg(A \rightarrow B))^k[l, 1'_i, \ldots, n'_i \setminus j], \Gamma \Rightarrow \Delta, C^l[k]} \; \text{ for all } \gamma \in \Gamma, \, \gamma \notin []^j$$

Let us illustrate the connections due to an example, a derivation of (†) in $\mathbf{cGh_p^+}$:

$$\cfrac{\cfrac{\cfrac{\cfrac{P^i[j] \Rightarrow P^j[i]}{P^i[j] \Rightarrow P^j[i], (\neg R)^k} \qquad Q^l[m] \Rightarrow Q^m[l]}{(P \to Q)^n[m], P^i[m] \Rightarrow Q^m[n,i], (\neg R)^k}}{\cfrac{(\neg\neg(P \to Q))^n[m], P^i[m] \Rightarrow Q^m[n,i], (\neg R)^k}{(\neg(R \to \neg(P \to Q)))^o[m,p], P^i[m] \Rightarrow Q^m[o,i], Q^p[o]}}}{(\neg(R \to \neg(P \to Q)))^o[m], P^i[m] \Rightarrow Q^m[o,i]} \; (\mathtt{L}\neg\to^+)$$

The application of $(\mathtt{L}\neg\to^+)$ is justified since the formula occurrence $(\neg D)^k$ at the upper sequent is not connected to A^i.

3.3.   Cut Elimination for $\mathbf{cGh_p^+}$

In this section we work with a variation of $\mathbf{cGh_p^+}$, where cut is replaced by the following more general mix rule. In the following Δ_1, Π_1 are multisets of a single formula, the mix formula M. The sets Δ_1, Π_1 are allowed to be empty and the mixformula M is allowed to appear in Π_0, Δ_0, such that not all occurrences of M in the upper sequents have to be deleted.[12]

$$(\mathtt{mix}) \; \frac{\Gamma \Rightarrow \Delta_0, \Delta_1 \qquad \Pi_1, \Pi_0 \Rightarrow \Lambda}{\Gamma, \Pi_0 \Rightarrow \Delta_0, \Lambda}$$

With the sets of connections at the lower sequent defined as:

$$n' \in []^{\gamma^{l'}} \text{ iff } \begin{cases} n' \in \Delta_0 \;\&\; n \in []^{\gamma^l} \text{ at lus, or} \\ n' \in \Lambda \quad \&\; \exists M^i \in \Delta_1 \text{ at lus } \exists M^j \in \Pi_1 \text{ at rus,} \\ \qquad \text{with } i \in []^{\gamma^l} \&\, n \in []^j; \end{cases}$$

$n' \in []^{\pi^{l'}}$ iff n' in Λ and $n \in []^{\pi^l}$ at rus.

As usual we define the grade of the mixformula to be the logical complexity of the formula, in our case it is its positive depth.

In an inference all the side formula occurrences in the lower sequent have a corresponding formula occurrence in an upper sequent, called its predecessors. In the case of contraction we call both active formula occurrences the predecessors of the principal formula occurrence.

As for the rank of a mix inference we use the sum of the left rank and the right rank as for example in Takeuti [10], p.23. The difference to Takeuti's version stems from the fact that we allow for our multisets of the mixformula to be empty as well as occurrences of the mixformula in Δ_0 and Π_0. To adjust our system we define the left (right) rank of a mix to be the maximum length of subsequent sequents in the subderivation $\mathcal{D}$ of the left upper sequent

---

[12]Kashima and Shimura [3] use such a rule as well as Ono [8] and the cut elimination follows the strategy of Kashima and Shimura [3].

(right upper sequent), that contain a predecessor of one of the mixformula occurrences in Δ_1 (Π_1). In the case that the set of mixformulas on the left (right) is empty we let the left (right) rank be 0.

If $\mathcal{D}$ is a derivation $\mathsf{last}(\mathcal{D})$ refers to the last inference of $\mathcal{D}$ and $\mathsf{fs}(\mathcal{D})$ to the final sequent of $\mathcal{D}$. If there are two derivations $\mathcal{D}, \mathcal{D}'$ of the same final sequent $\Gamma \Rightarrow \Delta$, then we say that $\mathcal{D} \preccurlyeq \mathcal{D}'$ iff all formula occurrences that are connected in $\mathsf{fs}(\mathcal{D})$ are also connected in $\mathsf{fs}(\mathcal{D}')$. $\mathcal{D} \simeq \mathcal{D}'$ iff $\mathcal{D} \preccurlyeq \mathcal{D}'$ and $\mathcal{D}' \preccurlyeq \mathcal{D}$

The following lemma is a variant of Kashima and Shimura's Lemma 3.2. that states, that if $A, \Gamma \Rightarrow \Delta$ is cut-free derivable and A has no connections to Δ, then $\Gamma \Rightarrow \Delta$ is also cut-free derivable. In our case this would not hold without the additional formula occurrence C^l in the lower sequent of the conditional rules, due to the simple fact that we could have 'anti-theorems', i.e. sequents with empty succedent.[13] To avoid empty succedents was one of the reasons for Kashima and Shimura to use initial sequents of the form $\bot \Rightarrow A^i$. In our case we avoid empty cedents completely. Theorems have the form $\top \Rightarrow A$ and anti-theorems $B \Rightarrow \bot$.

For the following lemma we slightly adapt the notion of depth of a derivation $\mathcal{D}$ by counting multiple successive applications of weakening as a single application and labelling this measure as $\mathsf{d}(\mathcal{D})$.

LEMMA 13.

(i) *If there is a cut- and mixfree derivation $\mathcal{D}$ of $A^i, \Gamma \Rightarrow \Delta$ in $\mathbf{cGh_p^+}$, such that $\delta \notin []^i$ in $\mathsf{last}(\mathcal{D})$ for all $\delta \in \Delta$, then there is a cut- and mixfree derivation $\mathcal{D}'$ of $\Gamma \Rightarrow \Delta$ in $\mathbf{cGh_p^+}$, with $\mathcal{D}' \preccurlyeq \mathcal{D}$ and $\mathsf{d}(\mathcal{D}') \leq \mathsf{d}(\mathcal{D})$.*

(ii) *If there is a cut- and mixfree derivation $\mathcal{D}$ of $\Gamma \Rightarrow \Delta, A^i$ in $\mathbf{cGh_p^+}$, such that $\gamma \notin []^i$ in $\mathsf{last}(\mathcal{D})$ for all $\gamma \in \Gamma$, then there is a cut- and mixfree derivation $\mathcal{D}'$ of $\Gamma \Rightarrow \Delta$ in $\mathbf{cGh_p^+}$, with $\mathcal{D}' \preccurlyeq \mathcal{D}$ and $\mathsf{d}(\mathcal{D}') \leq \mathsf{d}(\mathcal{D})$.*

PROOF. We prove (i) by induction on $\mathsf{d}(\mathcal{D})$. The case of initial sequents is impossible.

For the rules we distinguish the case where A is principal or nonprincipal. The case that A is an auxiliary is not possible.

Weakening is in both cases trivial. Contraction is also simple in both cases.

The propositional cases in which the active formulas are on the same side as the principal are straightforward.

---

[13] Thanks to a referee for demanding a clarification of this point.

Case (L $\rightarrow$).

$$\frac{\Theta \Rightarrow \Lambda, B^l \qquad C^j, \Pi \Rightarrow \Sigma}{(B \rightarrow C)^k, \Theta, \Pi \Rightarrow \Lambda, \Sigma}$$

If $A = B \rightarrow C$ is principal and $(B \rightarrow C)^k$ is not connected to Λ, Σ, then C^j is also not connected to Σ in the right upper sequent and we get the desired by IH and weakening.

If A is nonprincipal, then we have two subcases:

- Subcase 1:
  $A^i \in \Theta$. If $j \notin []^\sigma$ for all $\sigma \in \Sigma$ at the right upper sequent, then we use the IH on the right upper sequent and weakening. If $j \in \sigma$ for some $\sigma \in \Sigma$ at the right upper sequent, then $\lambda \notin []^i$ and $l \notin []^i$ at the left upper sequent, as otherwise A^i would be connected to some $\sigma \in \Sigma$ at the lower sequent, in contrast to the assumption. Then we can use the IH on the left upper sequent and then (L $\rightarrow$).
- Subcase 2:
  $A^i \in \Pi$, then at the right upper sequent A^i is not connected to any $\sigma \in \Sigma$ and we can use the IH and then (L $\rightarrow$).

Case (R $\rightarrow^+$).
Then A can only be in Γ and then we can use the IH and then (R $\rightarrow^+$).

Case (L$\neg\rightarrow^+$).

$$\frac{(\neg B)^o[1_i, \dots, n_i], \Gamma \Rightarrow \Delta, (\neg D)^j[1_j, \dots, m_j]}{(\neg(D \rightarrow B))^k[l, 1'_i, \dots, n'_i \setminus j], \Gamma \Rightarrow \Delta, C^l[k]} \text{ for all } \gamma \in \Gamma, \gamma \notin []^j$$

Then A^i can only be nonprincipal and we can use IH, since A^i is neither connected to any $\delta \in \Delta$ nor $(\neg D)^j$ at the upper sequent.

Case (R$\neg\rightarrow$).
Here A^i is also only in the context with two subcases. In both of them we can use the IH.

For (*ii*) we argue analogously. ∎

In the next lemma, which corresponds to Lemma 3.3. in [3], we show that cuts for the initial sequents (L$\perp'$) and (R$\neg\perp'$) are eliminable.

LEMMA 14.

(i) *If there is a cut- and mixfree derivation $\mathcal{D}$ of $\Gamma \Rightarrow \Delta, \perp^j$ in $\mathbf{cGh_p^+}$, then there is a cut- and mixfree derivation $\mathcal{D}'$ of $\Gamma \Rightarrow \Delta, A^j$ in $\mathbf{cGh_p^+}$, such that $[]^\gamma$ at $\mathsf{fs}(\mathcal{D}')$ is the same as $[]^\gamma$ at $\mathsf{fs}(\mathcal{D})$ for all $\gamma \in \Gamma$ and $|\mathcal{D}'| \leq |\mathcal{D}|$;*

(ii) if there is a cut- and mixfree derivation $\mathcal{D}$ of $(\neg\bot)^j, \Gamma \Rightarrow \Delta$ in $\mathbf{cGh_p^+}$, then there is a cut- and mixfree derivation $\mathcal{D}'$ of $A^j, \Gamma \Rightarrow \Delta$ in $\mathbf{cGh_p^+}$, such that $[]^\delta$ at $\mathsf{fs}(\mathcal{D}')$ is the same as $[]^\delta$ at $\mathsf{fs}(\mathcal{D})$ for all $\delta \in \Delta$ and $|\mathcal{D}'| \leq |\mathcal{D}|$.

PROOF. (*i*) is proved by induction on the depth of the derivation $\mathcal{D}$.

If $\Gamma \Rightarrow \Delta, \bot^j$ is an initial sequent, then it has to be of the form $\bot^k \Rightarrow \bot^j$ and we get $\bot^k \Rightarrow A^j$ as an instance of $(\mathtt{L}\bot')$.

If $\mathsf{fs}(\mathcal{D})$ is not an initial sequent, then $\bot^j$ is principal, auxiliary or neither.

In the principal case it is a structural inference. Weakening is obvious. Contraction is handled by using the induction hypothesis twice and then contraction.

In the auxiliary case it is $(\mathtt{L}\neg\rightarrow^+)$ with $C = \bot$. In this case, we use the rule $(\mathtt{L}\neg\rightarrow^+)$ on the upper sequent with $C = A$.

In the other cases we argue by employing the induction hypothesis on the upper sequents of $\mathsf{last}(\mathcal{D})$. (*ii*) is similar. ∎

LEMMA 15. *If there are cut- and mixfree derivations $\mathcal{D}$ of $\Gamma \Rightarrow \Delta_0, \Delta_1$ in $\mathbf{cGh_p^+}$ and $\mathcal{E}$ of $\Pi_1, \Pi_0 \Rightarrow \Lambda$ in $\mathbf{cGh_p^+}$ and Δ_1, Π_1 are (possibly empty) sequences consisting only of a formula M, then there is a cut- and mixfree derivation $\mathcal{F}$ of $\Gamma, \Pi_0 \Rightarrow \Delta_0, \Lambda$ in $\mathbf{cGh_p^+}$, such that all the connections at $\mathsf{fs}(\mathcal{F})$ are connections at the sequent that would result in an application of mix on $\mathsf{fs}(\mathcal{D})$ and $\mathsf{fs}(\mathcal{E})$, i.e.*

(i) *if $\delta_0 \in []^\gamma$ at $\mathsf{fs}(\mathcal{F})$, then $\delta_0 \in []^\gamma$ at $\mathsf{fs}(\mathcal{D})$;*

(ii) *if $\lambda \in []^{\pi_0}$ at $\mathsf{fs}(\mathcal{F})$, then $\lambda \in []^{\pi_0}$ at $\mathsf{fs}(\mathcal{E})$;*

(iii) *$\delta_0 \notin []^{\pi_0}$ at $\mathsf{fs}(\mathcal{F})$;*

(iv) *if $\lambda \in []^\gamma$ at $\mathsf{fs}(\mathcal{F})$, then there is a δ_1 in Δ_1 with $\delta_1 \in []^\gamma$ at last $\mathsf{fs}(\mathcal{D})$ and there is a π_1 in Π_1 with $\lambda \in []^{\pi_1}$ at last $\mathsf{fs}(\mathcal{E})$.*

PROOF. By an induction on the grade of the mixformula M with a side induction on the rank. So the last inference that we want to replace is of the form

$$(\mathtt{mix}) \frac{\overset{\mathcal{D}}{\Gamma \Rightarrow \Delta_0, \Delta_1} \quad \overset{\mathcal{E}}{\Pi_1, \Pi_0 \Rightarrow \Lambda}}{\Gamma, \Pi_0 \Rightarrow \Delta_0, \Lambda}$$

We distinguish the cases:

(I) left rank = 0 or right rank = 0: then either Δ_1 or Π_1 are empty. These cases are easily treated by weakening.

(II) rank = 2 and left rank $\neq 0$ and right rank $\neq 0$:

If one of the sequents is an initial sequent, then

(a) $\mathsf{last}(\mathcal{D})$ is an instance of (ID), say $A^i \Rightarrow A^j$, then Δ_0 is empty and Π_1 is nonempty. If Π_1 consists of a single occurrence of A, we can use $\mathcal{E}$. Otherwise we use contraction on $\mathsf{fs}(\mathcal{E})$.

(b) $\mathsf{last}(\mathcal{D})$ is an instance $A^i \Rightarrow (\neg\bot)^j$ of (R$\neg\bot'$) and Π_1 is a nonempty multiset of $\neg\bot$. By contraction and Lemma 14 (ii) we can transform $\mathcal{E}$ into a cut- and mixfree derivation of $A^i, \Pi_0 \Rightarrow \Lambda$.

(c) $\mathsf{last}(\mathcal{D})$ is an instance $\bot^i \Rightarrow A^j$ of (L$\bot'$) and Π_1 is a nonempty multiset of A's. If none of the mixformulas in Π_1 are connected to some $\lambda \in \Lambda$ at $\mathsf{fs}(\mathcal{E})$, then we use Lemma 13 (i) and weakening. If some mixformula M^l is connected to some λ^k, then we use $\bot^i \Rightarrow \lambda^k$ as an instance of (L$\bot'$) and weakening.

(d) $\mathsf{last}(\mathcal{E})$ is an instance of (ID), similar to (a).

(e) $\mathsf{last}(\mathcal{E})$ is an instance of (L$\bot'$), similar to (b).

(f) $\mathsf{last}(\mathcal{E})$ is an instance of (R$\neg\bot'$), similar to (c).

If none of the sequents is an initial sequent, then on both sides we have a single mix formula that is either principal or auxiliary. If $\mathsf{last}(\mathcal{D})$ or $\mathsf{last}(\mathcal{E})$ is an instance of weakening, then we get the result also by weakening. So we focus on the propositional cases.

If $\mathsf{last}(\mathcal{E})$ is (R $\rightarrow^+$) with the auxiliary C being our mix formula, then our proof has the form

$$
(\mathtt{mix}) \frac{\begin{array}{c} \mathcal{D} \\ \Gamma \Rightarrow \Delta_0, C^m \end{array} \qquad \dfrac{\begin{array}{c} \mathcal{E}_0 \\ A^i, \Pi_0 \Rightarrow \Lambda, B^j \end{array}}{C^l, \Pi_0 \Rightarrow \Lambda, (A \rightarrow B)^k}}{\Gamma, \Pi_0 \Rightarrow \Delta_0, \Lambda, (A \rightarrow B)^k}
$$

Then we can argue as follows: If Γ has no connection to C^m at $\mathsf{fs}(\mathcal{D})$, then we can use Lemma 13 to get $\Gamma \Rightarrow \Delta_0$ and then use weakening. If there is a formula $\gamma \in \Gamma$, such that $m \in []^\gamma$ at $\mathsf{fs}(\mathcal{D})$, then we use this γ as our C in an application of (R $\rightarrow^+$) to $\mathsf{fs}(\mathcal{E}_0)$ and then weakening.

The case of (L$\neg\rightarrow^+$) as $\mathsf{last}(\mathcal{D})$ with C^l the mixformula is analogous.

In the case of the mix being principal the unique readability and the symmetry of the calculus allow us to argue standardly, so we only discuss the case of a negated conditional:

(e) Let us consider the case of $\neg\rightarrow$, then $\mathsf{last}(\mathcal{D})$ is of the form

$$\frac{\begin{array}{cc}\mathcal{D}_0 & \mathcal{D}_1\\ \Gamma' \Rightarrow \Delta_0', (\neg B)^i & (\neg A)^j, \Gamma'' \Rightarrow \Delta_0''\end{array}}{\Gamma \Rightarrow \Delta_0, (\neg(A \to B))^k}$$

where $\Gamma = \Gamma', \Gamma''$ and $\Delta_0 = \Delta_0', \Delta_0''$.

On the right hand $\mathsf{last}(\mathcal{E})$ is of the form

$$\frac{\overset{\mathcal{E}_0}{(\neg B)^m \Pi_0 \Rightarrow \Lambda, (\neg A)^n}}{(\neg(A \to B))^o, \Pi_0 \Rightarrow \Lambda, C^l} \quad \pi \notin []^n \text{ for all } \pi \in \Pi_0$$

Since rank is 2 there are no other occurrences of the mixformula relevant. Then our derivation $\mathcal{F}$ is

$$\frac{\frac{\overset{\mathcal{D}_0}{\Gamma' \Rightarrow \Delta_0', (\neg B)^i} \quad \overset{\mathcal{E}_0}{(\neg B)^m, \Pi_0 \Rightarrow \Lambda, (\neg A)^n}}{(*) \ \Gamma', \Pi_0 \Rightarrow \Delta_0', \Lambda, (\neg A)^n} \quad \frac{\overset{\mathcal{D}_1}{}}{(\neg A)^j, \Gamma'' \Rightarrow \Delta_0''}}{\frac{\Gamma, \Pi_0 \Rightarrow \Delta_0, \Lambda}{\Gamma, \Pi_0 \Rightarrow \Delta_0, \Lambda, C^l}}$$

It is simple to check that all the connections in the last sequent of $\mathcal{F}$ are connections in the sequent that we would get by applying mix.

(i) If $\delta_0 \in []^\gamma$ at $\mathsf{fs}(\mathcal{F})$, then $\delta_0 \in []^\gamma$ at $\mathsf{fs}(\mathcal{D}_0)$ or $\delta_0 \in []^\gamma$ at $\mathsf{fs}(\mathcal{D}_1)$ or $n \in []^\gamma$ at $(*)$ and $\delta_0 \in []^j$ at $\mathsf{last}(\mathcal{D}_1)$.
In the first two cases $\gamma \sim \delta_0$ at $\mathsf{fs}(\mathcal{D})$.
In the third case by IH $n \in []^\gamma$ at $(*)$ implies that $i \in []^\gamma$ at $\mathsf{fs}(\mathcal{D}_0)$ and $n \in []^m$ at $\mathsf{fs}(\mathcal{E}_0)$. But then $\delta_0 \in []^\gamma$ at $\mathsf{last}(\mathcal{D})$.

(ii) If $\lambda \in []^{\pi_0}$ at $\mathsf{fs}(\mathcal{F})$, then $\lambda \in []^{\pi_0}$ at $\mathsf{fs}(\mathcal{E}_0)$ and at $\mathsf{fs}(\mathcal{E})$.

(iii) $\delta_0 \notin []^{\pi_0}$ since $\pi_0 \notin []^n$ for all $\pi_0 \in \Pi_0$ at $\mathsf{fs}(\mathcal{E}_0)$.

(iv) If $\lambda \in []^\gamma$ at $\mathsf{fs}(\mathcal{F})$, then $\lambda \in []^\gamma$ at $(*)$. By IH $i \in []^\gamma$ at $\mathsf{fs}(\mathcal{D}_0)$, and $\lambda \in []^m$ at $\mathsf{fs}(\mathcal{E}_0)$. Then $k \in []^\gamma$ at $\mathsf{fs}(\mathcal{D})$ and $\lambda \in []^o$ at $\mathsf{fs}(\mathcal{E})$.

(v) $\gamma, \pi_0 \nsim C^l$.

(III) rank > 2 and left rank $\neq 0$ and right rank $\neq 0$:

If left rank > 1 we distinguish cases depending on $\mathsf{last}(\mathcal{D})$.

If right rank > 1 we distinguish cases depending on $\mathsf{last}(\mathcal{E})$.

Except for the conditional cases with restricted contexts the standard strategy works.

We discuss the interesting case of restricted contexts

(a) $last(\mathcal{E})$ is $(\mathtt{R} \to^+)$ and all the mixformulas in Π_1 are nonprincipal:

$$\frac{\overset{\mathcal{E}_0}{\Pi_1, \Pi_0, A^i \Rightarrow \Lambda, B^j}}{C^l, \Pi_1, \Pi_0 \Rightarrow \Lambda, (A \to B)^k} \quad \lambda \notin []^i \text{ for all } \lambda \in \Lambda$$

and $\mathsf{fs}(\mathcal{D})$ is $\Gamma \Rightarrow \Delta_0, \Delta_1$.

If C^l is not one of the mixformulas in Π_1 we can argue

$$\dfrac{\overset{\mathcal{D}}{\Gamma \Rightarrow \Delta_0, \Delta_1} \qquad \overset{\mathcal{E}_0}{\Pi_1, \Pi_0, A^i \Rightarrow \Lambda, B^j}}{\dfrac{(*)\ \Gamma, \Pi_0, A^i \Rightarrow \Delta_0, \Lambda, B^j}{C^l, \Gamma, \Pi_0 \Rightarrow \Delta_0, \Lambda, (A \rightarrow B)^k}} \text{ by IH}$$

The application of $(\mathtt{R} \rightarrow^+)$ is justified, since at $(*)$, for all $\lambda \in \Lambda$, $\lambda \notin []^i$ by assumption. But also $\delta_0 \notin []^i$, for all $\delta_0 \in \Delta_0$ by IH and the fact that in an application of mix no formulas of the contexts of the succedent of the left upper sequent and the antecedent of the right upper sequent get connected.

If $k \in []^\gamma$ at $\mathsf{fs}(\mathcal{F})$, then there is some $\delta_1 \in \Delta_1$ with $\delta_1 \in []^\gamma$ at $\mathsf{fs}(\mathcal{D})$ and some $\pi_1 \in \Pi_1$ with $j \in []^{\pi_1}$ at $\mathsf{fs}(\mathcal{E}_0)$. Then also $k \in []^{\pi_1}$ at $\mathsf{fs}(\mathcal{E})$. The other connections are easy to check.

If C^l is one of the mixformulas in Π_1, then we argue as before. In the case that there is no connection between Γ and Δ_1 in $\mathsf{fs}(\mathcal{D})$ we use Lemma 13 and weakening. Otherwise we use the IH and then choose C in an application of $(\mathtt{R} \rightarrow^+)$ to be one of the γ's that are connected to a mixformula in Δ_1.

(b) $\mathsf{last}(\mathcal{E})$ is $(\mathtt{L}\neg\rightarrow^+)$ and all of the mixformulas of Π_1 are nonprincipal:

$$\dfrac{\overset{\mathcal{E}_0}{\Pi_1, \Pi_0, (\neg B)^i \Rightarrow \Lambda, (\neg A)^j}}{\Pi_1, \Pi_0, (\neg(A \rightarrow B))^k \Rightarrow \Lambda, C^l} \ j \notin []^\pi \text{ for all } \pi \in \Pi_1, \Pi_0$$

and $\mathsf{fs}(\mathcal{D})$ is $\Gamma \Rightarrow \Delta_0, \Delta_1$ with Δ_1 containing only mixformulas. Then the derivation looks like

$$\dfrac{\overset{\mathcal{D}}{\Gamma \Rightarrow \Delta_0, \Delta_1} \qquad \overset{\mathcal{E}_0}{\Pi_1, \Pi_0, (\neg B)^i \Rightarrow \Lambda, (\neg A)^j}}{\dfrac{(*)\ \Gamma, \Pi_0, (\neg B)^i \Rightarrow \Delta_0, \Lambda, (\neg A)^j}{\Gamma, \Pi_0, (\neg(A \rightarrow B))^k \Rightarrow \Delta_0, \Lambda, C^l}} \text{ by IH}$$

The application of $(\mathtt{L}\neg\rightarrow^+)$ is justified, since $j \notin []^\pi$ for all $\pi \in \Pi_0$ at $(*)$. Also $j \notin []^\gamma$ for all $\gamma \in \Gamma$ at $(*)$, because $j \notin []^\pi$ for all $\pi \in \Pi_1$ at $\mathsf{fs}(\mathcal{E}_0)$.

For the connections we notice $\delta_0 \notin []^k$ at $\mathsf{fs}(\mathcal{F})$. If $\lambda \in []^k$ at $\mathsf{fs}(\mathcal{F})$, then $\lambda \in []^k$ at $\mathsf{fs}(\mathcal{E})$. Also $l \in []^k$ at $\mathsf{fs}(\mathcal{E})$.

If one of the mixformulas is principal and there are other occurrences of the mixformula in Π_1 that are nonprincipal, then by IH the following is also mixfree derivable:

$$\frac{\begin{array}{cc} \mathcal{D} & \mathcal{E}_0 \\ \Gamma \Rightarrow \Delta_0, \Delta_1 \quad \Pi_1, \Pi_0, (\neg B)^i \Rightarrow \Lambda, (\neg A)^j \end{array}}{\begin{array}{c} \dfrac{\mathcal{D}}{\Gamma \Rightarrow \Delta_0, \Delta_1} \quad \dfrac{\Gamma, \Pi_0, (\neg B)^i \Rightarrow \Delta_0, \Lambda, (\neg A)^j}{(*) \; (\neg(A \to B))^k, \Gamma, \Pi_0 \Rightarrow \Delta_0, \Lambda, C^l} \\ \hline \dfrac{\Gamma, \Gamma, \Pi_0 \Rightarrow \Delta_0, \Delta_0, \Lambda, C^l}{\Gamma, \Pi_0 \Rightarrow \Delta_0, \Lambda, C^l} \end{array}}$$

The application of $(\mathtt{L}\neg\!\to^+)$ is again justified. Moreover, we have reduced the right rank to 1, so that we can use the IH again.

If $\delta_0 \in [\,]^\gamma$ at $\mathsf{fs}(\mathcal{F})$, then $\delta_0 \in [\,]^\gamma$ at $\mathsf{fs}(\mathcal{D})$.
If $\lambda \in [\,]^\gamma$ at $\mathsf{fs}(\mathcal{F})$, then $m \in [\,]^\gamma$ for some mixformula M^m in Δ_1 at $\mathsf{fs}(\mathcal{D})$ and either $\lambda \in [\,]^k$ at $(*)$ or $\lambda \in [\,]^{\pi_1}$ at $\mathsf{fs}(\mathcal{E}_0)$. In the first case $\lambda \in [\,]^k$ also at $\mathsf{fs}(\mathcal{E})$ and in the second case $\lambda \in [\,]^\gamma$ at $\mathsf{fs}(\mathcal{E})$.
If $l \in [\,]^\gamma$ at $\mathsf{fs}(\mathcal{F})$ then $n \in [\,]^\gamma$ for some mixformula M^n in Δ_1 at $\mathsf{fs}(\mathcal{D})$.
$\delta_0 \notin [\,]^{\pi_0}$ at $\mathsf{fs}(\mathcal{F})$.
If $\lambda \in [\,]^{\pi_0}$ at $\mathsf{fs}(\mathcal{F})$, then $\lambda \in [\,]^{\pi_0}$ at $\mathsf{fs}(\mathcal{E})$ and $l \notin [\,]^{\pi_0}$ at $\mathsf{fs}(\mathcal{F})$.

(c) $last(\mathcal{D})$ is $(\mathtt{R}\to^+)$ analogous to (b).
(d) $last(\mathcal{D})$ is $(\mathtt{L}\neg\!\to^+)$ analogous to (a).   ∎

THEOREM 2. *If $\Gamma \Rightarrow \Delta$ is derivable in $\mathbf{cGh_p^+}$, then there is a cut- and mixfree derivation of $\Gamma \Rightarrow \Delta$ in $\mathbf{cGh_p^+}$.*

4.   Equivalence of $\mathbf{cGh_p^+}$ and $\mathbf{G1h_p}$

The next step is to relate the system $\mathbf{cGh_p^+}$ to our more familiar systems $\mathbf{G1h_p}$ and $\mathbf{G3h_p}$. Obviously $\mathbf{cGh_p^+}$ is an extension of $\mathbf{cGh_p}$. We notice that in $\mathbf{cGh_p}$ the rules of contraposition $(\mathtt{Cp})$, $(\mathtt{ConCp})$ and $(\mathtt{ClCp})$ are admissible, similar to Lemmas 7 and 10. Then it is also simple to show that $\mathbf{cGh_p}$ and $\mathbf{G1h_p}$ are equivalent in the following sense:

LEMMA 16. *For all $\Gamma \Rightarrow \Delta$, $\mathbf{G1h_p} \vdash \Gamma \Rightarrow \Delta$ iff $\mathbf{cGh_p} \vdash \top, \Gamma \Rightarrow \Delta, \bot$.*

PROOF. For left to right we argue by induction on the depth of the derivation in $\mathbf{G1h_p}$. For the initial sequents $(\mathtt{L}\bot)$, we argue by $(\mathtt{L}\bot')$ and weakening.

For the cases with unrestricted contexts no problem appears, for example $(\mathtt{L}\to)$:

$$\frac{\Gamma \Rightarrow \Delta, A \quad\quad B, \Gamma \Rightarrow \Delta}{A \to B, \Gamma \Rightarrow \Delta}$$

By IH $\top, \Gamma \Rightarrow \Delta, A, \bot$ and $\top, B, \Gamma \Rightarrow \Delta, \top$ are derivable in $\mathbf{cGh_p}$. With $(\mathtt{L}\rightarrow)$ we get $\top, \top, A \rightarrow B, \Gamma, \Gamma \Rightarrow \Delta, \Delta, \top, \top$ and with contraction we are done.

In the case of an $(\mathtt{R}\rightarrow)$-inference,[14]

$$\frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow A \rightarrow B}$$

we use the IH to get

$$\frac{\dfrac{\top, A, \Gamma \Rightarrow B, \bot \qquad \bot \Rightarrow B}{\dfrac{\top, A, \Gamma \Rightarrow B, B}{\dfrac{\top, A, \Gamma \Rightarrow B}{\dfrac{\top, \top, \Gamma \Rightarrow A \rightarrow B}{\dfrac{\top, \Gamma \Rightarrow A \rightarrow B}{\top, \Gamma \Rightarrow A \rightarrow B, \bot}}}}}}{}$$

The case $(\mathtt{L}\neg\rightarrow)$ is analogous.

For the other direction by adapting the proof of Lemma 9 we notice that if $\mathbf{cGh_p} \vdash \Gamma \Rightarrow \Delta$, then $\mathbf{G1h_p} \vdash \Gamma \Rightarrow \Delta$. The extra occurrences of $\bot, \top$ can then be eliminated with $(\mathtt{Cut})$ using the initial sequents $(\mathtt{L}\bot)$ and the $\mathbf{G1h_p}$-derivable the sequents $\Rightarrow \top$. ∎

It is more demanding to establish the admissibility of the more general rules $(\mathtt{R}\rightarrow^+)$ and $(\mathtt{L}\neg\rightarrow^+)$ in $\mathbf{cGh_p}$ in the presence of $(\mathtt{Cut})$. The task is to provide an inductive argument on the depth of the derivation that shows that we can basically always work with sequents in which all the formulas without the relevant connections are already in conditional form.[15]

We call $\Gamma_0; \Gamma_1 \Rightarrow \Delta_0; \Delta_1$ a partition of $\Gamma \Rightarrow \Delta$ if Γ_0 and Γ_1 are (multi-)subsets of Γ and Δ_0 and Δ_1 are (multi-)subsets of Δ with $\Gamma = \Gamma_0 \cup \Gamma_1$ and $\Delta = \Delta_0 \cup \Delta_1$. If in a partition $\Gamma_0; \Gamma_1 \Rightarrow \Delta_0; \Delta_1$ one of the Γ_i is empty we understand it as $\top$ and if one of the Δ_i is empty, then it is $\bot$.

We will make use of the shorthands $\bigwedge \Gamma$ and $\bigvee \Delta$, where if Λ is empty, then $\bigwedge \Lambda$ is just $\top$, and $\bigvee \Lambda$ is just $\bot$. For convenience we use $(\Gamma; \Delta)$ as an abbreviation for the formula $\bigwedge \Gamma \rightarrow \bigvee \Delta$.

---

[14]Thanks to a referee for suggesting the following simplification.

[15]Again we follow the basic strategy of Kashima and Shimura [3].

We first provide an auxiliary Lemma that helps us argue with $(\Gamma; \Delta)$. For this we introduce the class of special sequents defined inductively: The following are special sequents:

$$(\Gamma; \Delta) \Rightarrow (A, \Gamma; \Delta); \qquad\qquad (\Gamma; \Delta) \Rightarrow (\Gamma; \Delta, A) \qquad (9)$$

$$(A, A, \Gamma; \Delta) \Rightarrow (A, \Gamma; \Delta); \qquad (\Gamma; \Delta, A, A) \Rightarrow (\Gamma; \Delta, A) \qquad (10)$$

$$(A, \Gamma; \Delta) \Rightarrow (\neg\neg A, \Gamma; \Delta); \qquad (\Gamma; \Delta, A) \Rightarrow (\Gamma; \Delta, \neg\neg A) \qquad (11)$$

$$(\Gamma; \Delta, A) \Rightarrow (\Gamma; \Delta, A \vee B); \qquad (\Gamma; \Delta, B) \Rightarrow (\Gamma; \Delta, A \vee B) \qquad (12)$$

$$(\neg A, \Gamma; \Delta) \Rightarrow (\neg(A \vee B), \Gamma; \Delta); \qquad (\neg B, \Gamma; \Delta) \Rightarrow (\neg(A \vee B), \Gamma; \Delta) \qquad (13)$$

$$(A, \Gamma; \Delta) \wedge (B, \Pi; \Lambda) \Rightarrow (A \vee B, \Gamma, \Pi; \Delta, \Lambda) \qquad (14)$$

$$(\Gamma; \Delta, \neg A) \wedge (\Pi; \Lambda, \neg B) \Rightarrow (\Gamma, \Pi; \Delta, \Lambda, \neg(A \vee B)) \qquad (15)$$

$$(\Gamma; \Delta, A) \wedge (B, \Pi; \Lambda) \Rightarrow (A \rightarrow B, \Gamma, \Pi; \Delta, \Lambda) \qquad (16)$$

$$(\Gamma; \Delta, \neg B) \wedge (\neg A, \Pi; \Lambda) \Rightarrow (\Gamma, \Pi; \Delta, \Lambda, \neg(A \rightarrow B)) \qquad (17)$$

$$(\Gamma; \Delta, C) \wedge (C, \Pi; \Lambda) \Rightarrow (\Gamma, \Pi; \Delta, \Lambda) \qquad (18)$$

The class of special sequents is also closed under the following inductive clause:

If $A \Rightarrow B$ is a special sequent, then $\neg B \Rightarrow \neg A$ is also a special sequent.

It will be sufficient to use a special form of cut on special sequents:

$$(\texttt{sp-cut}^1) \; \frac{\Gamma \Rightarrow \Delta, A \qquad A \Rightarrow B}{\Gamma \Rightarrow \Delta, B} \; \text{for } A \Rightarrow B \text{ a special sequent}$$

$$(\texttt{sp-cut}^2) \; \frac{\neg A, \Gamma \Rightarrow \Delta \qquad A \Rightarrow B}{\neg B, \Gamma \Rightarrow \Delta} \; \text{for } \neg B \Rightarrow \neg A \text{ a special sequent}$$

$$(\texttt{sp-cut}^3) \; \frac{\Gamma \Rightarrow \Delta, (A, \Pi; \Lambda) \qquad (A, \Pi; \Lambda), A \Rightarrow (\Pi; \Lambda)}{\Gamma, A \Rightarrow \Delta, (\Pi; \Lambda)}$$

$$(\texttt{sp-cut}^4) \; \frac{\neg(\Pi; \Lambda) \Rightarrow \neg(A, \Pi; \Lambda), \neg A \qquad \neg(A, \Pi; \Lambda), \Gamma \Rightarrow \Delta}{\neg(\Pi; \Lambda), \Gamma \Rightarrow \Delta, \neg A}$$

With the derivability of the special sequents as well as $(A, \Pi; \Lambda), A \Rightarrow (\Pi; \Lambda)$ we can reformulate the rules also as:

$$(\text{sp-cut}_1) \ \frac{\Gamma \Rightarrow \Delta, A^i}{\Gamma \Rightarrow \Delta, B^j} \qquad\qquad (\text{sp-cut}_2) \ \frac{(\neg A)^i, \Gamma \Rightarrow \Delta}{(\neg B)^j, \Gamma \Rightarrow \Delta}$$

for $A \Rightarrow B$ a special sequent,   for $\neg B \Rightarrow \neg A$ a special sequent,
where if $\gamma \sim A^i$, then $\gamma \sim B^j$.   where if $(\neg A)^i \sim \delta$, then $(\neg B)^j \sim \delta$.

$$(\text{sp-cut}_3) \ \frac{\Gamma \Rightarrow \Delta, (A, \Pi; \Lambda)^j}{A^i, \Gamma \Rightarrow \Delta, (\Pi; \Lambda)^j} \qquad (\text{sp-cut}_4) \ \frac{(\neg(A, \Pi; \Lambda))^i, \Gamma \Rightarrow \Delta}{(\neg(\Pi; \Lambda))^i, \Gamma \Rightarrow \Delta, (\neg A)^j}$$

where the connections are as to be expected, i.e. in (sp-cut_3) we have $A^i \sim (\Pi; \Lambda)^j$, $\gamma \sim \delta$ in the lower sequent iff $\gamma \sim \delta$ in the upper sequent and $\gamma \sim (\Pi; \Lambda)^j$ iff $\gamma \sim (A, \Pi; \Lambda)^j$. (sp-cut_4) is analogous with $(\neg(\Pi; \Lambda))^i \sim (\neg A)^j$.

LEMMA 17. *The special sequents are cutfree derivable in* $\mathbf{cGh_p}$*, where the cedents are connected. The rules* (sp-cut_h) *for* $1 \leq h \leq 4$ *are admissible in* $\mathbf{cGh_p}$.

PROOF. The base cases are tedious but simple and the inductive clause by admissible contraposition.

We focus on the rules: For (sp-cut_3) we argue as follows: We derive cut-free $(A, \Pi; \Lambda), A \Rightarrow (\Pi; \Lambda)$ and then use

$$\frac{\Gamma \Rightarrow \Delta, (A, \Pi; \Lambda) \qquad \dfrac{\dfrac{\dfrac{A \Rightarrow A \qquad \bigwedge \Pi \Rightarrow \bigwedge \Pi}{A, \bigwedge \Pi \Rightarrow A \wedge \bigwedge \Pi} \qquad \bigvee \Lambda \Rightarrow \bigvee \Lambda}{(A, \Pi; \Lambda), A, \bigwedge \Pi \Rightarrow \bigvee \Lambda}}{(A, \Pi; \Lambda), A \Rightarrow (\Pi; \Lambda)}}{A, \Gamma \Rightarrow \Delta, (\Pi; \Lambda)}$$

For the case (sp-cut_4) we argue by admissible contraposition that we get $\neg(\Pi; \Lambda) \Rightarrow \neg(A, \Pi; \Lambda), \neg A$. ∎

In addition to the special sequents we have the following: We collect some properties for cut-free derivable sequents:

LEMMA 18. *The following sequents are cut-free derivable in* $\mathbf{cGh_p}$*:*

$$(\Pi; \Lambda) \vee A \Rightarrow (\Pi; \Lambda, A) \tag{19}$$

$$\neg(\Pi; \Lambda, A) \Rightarrow \neg((\Pi; \Lambda) \vee A) \tag{20}$$

The following rules are admissible in $\mathbf{cGh_p}$*:*

$$\frac{\Gamma \Rightarrow \Delta, (\Pi; \Lambda), A}{\Gamma \Rightarrow \Delta, (\Pi; \Lambda, A)} \qquad \frac{\neg(\Pi; \Lambda), A, \Gamma \Rightarrow \Delta}{\neg(\Pi; \Lambda, \neg A), \Gamma \Rightarrow \Delta}$$

In the rest of the section we prove the admissibility of $(\text{R} \to^+)$ and $(\text{L}\neg \to^+)$ in $\mathbf{cGh_p}$, which will be a corollary of the following lemma. The idea is to replace all subderivations that have only a single application of $(\text{R} \to^+)$ or $(\text{L}\neg \to^+)$ as the last inference with a derivation without using $(\text{R} \to^+)$ or $(\text{L}\neg\to^+)$. We have to ensure that due to the replacement no new connections are introduced, i.e. we have to insure that all the connections in the new derivation would have been connections at the sequent that would result in an application of $(\text{R} \to^+)$ or $(\text{L}\neg \to^+)$. We will make use of the abbreviation $\Gamma \nsim \Delta$ as short for: for all $\gamma \in \Gamma$ and for all $\delta \in \Delta$, $\gamma \nsim \delta$.

LEMMA 19. *Let $\mathcal{D}$ be a derivation of $\Gamma \Rightarrow \Delta$ in $\mathbf{cGh_p}$ and $\Gamma_0; \Gamma_1 \Rightarrow \Delta_0; \Delta_1$ a partition, such that $\Gamma_1 \nsim \Delta_0$ at $\mathsf{fs}(\mathcal{D})$.*

Let S_1 be a sequent $C^l, \Gamma_0 \Rightarrow \Delta_0, (\Gamma_1; \Delta_1)^i$ that would result in an application of $(\text{R} \to^+)$ and S_2 a sequent $(\neg(\neg\Delta_0; \neg\Gamma_0))^i, \Gamma_1 \Rightarrow \Delta_1, C^l$ that would result in an application of $(\text{L}\neg\to^+)$, then there are derivations $\mathcal{D}'$ and $\mathcal{D}''$ in $\mathbf{cGh_p}$, such that:

(i) $C^l, \Gamma_0 \Rightarrow \Delta_0, (\Gamma_1; \Delta_1)^i$ *is $\mathsf{fs}(\mathcal{D}')$, with*

  if $n \in []^l$ at $\mathsf{fs}(\mathcal{D}')$, then $n \in []^l$ at S_1 and

  if $n \in []^{\gamma_0}$ at $\mathsf{fs}(\mathcal{D}')$, then $n \in []^{\gamma_0}$ at S_1;

(ii) $(\neg(\neg\Delta_0; \neg\Gamma_0))^i, \Gamma_1 \Rightarrow \Delta_1, C^l$ *is $\mathsf{fs}(\mathcal{D}'')$, with*

  if $n \in []^i$ at $\mathsf{fs}(\mathcal{D}'')$, then $n \in []^i$ at S_2 and

  if $n \in []^{\gamma_1}$ at $\mathsf{fs}(\mathcal{D}'')$, then $n \in []^{\gamma_1}$ at S_2.

PROOF. By induction on the depth of the derivation $\mathsf{d}(\mathcal{D})$ we show (i) and (ii).

The cases of initial sequents are simple.

(a) For (ID) we have four partitions to consider:

  1. In the case of $A^i; \Rightarrow A^j;$ for (i) we get $C^l, A^i \Rightarrow A^j, (\top; \bot)$ from (ID) by weakening. For (ii) we get $\neg(\neg A; \neg A) \Rightarrow C^l$ from (ID) with $(\text{L}\neg\to')$.
  2. In the case of $A^i; \Rightarrow; A^j$ we have that Γ_1 is empty. For (i) we get $C^l, A^i \Rightarrow (\top; A^j)$ from (ID) by $(\text{R} \to')$ and weakening. For (ii) we get $(\neg(\neg\bot; \neg A))^k \Rightarrow A^j, C^l$ from (ID) by $(\text{L}\neg\neg)$, (RW) and $(\text{L}\neg\to')$.
  3. The case $; A^i \Rightarrow A^j;$ is not possible by our assumption.
  4. The remaining case is $; A^i \Rightarrow; A^j$. For (i) we get $C^l \Rightarrow (A; A)^k$ from (ID) by $(\text{R} \to')$. For (ii) we get $\neg(\neg\bot; \neg\top), A^i \Rightarrow A^j$ by weakening from (ID).

(b) For $(\text{L}\bot')$ we also have three possible partitions.

1. In the case of $\perp^i; \Rightarrow A^j;$. For (i) we have $C^l, \perp \Rightarrow A, (\top; \perp)^k$ by $(\mathtt{L}\perp')$ and weakening. For (ii) we get $\neg(\neg A; \neg\perp)^k \Rightarrow C^l$ from $\perp \Rightarrow \neg\neg A$ by $(\mathtt{L}\neg\neg)$ and $(\mathtt{L}\neg\rightarrow')$.
2. In the case $\perp^i; \Rightarrow; A^j$ we have $C^l, \perp \Rightarrow (\top; A)^k$ by $(\mathtt{L}\perp')$ and weakening and $\neg(\neg\perp; \neg\perp)^k \Rightarrow A^j, C^l$ by $(\mathtt{ID})$, $(\mathtt{L}\neg\rightarrow')$ and weakening.
3. In the case of $; \perp^i \Rightarrow; A^j$ we have $C^l \Rightarrow (\perp; A)^k$ by $(\mathtt{L}\perp')$ and $(\mathtt{R}\rightarrow')$. $\neg(\neg\perp; \neg\top)^k, \perp^i \Rightarrow A^j$.

(c) The case $(\mathtt{R}\neg\perp')$ is analogous.

We distinguish cases according to $\mathsf{last}(\mathcal{D})$ starting with structural inferences:

(d) In the case of weakening we consider

$$\frac{\Gamma \Rightarrow \Delta}{\Pi, \Gamma \Rightarrow \Delta, \Lambda}$$

with n applications of $(\mathtt{LW})$ and m applications of $(\mathtt{RW})$ and all A_i in Π and all B_j in Λ in the lower sequent have no connections. For simplification we do mention the labels.

We have to show (i) and (ii) for arbitrary permutations $\Pi_0, \Gamma_0; \Pi_1, \Gamma_1 \Rightarrow \Lambda_0, \Delta_0; \Delta_1, \Lambda_1$.

By IH we have the derivations $\mathcal{D}'$ of $C^l, \Gamma_0 \Rightarrow \Delta_0, (\Gamma_1; \Delta_1)^k$ and $\mathcal{D}''$ of $(\neg(\neg\Delta_0; \neg\Gamma_0'))^k, \Gamma_1 \Rightarrow \Delta_1, C^l$.

For (i) we argue as follows: we get $C^l, \Pi_0, \Gamma_0 \Rightarrow \Delta_0, \Lambda_0, (\Gamma_1; \Delta_1)^k$ by weakening from $\mathcal{D}'$ and then $C^l, \Pi_0, \Gamma_0 \Rightarrow \Delta_0, \Lambda_0, (\Pi_1, \Gamma_1; \Delta_1, \Lambda_1)^k$ by internal weakening (9) and $(\mathtt{sp\text{-}cut}_1)$.

For (ii) we argue similarly.

(e)–(f) $(\mathtt{LC})$, $(\mathtt{RC})$ are similar.

The next cases are the propositional ones. The strategy is to mimik the application of the rule with the corresponding special sequents. We discuss only the conditional cases:

(g) $(\mathtt{L}\rightarrow)$: So the last inference is of the form

$$\frac{\begin{array}{cc} \mathcal{D}_0 & \mathcal{D}_1 \\ \Gamma \Rightarrow \Delta, A^i & B^j, \Pi \Rightarrow \Lambda \end{array}}{(A \rightarrow B)^k, \Gamma, \Pi \Rightarrow \Delta, \Lambda}$$

Then we have two partitions of $\mathsf{fs}(\mathcal{D})$ to consider:

1. $(A \to B)^k, \Gamma_0, \Pi_0; \Gamma_1, \Pi_1 \Rightarrow \Delta_0, \Lambda_0; \Delta_1, \Lambda_1$

   For (i) we make a case distinction:

   – if $\Gamma_1 \nsim A^i$ at $\mathsf{fs}(\mathcal{D}_0)$, then there is a partition $\Gamma_0; \Gamma_1 \Rightarrow \Delta_0, A^i; \Delta_1$ of the left upper sequent with $\Gamma_1 \nsim A^i$ and a partition $B^j, \Pi_0; \Pi_1 \Rightarrow \Lambda_0, \Lambda_1$ with $\Pi_1 \nsim \Lambda_0$. So by IH we have

$$
\begin{array}{cc}
\mathcal{D}_0' & \mathcal{D}_1'
\end{array}
$$
$$
\cfrac{\cfrac{C^{l_0}, \Gamma_0 \Rightarrow \Delta_0, A^i, (\Gamma_1; \Delta_1) \qquad B^j, C^{l_1}, \Pi_0 \Rightarrow \Lambda_0, (\Pi_1; \Lambda_1)}{(A \to B)^k, C^{l_0}, C^{l_1}, \Gamma_0, \Pi_0 \Rightarrow \Delta_0, (\Gamma_1; \Delta_1), \Lambda_0, (\Pi_1; \Lambda_1)} \ (\mathtt{L} \to)}{(A \to B)^k, \Gamma_0, \Pi_0, C^l \Rightarrow \Delta_0, \Lambda_0, (\Gamma_1, \Pi_1; \Delta_1, \Lambda_1).}
$$

   The double line inference is due to special sequents for internal weakening and contraction and $(\mathsf{sp\text{-}cut}_1)$.

   The connections are ok since we can trace back the connections of $(A \to B)^k$ to the connections of B^j which by IH will also be connected in the case of the original proof.

   – if $\Gamma_1 \sim A^i$ at $\mathsf{fs}(\mathcal{D}_0)$, then $B^j \nsim \Lambda_0$ at $\mathsf{fs}(\mathcal{D}_1)$ since otherwise $\Gamma_1 \sim \Lambda_0$ at $\mathsf{fs}(\mathcal{D})$.
   If $B^j \sim \Lambda_1$ at $\mathsf{fs}(\mathcal{D}_1)$, then we have

$$
\begin{array}{cc}
\mathcal{D}_0' & \mathcal{D}_1'
\end{array}
$$
$$
\cfrac{\cfrac{\cfrac{C^{l_0}, \Gamma_0 \Rightarrow \Delta_0, (\Gamma_1; \Delta_1, A) \qquad C^{l_1}, \Pi_0 \Rightarrow \Lambda_0, (B, \Pi_1; \Lambda_1)}{C^{l_0}, C^{l_1}, \Gamma_0, \Pi_0 \Rightarrow \Delta_0, \Lambda_0, (\Gamma_1; \Delta_1, A) \wedge (B, \Pi_1; \Lambda_1)}}{C^l, \Gamma_0, \Pi_0 \Rightarrow \Delta_0, \Lambda_0, (A \to B, \Gamma_1, \Pi_1; \Delta_1, \Lambda_1)} \ (\mathsf{sp\text{-}cut}_1)}{(A \to B)^k, C^l, \Gamma_0, \Pi_0 \Rightarrow \Delta_0, \Lambda_0, (\Gamma_1, \Pi_1; \Delta_1, \Lambda_1)} \ (\mathsf{sp\text{-}cut}_3)
$$

   Also in this case the connections are ok. $(A \to B)^k$ is only connected to $(\Gamma_1, \Pi_1; \Delta_1, \Lambda_1)$ at $\mathsf{fs}(\mathcal{D}')$, but B^j was connected to Λ_1 at $\mathsf{fs}(\mathcal{D}_1)$ and so $(A \to B)^k$ would also be connected to $(\Gamma_1, \Pi_1; \Delta_1, \Lambda_1)$ after an $(\mathtt{R} \to^+)$ inference.

   If $B^j \nsim \Lambda_1$ at $\mathsf{fs}(\mathcal{D}_1)$, then we can use Lemma 13 to infer $\Pi \Rightarrow \Lambda$ with a proof of depth less or equal to $\mathcal{D}_1$ and with $\Pi_1 \nsim \Lambda_0$. Then we can use the IH to get $C^l, \Pi_0 \Rightarrow \Lambda_0, (\Pi_1; \Lambda_1)$ and then argue by weakening and the special sequents for weakening.

   For (ii) we consider the cases

   – If $A^i \nsim \Gamma_1$ at $\mathsf{fs}(\mathcal{D}_0)$, then with the IH applied to the partition $\Gamma_0; \Gamma_1 \Rightarrow A^i, \Delta_0; \Delta_1$ we have

$$
\cfrac{\cfrac{\neg(\neg\Lambda_0; \neg\Pi_0, \neg B), \Pi_1 \Rightarrow \Lambda_1, C^{l_0} \qquad \neg(\neg A, \neg\Delta_0; \neg\Gamma_0), \Gamma_1 \Rightarrow \Delta_1, C^{l_1}}{\neg((\neg\Lambda_0; \neg\Pi_0, \neg B) \wedge (\neg A, \neg\Delta_0; \neg\Gamma_0)), \Gamma_1, \Pi_1 \Rightarrow \Delta_1, \Lambda_1, C^l} \ (\mathtt{L}\neg\wedge)}{\neg(\neg\Lambda_0, \neg\Delta_0; \neg\Pi_0, \Gamma_0, \neg(A \to B)), \Gamma_1, \Pi_1 \Rightarrow \Delta_1, \Lambda_1, C^l} \ (\mathsf{sp\text{-}cut}_2)
$$

– if $A^i \sim \Gamma_1$ at $\mathsf{fs}(\mathcal{D}_1)$, then $B^j \not\sim \Lambda_0$ and by IH

$$\dfrac{\neg(\neg\Delta_0; \neg\Gamma_0), \Gamma_1 \Rightarrow \Delta_1, A^i, C^{l_0} \qquad \neg(\neg\Lambda_0; \neg\Pi_0), B^j, \Pi_1 \Rightarrow \Lambda_1, C^{l_1}}{\dfrac{\neg(\neg\Delta_0; \neg\Gamma_0), \neg(\neg\Lambda_0; \neg\Pi_0), \Gamma_1, \Pi_1, (A \to B)^k \Rightarrow \Delta_1, \Lambda_1, C^{l_0}, C^{l_1}}{\neg(\neg\Delta_0, \neg\Lambda_0; \neg\Gamma_0, \neg\Pi_0), \Gamma_1, \Pi_1, (A \to B)^k \Rightarrow \Delta_1, \Lambda_1, C^l}}$$

Using Lemma 18 we get

$$\neg(\neg\Delta_0, \neg\Lambda_0; \neg\Gamma_0, \neg\Pi_0, \neg(A \to B)), \Gamma_1, \Pi_1 \Rightarrow \Delta_1, \Lambda_1, C^l.$$

The connections are ok

2. $\Gamma_0, \Pi_0; (A \to B)^k, \Gamma_1, \Pi_1 \Rightarrow \Delta_0, \Lambda_0; \Delta_1, \Lambda_1$.

By assumption $(A \to B)^k \not\sim \Lambda_0$ at $\mathsf{fs}(\mathcal{D})$ and therefore also $B^j \not\sim \Lambda_0$ at $\mathsf{fs}(\mathcal{D}_1)$.

For (i) similar to the first partition, but simpler. By IH we have

$$\dfrac{C^{l_0}, \Gamma_0 \Rightarrow \Delta_0, (\Gamma_1; \Delta_1, A) \qquad C^{l_1}, \Pi_0 \Rightarrow \Lambda_0, (B, \Pi_1; \Lambda_1)}{\dfrac{C^{l_0}, C^{l_1}, \Gamma_0, \Pi_0 \Rightarrow \Delta_0, \Lambda_0, (\Gamma_1; \Delta_1, A) \wedge (B, \Pi_1; \Lambda_1)}{C^l, \Gamma_0, \Pi_0 \Rightarrow \Delta_0, \Lambda_0, (A \to B, \Gamma_1, \Pi_1; \Delta_1, \Lambda_1)}} \; (\mathsf{sp\text{-}cut}_1)$$

For (ii) we can argue with the IH to get

$$\dfrac{\neg(\neg\Delta_0; \neg\Gamma_0), \Gamma_1 \Rightarrow \Delta_1, A^i, C^{l_1} \qquad \neg(\neg\Lambda_0; \neg\Pi_0), B^j, \Pi_1 \Rightarrow \Lambda_1, C^{l_0}}{\dfrac{\neg(\neg\Delta_0; \neg\Gamma_0) \, , \, \neg(\neg\Lambda_0; \neg\Pi_0), (A \to B)^k, \Gamma_1, \Pi_1 \Rightarrow \Delta_1, \Lambda_1, C^{l_0}, C^{l_1}}{\neg(\neg\Delta_0, \neg\Lambda_0; \neg\Gamma_0, \neg\Pi_0), (A \to B)^k, \Gamma_1, \Pi_1 \Rightarrow \Delta_1, \Lambda_1, C^l}} \; (\mathsf{L} \to)$$

(h) $(\mathsf{R} \to')$: Then we have an inference of the form:

$$\frac{A^i, \Gamma \Rightarrow B^j}{D^l, \Gamma \Rightarrow (A \to B)^k}$$

We have the two partitions:

1. $D^l, \Gamma_0; \Gamma_1 \Rightarrow (A \to B)^k;$. In this case we have $\Gamma_1 \not\sim (A \to B)^k$ in the lower sequent, therefore $\Gamma_1 \not\sim B^j$ in the upper sequent. Then we can use Lemma 13 to infer $A^i, \Gamma_0 \Rightarrow B^j$ and with $(\mathsf{R} \to')$ we get $D^l, \Gamma_0 \Rightarrow (A \to B)^k$.

   For (i) we use weakening to get $C^n, D^l, \Gamma_0 \Rightarrow (A \to B)^k, (\Gamma_1; \bot);$

   For (ii) we use double negation introduction and $(\mathsf{L}\neg \to')$ to infer $\neg(\neg(A \to B); \neg D, \neg\Gamma_0) \Rightarrow C^n$. Then with weakening we get the desired $\neg(\neg(A \to B); \neg D, \neg\Gamma_0), \Gamma_1 \Rightarrow C^n$.

2. If $(A \to B)^k$ is part of Δ_1, then D^l can be part of Γ_0 or Γ_1

   For (i) we can in both cases use $(\mathsf{R} \to')$ and weakening to infer either $C^n, D^l, \Gamma_0 \Rightarrow (\Gamma_1; A \to B)$ or $C^n, \Gamma_0 \Rightarrow (D, \Gamma_1; A \to B)$.

   For (ii) we argue as follows:

If $B^j \sim \Gamma_0$ at the upper sequent, then by IH
$\neg(\neg\bot; \neg\Gamma_0), \Gamma_1, A^i \Rightarrow B^j, B^m$. Contracting and then using $(\mathtt{R} \rightarrow')$ gives
us $\neg(\neg\bot; \neg\Gamma_0), D^l, \Gamma_1 \Rightarrow (A \rightarrow B)^k$ and we can use weakening. The
connections are ok by our assumption.
If $B^j \nsim \Gamma_0$ at the upper sequent, then we use Lemma 13 to infer
$A^i, \Gamma_1 \Rightarrow B^j$ and then $D^l, \Gamma_1 \Rightarrow (A \rightarrow B)^k$ by $(\mathtt{R} \rightarrow')$. We get $\neg(\neg\bot;$
$\neg\Gamma_0), D^l, \Gamma_1 \Rightarrow (A \rightarrow B)^k, C^m$ by weakening.
In the case that D^l is part of Γ_1 we are done. In the other case we use
Lemma 18 to infer $\neg(\neg\bot; \neg\Gamma_0, \neg D), \Gamma_1 \Rightarrow (A \rightarrow B)^k, C^m$.

    The other cases, including $(\mathtt{Cut})$, are similar.   ∎

We sum up the result in the following lemma:[16]

LEMMA 20. *If* $\mathbf{cGh_p^+} \vdash \Gamma \Rightarrow \Delta$, *then* $\mathbf{cGh_p} \vdash \Gamma \Rightarrow \Delta$.

PROOF. Let $\mathcal{D}$ be a derivation of $\Gamma \Rightarrow \Delta$ in $\mathbf{cGh_p^+}$. By the Cut-elimination
Theorem 2 we can assume that it is a cut-free derivation. Let D_i be a
subderivation of $\mathcal{D}$ with the only application of $(\mathtt{R} \rightarrow^+)$ or $(\mathtt{L}\neg \rightarrow^+)$ as last
inference. Then we can use the corresponding derivations given in Lemma
19. In the case $(\mathtt{R} \rightarrow^+)$ we can directly replace the subderivation $\mathcal{D}_i$. In
the case $(\mathtt{L}\neg \rightarrow^+)$ we have to replace the negated conditionals of the form
$\neg(\neg\neg A \rightarrow \neg\neg B)$ by $\neg(A \rightarrow B)$. In both cases the replacement does not
increase the connections and therefore we can successively replace all the
subderivations to get a $\mathbf{cGh_p}$ derivation $\mathcal{D}'$ of $\Gamma \Rightarrow \Delta$.   ∎

THEOREM 3. *The following are equivalent:*

(i) $\top, \Gamma \Rightarrow \Delta, \bot$ *is derivable in* $\mathbf{cGh_p^+}$;

(ii) $\top, \Gamma \Rightarrow \Delta, \bot$ *is cut-free derivable in* $\mathbf{cGh_p^+}$;

(iii) $\top, \Gamma \Rightarrow \Delta, \bot$ *is derivable in* $\mathbf{cGh_p}$;

(iv) $\Gamma \Rightarrow \Delta$ *is derivable in* $\mathbf{G1h_p}$;

(v) $\Gamma \Rightarrow \Delta$ *is derivable in* $\mathbf{G3h_p}$.

---

[16]In Kashima and Shimura [3] the rules of $(\mathtt{sp\text{-}cut})$ are used to establish a partial
cut-elimination. A similar strategy would probably work, if we define a calculus with the
$(\mathtt{sp\text{-}cut})$-rules instead of $(\mathtt{Cut})$. In this case we have make some adjustments, for example
replace the special sequent of internal cut (18) by the following:

$$(\Gamma; \Delta, (A, \Pi; \Lambda)) \Rightarrow (\Gamma, A; \Delta, (\Pi; \Lambda)) \tag{21}$$

$$((\Pi; \Lambda, \neg A), \Gamma; \Delta) \Rightarrow ((\Pi; \Lambda), \Gamma; \Delta, \neg A) \tag{22}$$

and add an additional inductive clause: If $A \Rightarrow B$ is a special sequent, then $(\Gamma; \Delta, A) \Rightarrow$
$(\Gamma; \Delta, B)$ and $(\neg A, \Gamma; \Delta) \Rightarrow (\neg B, \Gamma; \Delta)$ are also a special sequents.

PROOF. From (i) to (ii) by the Cut-elimination Theorem 2. From (ii) to (iii) by Lemma 20. From (iii) to (i) is obvious. The equivalence of (iii) and (iv) by Lemma 16 and the equivalence of (iv) and (v) by Theorem 1.   ∎

5.   Conservativity

We look at one application of the cut-elimination, the conservativity of HYPE over its intuitionistic part as well as its FDE-fragment. In Leitgeb [4] the conservativity considerations are semantic considerations and focus on the intuitionistic part.

First of all we have a subformula property for cut-free derivations for our suitably modified notion of subformula given in Definition 1.

LEMMA 21. *Let $\mathcal{D}$ be a cut-free derivation of $\Gamma \Rightarrow \Delta$ in $\mathbf{cGh_p^+}$ and $\Gamma' \Rightarrow \Delta'$ in $\mathcal{D}$. All the formulas in $\Gamma' \Rightarrow \Delta'$ are subformulas of $\bigwedge \Gamma \to \bigvee \Delta$;*

The cut elimination can be directly employed to establish the conservativity of HYPE over FDE[17] as well as intuitionistic logic:[18]

LEMMA 22. *(Conservativity)*

(i) *For every sequent $\Gamma \Rightarrow \Delta$ derivable in $\mathbf{G3h_p}$ and not containing $\to$, the sequent $\Gamma \Rightarrow \Delta$ is derivable in $\mathbf{G3fde_p}$;*

(ii) *for every sequent $\Gamma \Rightarrow \Delta$ derivable in $\mathbf{G3h_p}$ and not containing $\neg$, the sequent $\Gamma \Rightarrow \Delta$ is derivable in $\mathbf{G3i_p}$.*

PROOF. For (i): if $\Gamma \Rightarrow \Delta$ is derivable in $\mathbf{G3h_p}$ and not containing $\to$, then $\top, \Gamma \Rightarrow \Delta, \bot$ is cut-free derivable in $\mathbf{cGh_p^+}$. Without $\to$ the proof does not use the more general rules so that it is a cutfree proof in $\mathbf{cGh_p}$ without the rules for $\to$. It is an easy exercise to check that it is then derivable in $\mathbf{G3fde_p}$.

for (ii): if $\Gamma \Rightarrow \Delta$ is derivable in $\mathbf{G3h_p}$ and not containing $\neg$, then $\top, \Gamma \Rightarrow \Delta, \bot$ is cut-free derivable in $\mathbf{cGh_p^+}$. Since it does not contain $\neg$, it is derivable in the propositional part of Kashima and Shimura's $\mathbf{CLD}$ in [3]. With their Corollary 4.4 it is derivable in the intuitionistic calculus $\mathbf{G3i_p}$. ∎

There is an interesting aspect for the issue of conservativity. For the axiomatic calculus the conservativity extends to theories conceived as sets

---

[17]For a sequent formulation of FDE compare Anderson and Belnap §17 [1]. The addition of $\bot$ is unproblematic.

[18]Compare Leitgeb [4], p. 380.

of formulas, such that if $\Sigma \vdash_{\mathtt{H}} A$ and Σ, A in $\mathcal{L} \backslash \{\rightarrow\}$, then $\Sigma \vdash_{\mathtt{FDE}} A$.[19] This is not the case if we consider theories as sets of sequents. There are finite sets of sequents $\Gamma_1 \Rightarrow \Delta_1, \ldots, \Gamma_n \Rightarrow \Delta_n, \Gamma \Rightarrow \Delta$ not containing $\rightarrow$, such that there is a derivation of $\Gamma \Rightarrow \Delta$ from the initial sequents $\Gamma_1 \Rightarrow \Delta_1, \ldots, \Gamma_n \Rightarrow \Delta_n$ in HYPE, even so there is no derivation in FDE. An example is given in the arithmetical case by PA in a language expansion $\mathcal{L}^+$ containing a partial predicate, where induction is formulated as a rule. Whereas PA over FDE does not derive $\mathsf{TI}_{\mathcal{L}^+}(\omega^\omega)$, PA over HYPE does derive $\mathsf{TI}_{\mathcal{L}^+}(\omega^\omega)$. In the case of HYPE we can carry out the usual Gentzen proof for PA, but we do not have this option in FDE.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] ANDERSON, A., and N. BELNAP, *Entailment: The Logic of Relevance and Necessity*, Princeton University Press, Princeton, 1975.

---

[19]It might appear strange to talk about axiomatic systems for FDE as usually we do not have FDE-theorems. For finite Γ, however, the correspondence between the axiomatic turnstile and the sequent arrow are obvious and a reformulation of FDE 'axiomatically' is straightforward.

[2] KASHIMA, R., *Cut-Elimination Theorem for the Intermediate Logic CD. Research Reports on Information Sciences: Computer Science*, Tokio Institute of Technology, Meguro City, 1991.

[3] KASHIMA, R., and T. SHIMURA, Cut-elimination theorem for the logic of constant domains, *Mathematical Logic Quarterly* 40:153–172, 1994.

[4] LEITGEB, H., Hype: A system of hyperintensional logic (with an application to semantic paradoxes), *Journal of Philosophical Logic* 48:305–405, 2019.

[5] MOISIL, G., Logique modale, *Disquisitiones Mathematicae et Physica* 2:3–98, 1942.

[6] NEGRI, S., and J. V. PLATO, *Structural Proof Theory*, Cambridge University Press, Cambridge, 2001.

[7] ODINTSOV, S., and H. WANSING, Routley star and hyperintensionality, *Journal of Philosophical Logic* 50:33–56, 2020.

[8] ONO, H., *Proof Theory and Algebra in Logic*, Springer, Berlin, 2019.

[9] SPERANSKI, S. O., Negation as a modality in a quantified setting, *Journal of Logic and Computation*, https://doi.org/10.1093/logcom/exab025, 2021.

[10] TAKEUTI, G., *Proof Theory*, 2nd edn., North Holland, Amsterdam, 1987.

[11] TROELSTRA, A. S., and H. SCHWICHTENBERG, *Basic Proof Theory*, 2nd edn., Cambridge University Press, Cambridge, 2000.

M. FISCHER
Ludwig-Maximilians-Universität München
Munich
Germany
M.Fischer@lrz.uni-muenchen.de