

# Correspondence between Modal Hilbert Axioms and Sequent Rules with an Application to S5<sup>\*</sup>

Björn Lellmann<sup>1</sup> and Dirk Pattinson<sup>1,2</sup>

<sup>1</sup> Department of Computing, Imperial College London

<sup>2</sup> Research School of Computer Science, The Australian National University

**Abstract.** Which modal logics can be ‘naturally’ captured by a sequent system? Clearly, this question hinges on what one believes to be natural, i.e. which format of sequent rules one is willing to accept. This paper studies the relationship between the format of sequent rules and the corresponding syntactical shape of axioms in an equivalent Hilbert-system. We identify three different such formats, the most general of which captures most logics in the S5-cube. The format is based on restricting the context in rule premises and the correspondence is established by translating axioms into rules of our format and vice versa. As an application we show that there is no set of sequent rules of this format which is sound and cut-free complete for S5 and for which cut elimination can be shown by the standard permutation-of-rules argument.

## 1 Introduction

Syntactical descriptions of modal logics often are given in terms of Hilbert calculi. This allows for a simple, intuitive description of the logic and elegant completeness proofs via canonical models. But Hilbert calculi are not optimal to establish e.g. decidability and interpolation for which Sequent systems are far better suited. This raises the question of a precise correspondence between Hilbert and Sequent calculi, and more generally of a classification of the structural, proof-theoretic machinery such as labels, nested sequents and the format of sequent rules that is necessary to give complete (cut-free) sequent systems for a given modal logic. This clearly depends on the format of sequent rules as every logic  $L$  can be captured trivially by the system  $\{\Rightarrow A \mid A \in L\}$  where  $A$  ranges over the theorems of the logic.

Here, we consider rules that are close in spirit to the standard sequent rules for modal logics such as K, KT and S4 and restrict ourselves to pure two-sided sequents with the comma as only structural connective. In particular, we disallow labels [13], additional structural connectives as in Display Logic [10] or nested sequents [3, 15]. Our motivating question is thus *which modal logics can be captured by pure sequent calculi*, or more precisely, *what format of pure sequent rules is necessary to capture a given modal logic?* The rules considered here introduce precisely one layer of modalities and fix context formulae. Within this format, we

---

<sup>\*</sup> Supported by EPSRC-Project EP/H016317/1

consider three classes of rules that differ in the way context formulae are handled. The most general of these formats, *rules with context restrictions*, captures not only most standard normal logics but also a number of non-normal logics such as conditional logics, probabilistic or coalition logic, including all non-iterative logics, and is amenable to reasonably simple purely syntactical criteria sufficient for cut elimination and the subformula property [12].

Our main contribution is twofold. We establish a correspondence between syntactical shapes of axioms in a Hilbert system and rule formats for logical rules in a sequent calculus (with cut) so that both formalisms axiomatise the same logic. We then use this correspondence to obtain impossibility results. Our first such result merely illustrates the techniques and shows that  $K4$  cannot be axiomatised by shallow sequent rules (with cut, and *a posteriori*, also without cut). We then show that there cannot be a cut-free system of rules with context restrictions for  $S5$  where the rules additionally satisfy a permutability condition that we call *mixed cut closure*. This condition allows to permute cuts between principal and context formulae, and is present in virtually all pure sequent systems for modal logics (e.g. [16] or [9] for the dual case of tableau systems). While this does not show that there cannot be a complete and cut-free pure sequent system for  $S5$ , we still obtain a dichotomy result: either,  $S5$  cannot be axiomatised by rules with restrictions and a more general rule format and/or proof theoretic structure is needed, or else the rules of a complete, cut-free system do not allow to permute principal/context cuts (which we consider to be highly unlikely). While clearly more work is needed to establish a complete hierarchy of calculi and rule formats, together with associated impossibility results that we hope to inspire, we are not aware of any other formal impossibility results, or proof-theoretic hierarchy of modal sequent systems to date.

Technically, our methods are purely syntactical and we only require logics to be monotone. Our systems are extensions of the propositional system  $G3cp$  of [16] and we allow contraction throughout so that our results implicitly extend to systems where contraction has been absorbed into the modal rules (e.g.  $S4$  given in *op.cit.*). Rules with restrictions have been used previously in [12] which provides a formal translation from non-nested (Hilbert) axioms to rules and the idea for the translation of nested axioms as well as sufficient syntactic criteria for cut elimination. The formal translation for nested axioms, the converse translation and the impossibility results (for which the translation is crucial) are new.

*Related Work.* Our work is close to [5] in spirit where axioms for substructural logics are translated into (structural) sequent rules and a hierarchy of substructural logics is presented (together with a semantic cut-admissibility proof). For display logic, [10] gives a back-and-forth translation between so-called primitive axioms for normal modal logics and display rules using extra structural connectives and, implicitly, also tense logic. Similarly, [13] uses labels to capture first-order frame conditions over normal modal logics definable by geometric implications. Finally, [6] translates between paraconsistent logics and sequent calculi in a similar way, using different syntactical formats (and constructs cut-free sequent systems via non-deterministic semantics). We are not aware of any

translations between modal axioms and logical rules for pure sequent calculi or any formal impossibility results.

## 2 Preliminaries

We write  $\mathbb{N}$  for the set of non-negative integers and  $\mathfrak{P}(S)$  for the power set of a set  $S$ . We assume familiarity with the standard notions of modal logic as given e.g. in [2, 4]. Throughout the paper we fix a set  $\Lambda$  of *modalities* which we take to be unary for expository reasons, and a countable set  $\mathbf{Var}$  of propositional variables. The set  $\mathcal{F}$  of *formulae* is then defined using the modalities  $\heartsuit \in \Lambda$ , variables  $p \in \mathbf{Var}$  and the boolean connectives  $\perp, \wedge, \vee, \rightarrow$ :

$$\mathcal{F} \ni A ::= p \mid \perp \mid A \vee A \mid A \wedge A \mid A \rightarrow A \mid \heartsuit A$$

As usual we abbreviate  $A \rightarrow \perp$  by  $\neg A$ . Finite sequences  $A_1, \dots, A_n$  of formulae are denoted by  $\mathbf{A}$ . We write  $\bigwedge_{i=1}^n A_i$  or  $\bigwedge \mathbf{A}$  for the iterated conjunction  $A_1 \wedge \dots \wedge A_n$  and similarly for iterated disjunctions. The empty conjunction is  $\perp \rightarrow \perp$  and similarly, the empty disjunctions is  $\perp$ . If  $F$  is a set of formulae we write  $\Lambda(F)$  for the set  $\{\heartsuit A \mid \heartsuit \in \Lambda, A \in F\}$  and  $\mathbf{Prop}(\mathbf{p})$  for the set of propositional formulae in the variables  $\mathbf{p}$ . If  $\sigma : \mathbf{Var} \rightarrow \mathcal{F}$  is a substitution and  $A$  is a formula we write  $A\sigma$  for the result of uniformly substituting every variable in  $A$  according to  $\sigma$ . For a formula  $A$ ,  $\mathbf{Sf}(A)$  is the set of subformulae of  $A$  and  $\mathbf{var}(A)$  for the set of variables occurring in  $A$ . A formula is *rank-1* if every variable occurs under exactly one modality and *non-iterative* if every variable occurs under at most one modality.

If  $F$  is a set of formulae, a *multiset over  $F$*  is a map  $F \rightarrow \mathbb{N}$  with finite support. The notion of union between multisets extends the set-theoretic union in the obvious way and we write  $\Gamma, \Delta$  for the union of the multisets  $\Gamma$  and  $\Delta$ . The set  $\mathcal{S}(F)$  of *sequents over  $F$*  is the set of tuples of multisets over  $F$ , written as  $\Gamma \Rightarrow \Delta$ . Application of a substitution and the notion of the set of subformulae extend to multisets and sequents of formulae in the obvious way. We use the system  $\mathbf{G3cp}$  from [14, 16] with general axioms (see Table 1) for the underlying (classical) propositional logic and denote this system by  $\mathbf{G}$ . As structural rules, we consider  $\mathbf{Cut}$  and  $\mathbf{Con} = \{\mathbf{Con}_L, \mathbf{Con}_R\}$  (see Table 1) and we write  $\mathbf{G}[\mathbf{CutCon}]$  if a result holds for  $\mathbf{G}$  and any extension with  $\mathbf{Cut}$  or  $\mathbf{Con}$ . For  $\mathcal{A} \subseteq \mathcal{F}$  we write  $\mathbf{G}[\mathbf{CutCon}] + \mathcal{A}$  for the sequent system with *groundsequents*  $\Rightarrow A$  for  $A \in \mathcal{A}$ . *Derivability in  $\mathbf{G}[\mathbf{CutCon}] + \mathcal{A}$*  is defined as derivability in  $\mathbf{G}[\mathbf{CutCon}]$  from assumptions  $\{\Gamma \Rightarrow A\sigma, \Delta \mid A \in \mathcal{A}, \Gamma \Rightarrow \Delta \in \mathcal{S}(\mathcal{F}), \sigma \text{ a substitution}\}$  and denoted by  $\vdash_{\mathbf{G}[\mathbf{CutCon}] + \mathcal{A}}$ . In case  $\mathcal{A} = \emptyset$  we also write  $\vdash_{\mathbf{G}[\mathbf{CutCon}]}$ .

We consider modal logics given by a *Hilbert system*, i.e. containing set  $\mathcal{A} \subseteq \mathcal{F}$  of *axioms*, all (classical) propositional tautologies, and closed under modus ponens (from  $A \rightarrow B$  and  $A$  infer  $B$ ) and uniform substitution (from  $A$  infer  $A\sigma$ ) as well as monotonicity (from  $A \rightarrow B$  infer  $\heartsuit A \rightarrow \heartsuit B$ ) for all  $\heartsuit \in \Lambda$ .

## 3 Rules with Restrictions

We briefly recapitulate the notion of a rule with context restrictions from [12].

**Table 1.** Rules for the standard systems  $\mathsf{G}$ ,  $\mathcal{R}_K$  and  $\mathcal{R}_{S5}$  and the structural rules.

$\overline{\Gamma, A \Rightarrow A, \Delta} \mathcal{A}$	$\overline{\Gamma, \perp \Rightarrow \Delta} \perp_L$
$\frac{\Gamma, A \Rightarrow \Delta \quad \Gamma, B \Rightarrow \Delta}{\Gamma, A \vee B \Rightarrow \Delta} \vee_L$	$\frac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A \wedge B \Rightarrow \Delta} \wedge_L$
$\frac{\Gamma \Rightarrow A, B, \Delta}{\Gamma \Rightarrow A \vee B, \Delta} \vee_R$	$\frac{\Gamma \Rightarrow A, \Delta \quad \Gamma \Rightarrow B, \Delta}{\Gamma \Rightarrow A \wedge B, \Delta} \wedge_R$
$\frac{A_1, \dots, A_n \Rightarrow B}{\Gamma, \Box A_1, \dots, \Box A_n \Rightarrow \Box B, \Delta} R_{K_n}$	$\frac{\Gamma, A \Rightarrow \Delta}{\Gamma, \Box A \Rightarrow \Delta} R_T$
$\frac{\Gamma \Rightarrow A, \Delta \quad \Sigma, A \Rightarrow \Pi}{\Gamma, \Sigma \Rightarrow \Delta, \Pi} \text{Cut}$	$\frac{\Gamma, A \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta} \text{Con}_L$
$\frac{\Gamma, B \Rightarrow \Delta \quad \Gamma \Rightarrow A, \Delta}{\Gamma, A \rightarrow B \Rightarrow \Delta} \rightarrow_L$	$\frac{\Gamma, A \Rightarrow B, \Delta}{\Gamma \Rightarrow A \rightarrow B, \Delta} \rightarrow_R$
$\frac{\Box \Gamma \Rightarrow A, \Box \Delta}{\Sigma, \Box \Gamma \Rightarrow \Box A, \Box \Delta, \Pi} R_5$	$\frac{\Gamma \Rightarrow A, A, \Delta}{\Gamma \Rightarrow A, \Delta} \text{Con}_R$
$\mathsf{G} := \{\mathcal{A}, \perp_L, \vee_L, \vee_R, \wedge_L, \wedge_R, \rightarrow_L, \rightarrow_R\}$	
$\mathcal{R}_K := \{R_{K_n} \mid n \geq 0\}$	
$\mathcal{R}_{S5} := \{R_T, R_5\}$	

**Definition 1.** For a set  $F$  of formulae a *context restriction over  $F$*  (or *restriction*) is a tuple  $\langle F_1; F_2 \rangle \in \mathfrak{P}(F)^2$ . For a restriction  $\mathcal{C} = \langle F_1; F_2 \rangle$  and a sequent  $\Gamma \Rightarrow \Delta$  the *restriction of  $\Gamma \Rightarrow \Delta$  according to  $\mathcal{C}$*  is the sequent  $(\Gamma \Rightarrow \Delta)|_{\mathcal{C}} = \Gamma|_{F_1} \Rightarrow \Delta|_{F_2}$  consisting of all those formulae from  $\Gamma$  resp.  $\Delta$  which are substitution instances of formulae in  $F_1$  resp.  $F_2$ . We write  $\mathfrak{C}(F)$  for the set of all restrictions over  $F$ .

**Example 2.** Let  $\Gamma \Rightarrow \Delta = \Box(A \wedge B), C \vee D \Rightarrow \Box E, F$ . Then:

1. for  $\mathcal{C}_\emptyset := \langle \emptyset; \emptyset \rangle$  we have  $(\Gamma \Rightarrow \Delta)|_{\mathcal{C}_\emptyset} = \Rightarrow$
2. for  $\mathcal{C}_{\text{id}} := \langle \{p\}; \{p\} \rangle$  we have  $(\Gamma \Rightarrow \Delta)|_{\mathcal{C}_{\text{id}}} = \Gamma \Rightarrow \Delta$
3. for  $\mathcal{C}_{S5} := \langle \{\Box p\}; \{\Box p\} \rangle$  we have  $(\Gamma \Rightarrow \Delta)|_{\mathcal{C}_{S5}} = \Box(A \wedge B) \Rightarrow \Box E$ .

**Definition 3.** A *rule with context restrictions* (or simply a *rule*) is a tuple  $\mathcal{P}/\Sigma \Rightarrow \Pi$  where  $\mathcal{P} \subseteq \mathcal{S}(\text{Var}) \times \mathfrak{C}(\mathcal{F})$  is the set of *premisses* with associated context restrictions, and  $\Sigma \Rightarrow \Pi \in \mathcal{S}(\Lambda(\text{Var}))$  are the *principal formulae*, such that no variable occurs twice in the principal formulae and every variable occurs in the principal formulae iff it occurs in at least one of the premisses. An *instance* of a rule  $R$  is given by a substitution  $\sigma : \text{Var} \rightarrow \mathcal{F}$  and a *context*  $\Gamma \Rightarrow \Delta \in \mathcal{S}(\mathcal{F})$  and is written as

$$\frac{\{\Gamma|_{F_1}, \Theta\sigma \Rightarrow \Delta|_{F_2}, \Upsilon\sigma \mid (\Theta \Rightarrow \Upsilon; \langle F_1; F_2 \rangle) \in \mathcal{P}\}}{\Gamma, \Sigma\sigma \Rightarrow \Delta, \Pi\sigma}.$$

We assume that every set of rules is closed under injective renaming of variables (respecting the variable conditions) and for every  $\heartsuit \in \mathcal{A}$  includes the *monotonicity rule*  $\text{Mon} = \{(p \Rightarrow q; \mathcal{C}_\emptyset)\} / \heartsuit p \Rightarrow \heartsuit q$ . Notions concerning derivability in  $\mathsf{G}[\text{CutCon}] + \mathcal{A}$  are extended in the obvious way to a set  $\mathcal{R}$  of rules using the notation  $\vdash_{\mathsf{G}[\text{CutCon}]\mathcal{R} + \mathcal{A}}$ . A rule is called *shallow* if the only restrictions occurring in it are  $\mathcal{C}_\emptyset$  or  $\mathcal{C}_{\text{id}}$ . It is *one-step* if only the restriction  $\mathcal{C}_\emptyset$  occurs in it.

**Example 4.** Most standard modal rules fit the format of rules with restrictions:

1.  $R_{K_n}$  is the rule with restrictions  $\{(p_1, \dots, p_n \Rightarrow q; \mathcal{C}_\emptyset)\} / \Box p_1, \dots, \Box p_n \Rightarrow \Box q$
2. The rule  $R_\top$  is the rule with restrictions  $\{(p \Rightarrow ; \mathcal{C}_{id})\} / \Box p \Rightarrow$
3. The rule  $R_5$  is the rule with restrictions  $\{(\Rightarrow p; \mathcal{C}_{S5})\} / \Rightarrow \Box p$ .

The rules  $R_\top$  and  $R_{K_n}$  are also shallow rules, rule  $R_{K_n}$  also is a one-step rule.

By additionally copying the principal formulae into the premisses in the spirit of G3-systems such as the system G3s for S4 in [16] Contraction can be made admissible in cut-free rule sets [12, Thm. 16].

## 4 From Axioms to Rules

We now give an extension and formalisation of the translation from axioms in a Hilbert-system to rules in a sequent system from [11, 12] which uses the following standard argument.

**Lemma 5.** *For every set  $\mathcal{A}$  of axioms and sequent  $\Gamma \Rightarrow \Delta$  we have  $\vdash_{\text{GCutCon}+\mathcal{A}} \Gamma \Rightarrow \Delta$  iff  $\vdash_{\mathcal{H}\mathcal{A}} \bigwedge \Gamma \rightarrow \bigvee \Delta$ .*

The following equivalence between rules and axioms therefore suffices:

**Definition 6.** If  $\mathcal{R}$  is a set of rules, then a set  $\mathcal{A}$  of axioms and a set  $\mathcal{R}_\mathcal{A}$  of rules are *equivalent over  $\text{GR}$*  if for every sequent  $\Gamma \Rightarrow \Delta$  we have  $\vdash_{\text{GRCutCon}+\mathcal{A}} \Gamma \Rightarrow \Delta$  iff  $\vdash_{\text{GRCutCon}\mathcal{R}_\mathcal{A}} \Gamma \Rightarrow \Delta$ .

The main idea for the translation is to take a substitution instance of a non-iterative axiom, where the substitution formulae satisfy certain restrictions, and use the techniques from [11] to turn the non-iterative axiom into a sequent rule. The restrictions guarantee that when substituting the formulae in the rule we get a so-called proto rule, i.e. a rule with a fixed number of context formulae that can be turned into a rule under certain conditions. While this class of axioms might seem restrictive at first, we show in Section 5 that it is sufficient to construct every rule with restrictions. In a first step inspired by [5] we consider conjunctive normal forms of formulae where the polarities of subformulae are controlled.

**Definition 7.** Let  $C_\ell, C, C_r$  be sets of formulae. The sets  $\mathcal{F}_\ell(C_\ell, C, C_r)$  and  $\mathcal{F}_r(C_\ell, C, C_r)$  of *left resolvable* (resp. *right resolvable*) formulae for  $(C_\ell, C, C_r)$  are given by the following grammar with starting variable  $P_\ell$  (resp.  $P_r$ ):

$$\begin{aligned} P_\ell &::= P_\ell \vee P_\ell \mid P_\ell \wedge P_\ell \mid P_r \rightarrow P_\ell \mid A_\ell \mid B \mid \perp & \text{where } A_\ell \in C_\ell, B \in C \\ P_r &::= P_r \vee P_r \mid P_r \wedge P_r \mid P_\ell \rightarrow P_r \mid A_r \mid B \mid \perp & \text{where } A_r \in C_r, B \in C. \end{aligned}$$

Using invertibility of the propositional rules in every calculus containing GCutCon it is then not hard to see that we can decompose a right resolvable formula into clauses resp. sequents in the following way.

**Lemma and Definition 8.** *If  $A$  is a right (resp. left) resolvable formula with context formulae in  $(C_\ell, C, C_r)$ , then  $A$  (resp.  $A \rightarrow \perp$ ) is equivalent to a set  $\mathcal{C}_A$  of clauses of the form  $\bigwedge_{i=1}^n A_i \rightarrow \bigvee_{j=1}^m B_j$  with  $n, m \geq 0$  and  $A_i \in C_\ell \cup C$  and  $B_j \in C_r \cup C$  corresponding to sequents  $A_1, \dots, A_n \Rightarrow B_1, \dots, B_m$ . We then call  $\bigwedge \mathcal{C}_A$  a  $(C_\ell, C, C_r)$ -normal form of  $A$ .  $\square$*

The notions of resolvable formulae allow us now to specify the syntactic requirement for a formula to be expressible as a rule. For conjunctions of such formulae we treat each clause separately.

**Definition 9.** Let  $V \subseteq \text{Var}$  and let  $C_\ell, C_r \subseteq \mathcal{F}$  such that for all formulae  $A, B \in C_\ell \cup C_r$  we have  $\text{var}(A) \cap \text{var}(B) = \emptyset$  and  $\text{var}(A) \cap V = \emptyset$ . A formula is *translatable with context formulae in  $(C_\ell, C_r)$  and variables in  $V$*  if it has the form  $\bigwedge_{i=1}^n A_i \rightarrow \bigvee_{j=1}^m B_j$  with  $A_i \in C_\ell \cup \Lambda(\mathcal{F}_r(C_\ell, C, C_r)) \cup V$  for all  $i \leq n$  and  $B_j \in C_r \cup \Lambda(\mathcal{F}_\ell(C_\ell, C, C_r)) \cup V$  for all  $j \leq m$  and if furthermore every formula in  $C_\ell \cup C_r$  occurs at most once on the top level of the formula (i.e. not in the scope of a modality) and occurs on the top level iff it occurs in the scope of a modality. A formula is *translatable* if it is translatable for some sets  $C_\ell, C_r$  of context formulae and variables  $V$ .

Intuitively, a formula is translatable with context formulae in  $(C_\ell, C_r)$  and variables in  $V$  if it is a substitution instance of a non-iterative formula, where subformulae in  $C_\ell$  (resp.  $C_r$ ) are responsible for the nesting and behave in a way that they can be turned into context formulae on the left (resp. right) of the corresponding rule. The variables in  $V$  are used in the construction of the principal formulae of the rule. A formula might be translatable in more than one way.

**Example 10.** 1. The axiom (4)  $\Box q \rightarrow \Box \Box q$  is translatable with context formulae in  $(\{\Box q\}, \emptyset)$  and variables in  $\emptyset$ .  
 2. The axiom (5)  $\neg \Box q \rightarrow \Box \neg \Box q$  is translatable with context formulae in  $(\emptyset, \{\Box q\})$  or  $(\{\neg \Box q\}, \emptyset)$  and variables in  $\emptyset$ .  
 3. The axiom (T)  $\Box q \rightarrow q$  is translatable with context formulae in  $(\emptyset, \emptyset)$  and variables in  $\{q\}$ .

Since such an axiom contains a fixed number of context formulae, in the first step instead of rules we only get so-called proto-rules, i.e. rules with a fixed context.

**Definition 11.** Given a rule with restrictions  $R = \mathcal{P}/\Sigma \Rightarrow \Pi$  a *proto-rule for  $R$*  is a tuple  $(R; \Gamma \Rightarrow \Delta)$  given by a context  $\Gamma \Rightarrow \Delta \in \mathcal{S}(\mathcal{F})$  such that

1. no propositional variable occurs more than once in  $\Gamma \Rightarrow \Delta$
2. no propositional variable occurs both in  $\Gamma \Rightarrow \Delta$  and  $R$
3. if  $\Gamma \Rightarrow \Delta \neq \Rightarrow$ , then  $(\Gamma \Rightarrow \Delta)|_{\mathcal{C}} \neq \Rightarrow$  for every restriction  $\mathcal{C}$  of  $R$ .

We often leave the context implicit and write  $\widehat{R}$  for a proto-rule for  $R$ . An *instance* of a proto-rule  $\widehat{R} = (R; \Gamma \Rightarrow \Delta)$  is given by a substitution  $\sigma : \text{Var} \rightarrow \mathcal{F}$  and a context  $\Theta \Rightarrow \Xi$  where  $(\Theta \Rightarrow \Xi)|_{\mathcal{C}} = \Rightarrow$  for every restriction  $\mathcal{C}$  of  $R$ , and is the same as the instance of  $R$  with substitution  $\sigma$  and context  $\Gamma\sigma, \Theta \Rightarrow \Delta\sigma, \Xi$  according to Definition 3. Derivability using proto-rules is defined as expected.

Informally, the difference between rules and proto-rules is that in proto-rules the premisses *including the context* are fixed up to substitution, while in rules also the *number* of the context formulae in the premisses may vary.

**Theorem 12.** *Every translatable axiom  $A$  is equivalent over  $\mathbf{GMonCutCon}$  to a proto-rule.*

*Proof (Sketch).* We start with a translatable axiom  $A$ , i.e. a formula of the form

$$\bigwedge_{P \in \mathbf{P}_\ell} \heartsuit_P P \wedge \bigwedge_{q \in \mathbf{q}_\ell} q \wedge \bigwedge_{L \in \mathbf{L}} L \rightarrow \bigvee_{P \in \mathbf{P}_r} \heartsuit_P P \vee \bigvee_{q \in \mathbf{q}_r} q \vee \bigvee_{R \in \mathbf{R}} R$$

where for some  $C_\ell, C_r \subseteq \mathcal{F}$  and  $V \subseteq \mathbf{Var}$  satisfying the restrictions of Definition 9 we have  $\mathbf{P}_\ell \subseteq \mathcal{F}_r(C_\ell, V, C_r)$ ,  $\mathbf{P}_r \subseteq \mathcal{F}_\ell(C_\ell, V, C_r)$  and  $\mathbf{q}_\ell \cup \mathbf{q}_r \subseteq V$  and  $\mathbf{L} \subseteq C_\ell$ ,  $\mathbf{R} \subseteq C_r$ . After turning this into the ground sequent  $\Rightarrow A$  and resolving propositional logic using  $\mathbf{GCutCon}$  we replace the formulae  $P$  under the modalities with fresh variables  $t_P$  and add as premisses all the sequents  $t_P \Rightarrow P$  for  $P \in \mathbf{P}_\ell$  and  $P \Rightarrow t_P$  for  $P \in \mathbf{P}_r$  to get

$$\frac{\{t_P \Rightarrow P \mid P \in \mathbf{P}_\ell\} \quad \{P \Rightarrow t_P \mid t_P \in \mathbf{P}_r\}}{\{\heartsuit_P t_P \mid P \in \mathbf{P}_\ell\}, \mathbf{q}_\ell, \mathbf{L} \Rightarrow \{\heartsuit_P t_P \mid P \in \mathbf{P}_r\}, \mathbf{q}_r, \mathbf{R}}.$$

Now we replace the sequents  $t_P \Rightarrow P$  (resp.  $P \Rightarrow t_P$ ) by the sequents corresponding to the clauses of a  $(C_\ell, C, C_r)$ -normal form of the formulae  $t_P \rightarrow P$  (resp.  $P \rightarrow t_P$ ) according to Lemma 8. The lemma ensures that in this step a context formula ends up on the left (resp. right) hand side of one of these premisses if and only if it is in  $\mathbf{L}$  (resp.  $\mathbf{R}$ ). Now we apply a standard trick and introduce two fresh variables  $r, s$  and replace variables  $q \in \mathbf{q}_\ell$  (resp.  $\mathbf{q}_r$ ) with premisses  $r \Rightarrow q, s$  (resp.  $r, q \Rightarrow s$ ). Finally, we eliminate the occurrences of the variables  $\mathbf{q}_\ell \cup \mathbf{q}_r$  in the premisses by performing all possible cuts between sequents in the premisses with cut formula  $q \in \mathbf{q}_\ell \cup \mathbf{q}_r$ . The resulting proto-rule then is seen to be equivalent to the axiom  $A$ , where for the last step and the fact that premisses ensuring  $t_P \rightarrow P$  instead of  $t_P \leftrightarrow P$  suffice we make use of the monotonicity rules and Cut and the techniques of [12, Lemma 9].  $\square$

In order to produce rules instead of proto-rules we note that in presence of the propositional rules and Cut a rule where a formula  $A$  occurs in the left component of a context restriction is equivalent to a set of proto-rules with a conjunction  $\bigwedge_{i \leq n} A^i$  in place of  $A$ , where  $A^i$  results from  $A$  by renaming the variables to fresh ones, and similarly using disjunction for formulae in the right component. This motivates the next definition.

**Definition 13.** Let the formula  $B$  be translatable with context formulae in  $(\{C_1, \dots, C_n\}, \{D_1, \dots, D_m\})$  and variables in  $V$ . For  $s_1, \dots, s_n, t_1, \dots, t_m \geq 0$  the formula  $B_{s_1, \dots, s_n, t_1, \dots, t_m}$  is constructed from  $B$  by replacing every occurrence of a formula  $C_k$  with  $\bigwedge_{i=1}^{s_k} C_k^i$  and every occurrence of a formula  $D_\ell$  with  $\bigvee_{i=1}^{t_\ell} D_\ell^i$ , where the formulae  $C_k^i, D_\ell^i$  result from  $C_k$  (resp.  $D_\ell$ ) by injectively renaming its variables  $\mathbf{p}$  to fresh variables  $\mathbf{p}^i$ . Then an  $\omega$ -set for  $B$  is a set  $\{B_{s_1, \dots, s_n, t_1, \dots, t_m} \mid s_i \geq 0, t_j \geq 0 \text{ for } i \leq n, j \leq m\}$ .

**Example 14.** The set  $\{\neg \bigvee_{i=1}^n \Box q^i \rightarrow \Box \neg \bigvee_{i=1}^n \Box q^i \mid n \geq 0\}$  is an  $\omega$ -set for the axiom (5)  $\neg \Box q \rightarrow \Box \neg \Box q$ .

Intuitively the formulae in an  $\omega$ -set for  $B$  are constructed from  $B$  by substituting each context formula with finite (possibly empty) conjunctions resp. disjunctions of copies of this formula with fresh variables. By Theorem 12 each of these axioms translates into a proto-rule and since the context formulae are in the same positions together they are equivalent to a rule with restrictions.

**Corollary 15.** *Every  $\omega$ -set for a translatable formula is equivalent over the rule set  $\text{GMonCutCon}$  to a rule with restrictions.*  $\square$

**Example 16.** Translating the  $\omega$ -set  $\{\neg \bigvee_{i=1}^n \Box q^i \rightarrow \Box \neg \bigvee_{i=1}^n \Box q^i \mid n \geq 0\}$  for the axiom (5)  $\neg \Box q \rightarrow \Box \neg \Box q$  yields the rule  $\{(\Rightarrow p; \langle \emptyset; \{\Box p\} \rangle)\} / \Rightarrow \Box p$ .

We would like to translate axioms instead of  $\omega$ -sets of axioms. This is possible if the conjunctions resp. disjunctions can be pushed into the context formulae.

**Definition 17.** A formula  $A$  with free variables  $p_1, \dots, p_n = \mathbf{p}$  is *left normal* for a set  $\mathcal{R}$  of rules if for every  $k \geq 0$  there are formulae  $B_1, \dots, B_n$  such that  $\vdash_{\text{GCutCon}\mathcal{R}} \Rightarrow \bigwedge_{i=1}^k A^i \leftrightarrow A \sigma_{\mathbf{p}}^{\mathbf{B}}$  where  $A^i$  is the result of injectively renaming the propositional variables  $\mathbf{p}$  in  $A$  to fresh variables  $\mathbf{p}^i$  and  $\sigma_{\mathbf{p}}^{\mathbf{B}}$  is the substitution given by  $\sigma(p_j) = B_j$  and  $\sigma(x) = x$  for  $x \notin \mathbf{p}$ . A formula  $A$  is *right normal* if  $A \rightarrow \perp$  is left normal. A context restriction  $\langle F_1; F_2 \rangle$  is *normal* if every formula in  $F_1$  (resp.  $F_2$ ) is left (resp. right) normal.

Adding a translatable axiom where all context formulae are left resp. right normal for  $\mathcal{R}$  is equivalent over  $\mathcal{R}$  to adding its  $\omega$ -set, and thus we get:

**Theorem 18.** *Let  $B$  be a translatable formula with context formulae in  $(C_\ell, C_r)$  and variables in  $V$  such that the formulae in  $C_\ell$  are left normal in  $\mathcal{R}$  and those in  $C_r$  are right normal for  $\mathcal{R}$ . Then  $B$  is equivalent over  $\text{GRCutCon}$  to a rule with restrictions.*  $\square$

Since variables are both left and right normal for every rule set this immediately yields the translation result for non-iterative and rank-1 axioms from [11].

**Corollary 19.** *Every non-iterative (resp. rank-1) axiom is equivalent over the rule set  $\text{GMonCutCon}$  to a set of shallow (resp. one-step) rules.*  $\square$

**Example 20.** 1. The context formula  $\Box q$  is left normal for  $\mathcal{R}_K$  and thus translating the axiom (4)  $\Box q \rightarrow \Box \Box q$  using Theorem 18 yields the well-known rule  $R_4 = \{(\Rightarrow p; \langle \{\Box p\}; \emptyset \rangle)\} / \Rightarrow \Box p$ .

2. Similarly, translating the axiom (T) yields the standard rule  $R_T$  from Table 1.

3. By propositional reasoning and the axioms of  $K$  adding both axioms (4) and (5) is equivalent to adding the set  $\{\bigwedge_{i=1}^n \Box q_\ell^i \wedge \neg \bigvee_{j=1}^m \Box q_r^j \rightarrow \Box (\bigwedge_{i=1}^n \Box q_\ell^i \wedge \neg \bigvee_{j=1}^m \Box q_r^j) \mid m, n \geq 0\}$ , which is an  $\omega$ -set for the axiom  $\Box q_\ell \wedge \neg \Box q_r \rightarrow \Box (\Box q_\ell \wedge \neg \Box q_r)$  under translatability with context formulae in  $(\{\Box q_\ell\}, \{\Box q_r\})$ . By Corollary 15 this set translates into the standard rule  $R_5$  from Table 1.



## 5 From Rules to Axioms

The results of the previous section raise the question whether the format of  $\omega$ -sets for axioms is really necessary. It turns out that the format is both necessary and sufficient in the sense that an axiom can be translated into a rule with restrictions if and only if adding the axiom is equivalent to adding an  $\omega$ -set. We show this by translating rules with restrictions back into  $\omega$ -sets of axioms. The first step is to bring the premisses of the rules into a normal form.

**Lemma and Definition 21.** *Every rule with restrictions is equivalent to a rule in standard form over  $\text{GMonCutCon}$ , i.e. a rule with restrictions  $\mathcal{P}/\Sigma \Rightarrow \Pi$  where*

1. *if  $(\Rightarrow p; \mathcal{C}_\emptyset) \in \mathcal{P}$  then there is no premiss  $(\Gamma, p \Rightarrow \Delta; \mathcal{C}) \in \mathcal{P}$  and no premiss  $(\Rightarrow p; \mathcal{C}) \in \mathcal{P}$  with  $\mathcal{C} \neq \mathcal{C}_\emptyset$*
2. *for all  $q_1, \dots, q_n, p \in \text{Var}$ : if  $(q_1, \dots, q_n \Rightarrow p; \mathcal{C}) \in \mathcal{P}$  and  $(\Rightarrow q_i; \mathcal{C}_i) \in \mathcal{P}$  for all  $i \leq n$ , then  $(\Rightarrow p; \mathcal{C} \cup \bigcup_{i=1}^n \mathcal{C}_i) \in \mathcal{P}$ .*
3. *no variable occurs both on the left hand side of a premiss and on the right hand side of a (possibly different) premiss.*

*Proof (Sketch).* For the first claim if there is are premisses  $(\Rightarrow p; \mathcal{C}_\emptyset)$  and  $(\Gamma, p \Rightarrow \Delta; \mathcal{C})$  in  $\mathcal{P}$ , due to the presence of **Cut** we may replace the latter with  $(\Gamma \Rightarrow \Delta; \mathcal{C})$ . Also, premisses  $(\Rightarrow p; \mathcal{C})$  with  $\mathcal{C} \neq \mathcal{C}_\emptyset$  are derived by (admissible) Weakening from  $(\Rightarrow p; \mathcal{C}_\emptyset)$  and thus can be omitted. For the second claim we simply add the missing premisses, which in the presence of **Cut** yields an equivalent rule. For the last claim we use the fact that all our rule sets include **Mon** and replace a rule by the cut between this rule and **Mon**, see [12, Lemma 9].  $\square$

For the rest of this section we assume w.l.o.g. that all rules are in standard form. Again instead of translating rules directly we first work with sets of proto-rules. Given a proto-rule we now turn its premisses and conclusion into formulae.

**Definition 22.** Let  $R = \mathcal{P}/\Sigma \Rightarrow \Pi$  be a rule and  $\widehat{R}$  a proto-rule for  $R$  given by the context  $\Gamma \Rightarrow \Delta$ . The formulae  $\text{Prem}_{\widehat{R}}$  and  $\text{Concl}_{\widehat{R}}$  are defined by

$$\begin{aligned} \text{Prem}_{\widehat{R}} &= \bigwedge_{(\Theta \Rightarrow \Xi; \langle F_1; F_2 \rangle) \in \mathcal{P}} (\bigwedge \Gamma \upharpoonright_{F_1} \wedge \bigwedge \Theta \rightarrow \bigvee \Xi \vee \bigvee \Delta \upharpoonright_{F_2}) \\ \text{Concl}_{\widehat{R}} &= \bigwedge \Gamma \wedge \bigwedge \Sigma \rightarrow \bigvee \Pi \vee \bigvee \Delta \end{aligned}$$

Then by propositional reasoning it is clear that the premisses of a proto-rule  $\widehat{R}$  (resp. its conclusion) are derivable if and only if the sequent  $\Rightarrow \text{Prem}_{\widehat{R}}$  (resp.  $\Rightarrow \text{Concl}_{\widehat{R}}$ ) is derivable. To turn these formulae into an axiom we make use of the notion of a projective formula, see e.g. [8].

**Definition 23.** A formula  $A \in \mathcal{F}(\Lambda)$  is *projective* if there is a substitution  $\sigma : \text{Var} \rightarrow \mathcal{F}(\Lambda)$  such that  $\vdash_{\text{GMonCutCon}} \Rightarrow A\sigma$ ; and for all  $p \in \text{var}(A)$  we have  $\vdash_{\text{GMonCutCon}} A \Rightarrow p \leftrightarrow p\sigma$ . Such a substitution *witnesses projectivity* of  $A$ .

Given a proto-rule  $\widehat{R}$  once we have a substitution witnessing the projectivity of the formula  $\text{Prem}_{\widehat{R}}$  we are done using the following Lemma.

**Lemma 24.** *If  $\widehat{R}$  is a proto-rule and  $\sigma$  a substitution witnessing projectivity of  $\text{Prem}_{\widehat{R}}$ , then the axiom  $\text{Concl}_{\widehat{R}}\sigma$  is equivalent to  $\widehat{R}$  over every rule set  $\mathcal{R}$ .*

*Proof.* In a first step an induction on the complexity of the formula  $B$  shows that if  $\sigma$  witnesses projectivity of a formula  $A$ , then for every formula  $B$  we have  $\vdash_{\text{GMonCutCon}} A \Rightarrow B \leftrightarrow B\sigma$ . To see that the proto-rule  $\widehat{R}$  is derivable using the axiom assume that we have derivations of its premisses. Then by propositional logic we also have  $\vdash_{\text{GRCutCon}} \Rightarrow \text{Prem}_{\widehat{R}}$ . Thus by projectivity and Cut we get  $\vdash_{\text{GRCutCon}} \Rightarrow \text{Concl}_{\widehat{R}}\sigma \rightarrow \text{Concl}_{\widehat{R}}$  which together with the ground sequent  $\Rightarrow \text{Concl}_{\widehat{R}}\sigma$  yields  $\Rightarrow \text{Concl}_{\widehat{R}}$ . Now resolving the propositional connectives using  $\text{GCutCon}$  yields the conclusion of  $R$ . For the other direction by projectivity we have  $\vdash_{\text{GRCutCon}} \Rightarrow \text{Prem}_{\widehat{R}}\sigma$ , and resolving the propositional connectives and applying the rule  $R$  and propositional rules yields  $\vdash_{\text{GRCutCon}R} \Rightarrow \text{Concl}_{\widehat{R}}\sigma$ .  $\square$

Using standard techniques [8] we can always construct such a substitution:

**Definition 25** ( $\theta$ ). Let  $R = \mathcal{P}/\Sigma \Rightarrow \Pi$  be a rule in standard form and  $\widehat{R}$  a proto-rule for  $R$  given by  $\Gamma \Rightarrow \Delta$ . Define the substitution  $\theta_{\widehat{R}}$  by

$$\theta_{\widehat{R}}(p) := \begin{cases} \top & : (\Rightarrow p; \mathcal{C}_\emptyset) \in \mathcal{P} \\ \text{Prem}_{\widehat{R}} \rightarrow p & : (\Theta \Rightarrow p, \Xi; \mathcal{C}) \in \mathcal{P} \text{ for some } \mathcal{C} \neq \mathcal{C}_\emptyset \text{ and } \Theta \Rightarrow \Xi \\ \text{Prem}_{\widehat{R}} \wedge p & : (\Theta, p \Rightarrow \Xi; \mathcal{C}) \in \mathcal{P} \text{ for some } \Theta \Rightarrow \Xi \\ p & : \text{otherwise.} \end{cases}$$

**Lemma 26.**  *$\theta_{\widehat{R}}$  witnesses projectivity of  $\text{Prem}_{\widehat{R}}$  if the latter is satisfiable.*

The proof is by standard propositional reasoning and gives the desired translation when combined with Lemma 24.

**Theorem 27.** *Every proto-rule is equivalent to a translatable axiom.*

*Proof.* By Lemmata 24 and 26 we get equivalence of  $\text{Concl}_{\widehat{R}}\theta_{\widehat{R}}$  and  $\widehat{R}$ . Since by Lemma 21 the rule  $R$  was w.l.o.g. in standard form, every variable occurs either only on the left or on the right of premisses and conclusion, which ensures that the axiom  $\text{Concl}_{\widehat{R}}\theta_{\widehat{R}}$  is translatable.  $\square$

Moreover, since a rule  $R$  is equivalent to the set of proto-rules for  $R$  we can translate rules into sets of axioms.

**Proposition 28.** *Every rule  $R$  is equivalent to an  $\omega$ -set for an axiom  $A_R$ . If all context restrictions of  $R$  are normal, then it is equivalent to a single axiom  $A_R$ . If  $R$  is a shallow (resp. one-step) rule, then it is equivalent to a non-iterative (resp. rank-1) axiom.*

*Proof.* It is not hard to see that a rule  $R$  is equivalent to the set of proto-rules for it and that the set of translations of these proto-rules is an  $\omega$ -set. In case all the restrictions are normal, adding the  $\omega$ -set for the corresponding axiom  $A_R$  is equivalent to adding the axiom  $A_R$  itself. A close inspection of the translations of shallow and one-step rules together with the fact that all restrictions for such rules are normal yields the last statement.  $\square$

In conclusion we thus have the following correspondences between rules and axioms for monotone modalities and normal context formulae resp. restrictions:

rules with context restrictions	$\leftrightarrow$	translatable axioms
shallow rules	$\leftrightarrow$	non-iterative axioms
one-step rules	$\leftrightarrow$	rank-1 axioms

In the first case if not all context formulae are normal the correspondences hold only between rules and  $\omega$ -sets for translatable axioms.

**Example 29.** We apply the translation procedure to the rule  $R_5$  from Table 1. Proto-rules for  $R_5$  are given by sequents  $\Box r_1, \dots, \Box r_n \Rightarrow \Box s_1, \dots, \Box s_m$  with  $m, n \geq 0$ . Thus we get the formulae  $\text{Prem}_{\widehat{R}_5} = \bigwedge_{i=1}^n \Box r_i \rightarrow p \vee \bigvee_{j=1}^m \Box s_j$  and  $\text{Concl}_{\widehat{R}_5} = \bigwedge_{i=1}^n \Box r_i \rightarrow \Box p \vee \bigvee_{j=1}^m \Box s_j$  and the substitution  $\theta_{\widehat{R}_5}$  with  $\theta_{\widehat{R}_5}(p) = \text{Prem}_{\widehat{R}_5} \rightarrow p$  as well as  $\theta_{\widehat{R}_5}(r_i) = r_i$  and  $\theta_{\widehat{R}_5}(s_j) = s_j$ . Thus by Lemmata 24 and 26 the rule  $R_5$  is equivalent to the set of axioms

$$\left\{ \bigwedge_{i=1}^n \Box r_i \rightarrow \Box \left( \left( \bigwedge_{i=1}^n \Box r_i \rightarrow p \vee \bigvee_{j=1}^m \Box s_j \right) \rightarrow p \right) \vee \bigvee_{j=1}^m \Box s_j \mid m, n \geq 0 \right\},$$

which is an  $\omega$ -set for the axiom  $\Box r \rightarrow \Box((\Box r \rightarrow p \vee \Box s) \rightarrow p) \vee \Box s$  under translatability with context formulae in  $(\{\Box r\}, \{\Box s\})$  and variables in  $\{p\}$ . By propositional reasoning and normality of  $\Box$  this  $\omega$ -set is moreover equivalent to axioms of the form  $\Box r \rightarrow \Box((\Box r \wedge \neg \bigvee_{j=1}^m \Box s_j) \vee p) \vee \bigvee_{j=1}^m \Box s_j$ , and since  $\Box$  is monotone we may first replace  $p$  by  $\top$  and then omit it, yielding axioms  $\Box r \rightarrow \Box(\Box r \wedge \neg \bigvee_{j=1}^m \Box s_j) \vee \bigvee_{j=1}^m \Box s_j$ . Finally, these axioms are equivalent over  $\mathbf{K}$  to axioms (4)  $\Box r \rightarrow \Box \Box r$  and  $\{\neg \bigvee_{j=1}^m \Box s_j \rightarrow \Box \neg \bigvee_{j=1}^m \Box s_j \mid m \geq 0\}$ , where the latter is an  $\omega$ -set for the axiom (5), which over  $\mathbf{K}$  is equivalent to (5) itself.

Similarly, rule  $R_{\top}$  translates into the axiom  $r \wedge \Box((r \wedge p \rightarrow s) \wedge p) \rightarrow s$ . Again, by arguments as above this axiom can be seen to be equivalent (as an axiom) over  $\mathbf{K}$  to the standard axiom (T)  $\Box A \rightarrow A$ .

**Remark 30.** The translations show that Hilbert-style axioms actually correspond to proto-rules instead of rules. Thus from this perspective it would be more natural to consider sequent systems with proto-rules instead of rules.

## 6 Application: Limitative Results for K4 and S5

The correspondence results of the previous sections open up new possibilities for investigations into the expressive power of sequent rules of a certain format by investigating the limits of the corresponding class of axioms. We exemplify this by establishing two results: the logic K4 cannot be captured by shallow rules (not even with cut), and S5 cannot be captured in a cut-free way by rule with context restrictions with an additional permutation property. Note that the situation for both logics is different: K4 is not captured by any set of shallow rules, with or without cut. For S5 we have exhibited a set of rules that, together with cut,

is sound and complete for **S5**. Here, we show that there cannot be a cut-free complete set of rules which allows permutations of principal/context cuts into the premisses of the first rule. While the result for **K4** is intuitively obvious (but still non-trivial to establish formally), our result on **S5** can be interpreted to say that rules with restrictions for **S5** must be rather exotic, as they do not satisfy permutability which is present in virtually all cut-free rule sets. While this also implies that ‘standard’ proofs of cut-elimination that rely on propagating cut upwards fail, our considerations are independent of any such proof and just use permutability of rules. For this section we drop the assumption that the rule sets include the rule **Mon**.

**Theorem 31.** *There is no set  $\mathcal{R}$  of shallow or one-step rules such that  $\text{GRCutCon}$  is sound and complete for **K4**.*

*Proof.* If there was such a set  $\mathcal{R}$  of rules, then also  $\text{GRMonCutCon}$  would be sound and complete for **K4**. Then by Proposition 28 we would have a non-iterative axiomatisation  $\mathcal{A}$  of **K4**. Now consider the two frames defined by  $\mathfrak{F}_1 = (\{a, b, c\}, \{(a, b), (b, c), (a, c)\})$  and  $\mathfrak{F}_2 = (\{1, 2, 3, 4\}, \{(1, 2), (2, 3), (1, 4)\})$ . Clearly  $\mathfrak{F}_1$  is transitive while  $\mathfrak{F}_2$  is not. But given a non-iterative formula which is satisfied in one of the two frames we can find a world of the other frame and an assignment such that the formula is satisfied. That is, precisely the same non-iterative formulae are valid on both frames. On the other hand it is well-known that **K4** is the logic of transitive Kripke-frames which is modally definable by the formula  $(4)\Box p \rightarrow \Box\Box p$ . But then (4) must be derivable in  $\mathcal{HA}$  and would therefore be valid in  $\mathfrak{F}_2$ , a contradiction since  $\mathfrak{F}_2$  is not transitive.  $\square$

Along these lines it is also possible to show that shallow rules are not expressive enough to characterise symmetric frames or that one-step rules cannot capture reflexivity, thus establishing that the inclusions between the different formats are proper. The situation becomes more interesting if we want to show limitative results for rules with restrictions in general. It is well known that the properties mentioned above together characterise the logic **S5**, and there is a set of rules which together with the cut rule captures this logic. On the other hand every known cut-free sequent system for **S5** seems to involve some extensions of the rule format to facilitate a cut admissibility proof (see e.g. [15] for an overview). This seems to suggest that there is no set of rules with restrictions which is sound and cut-free complete for **S5**. As argued above, our translation implicitly involves the cut rule which is problematic for proving results about cut-free systems per se. On the other hand assuming a certain permutability of rules gives us more information about the rule set. The first step is to show that every rule set which is sound and cut-free complete for **S5** must include certain rules.

**Lemma 32.** *If  $\mathcal{R}$  is a set of modal rules such that  $\text{GRCon}$  is sound and complete for **S5**, then there are rules  $R_1 = \mathcal{P}_1/\Sigma_1 \Rightarrow \Pi_1$  and  $R_2 = \mathcal{P}_2/\Sigma_2 \Rightarrow \Pi_2$  with*

1.  $\Box p \in \Sigma_1$  and  $(\Rightarrow p)|_{\mathcal{C}} \Rightarrow p$  for a restriction  $\mathcal{C}$  of  $R_1$
2.  $\Box p \in \Sigma_2$  and  $\Box q \in \Pi_2$ ; or  $\Box q \in \Pi_2$  and  $(\Box p \Rightarrow)|_{\mathcal{C}} \Box p \Rightarrow$  for a restriction  $\mathcal{C}$  of  $R_2$ .

*Proof.* For the existence of  $R_1$  we inspect all possible derivations of the S5-valid sequent  $\Box p \Rightarrow p$ . In the last applied modal rule  $\Box p \Rightarrow$  must have been principal due to the rule format. But then  $(\Rightarrow p)|_{\mathcal{C}} = \Rightarrow p$  for a restriction  $\mathcal{C}$  of this rule since otherwise the sequent  $\Box p \Rightarrow$  would be derivable, in contradiction to the fact that  $\Box p \Rightarrow \perp$  is not S5-valid.

For the existence of  $R_2$  we consider the possible derivations of the S5-valid sequent  $\Box p \Rightarrow \Box(p \vee q)$ . If in such a derivation  $\Rightarrow \Box(p \vee q)$  is never principal, then since no rule decreases the complexity of a formula and since context formulae cannot change sides, every occurrence of  $\Box(p \vee q)$  must come from the weakening context of an axiom. But then essentially the same derivation is used to derive  $\Box p \Rightarrow$ , a contradiction since  $\Box p \Rightarrow \perp$  is not S5-valid. Thus we have a rule where  $\Rightarrow \Box(p \vee q)$  is principal. But then for some  $\Gamma \Rightarrow \Delta \in \{\Box p \Rightarrow, p \Rightarrow, \Rightarrow p\}$  we must have  $(\Gamma \Rightarrow \Delta)|_{\mathcal{C}} = \Gamma \Rightarrow \Delta$  for some restriction  $\mathcal{C}$  of this rule. In the first two cases we are done, and in the last case we can derive  $\Rightarrow p \vee \Box(p \vee q)$ , again a contradiction since the corresponding formula is not S5-valid.  $\square$

The additional permutation property which we demand of the rule sets is suggested by standard cut elimination proofs in the spirit of [7] which rely on a permutation-of-rules argument to transform a derivation with cut into one without. It demands essentially that cuts where the cut formula is principal in the last applied rule in the derivation of one of the premisses and contextual in that of the other premiss can be permuted into the premisses of the latter rule. If the main connective of the cut formula is propositional we also consider an alternative property, *G-invertibility*, which essentially ensures that we can show invertibility of the propositional connectives by permuting applications of the propositional rules below applications of the modal rules. It should be noted that while these properties are motivated by a certain proof technique, they are properties of *rule sets* and as such independent of the proof technique.

**Definition 33.** A rule set  $\mathcal{R}$  is *mixed-cut closed* if whenever  $R, Q \in \mathcal{R}$  and  $R$  has principal formulae  $\Gamma \Rightarrow \Delta, A$  (resp.  $A, \Gamma \Rightarrow \Delta$ ) such that for some restriction  $\mathcal{C}_Q$  of  $Q$  we have  $(\Rightarrow A)|_{\mathcal{C}_Q} = \Rightarrow A$  (resp.  $(A \Rightarrow)|_{\mathcal{C}_Q} = A \Rightarrow$ ), then  $(\Gamma \Rightarrow \Delta)|_{\mathcal{C}_Q} = \Gamma \Rightarrow \Delta$  and for every sequent  $\Sigma \Rightarrow \Pi$  and restriction  $\mathcal{C}_R$  of  $R$  we have  $(\Sigma \Rightarrow \Pi)|_{\mathcal{C}_R}|_{\mathcal{C}_Q} = (\Sigma \Rightarrow \Pi)|_{\mathcal{C}_R}$ . A rule set  $\mathcal{R}$  is *G-inverting* if for every restriction  $\langle F_0; F_1 \rangle$  of a rule in  $\mathcal{R}$  and  $i \in \{0, 1\}$  we have: whenever  $A \circ B \in F_i$  with  $\circ \in \{\wedge, \vee\}$ , then also  $A, B \in F_i$  and whenever  $A \rightarrow B \in F_i$ , then also  $A \in F_{1-i}$  and  $B \in F_i$ .

Using the two rules from Lemma 32 together with these properties we obtain a restricted format of the rules in such a rule set.

**Lemma 34.** *Let  $\mathcal{R}$  be a mixed-cut closed set of modal rules such that GRCon is sound and complete for S5 and such that GR is mixed-cut closed or  $\mathcal{R}$  is G-inverting. Then w.l.o.g. for every restriction  $\langle F_1; F_2 \rangle$  of a rule in  $\mathcal{R}$  we have  $F_1 \subseteq \{\Box p, p\}$  and  $F_2 \subseteq \{p\}$ .*

*Proof.* Suppose GR is mixed-cut closed. If e.g. for a restriction  $\langle F_1; F_2 \rangle$  of a rule in  $\mathcal{R}$  we have  $A \wedge B \in F_1$ , then by mixed-cut closure of GR applied to this rule

and the rule  $\wedge_R$  we also have  $p \in F_1$  and  $p \in F_2$  and thus w.l.o.g.  $F_1 = F_2 = \{p\}$ . Similarly for the other propositional connectives. If  $\Box A \in F_1$ , then using mixed-cut closure with  $R_2$  from the Lemma we have  $\Box p \in F_1$ . Similarly, if  $\Box A \in F_2$ , then also  $p \in F_2$  using  $R_1$  from the Lemma. If on the other hand  $\mathcal{R}$  is  $\mathbf{G}$ -inverting, then in a derivation in  $\mathbf{GRCon}$  we may permute applications of rules from  $\mathbf{G}$  below applications of rules from  $\mathcal{R}$  in such a way that in the new derivation only variables or modalised formulae occur as contextual formulae of modal rules. Furthermore using mixed-cut closure of  $\mathcal{R}$  and the rules  $R_1, R_2$  as above we may replace the modalised formulae in the restrictions by modalised variables. Thus w.l.o.g. we may replace  $\mathcal{R}$  by an equivalent set  $\mathcal{R}'$  of rules where only formulae of the desired format occur in the restrictions. This preserves mixed-cut closure of  $\mathcal{R}$  and  $\mathbf{G}$ -invertibility.  $\square$

Now we are in a position to use techniques similar to the ones used in proving Theorem 31 to show that the translations of rules in the format specified by the preceding Lemma cannot characterise  $\mathbf{S5}$ .

**Theorem 35.** *There is no mixed-cut closed set  $\mathcal{R}$  of modal rules such that  $\mathbf{GRCon}$  is sound and complete for  $\mathbf{S5}$  and such that  $\mathbf{GR}$  is mixed-cut closed or  $\mathcal{R}$  is  $\mathbf{G}$ -inverting.*

*Proof.* Adding the rules **Cut** and **Mon** to the rule set  $\mathcal{R}$  preserves soundness and completeness. Moreover, bringing the rule set into standard form preserves the restricted format of the rules guaranteed by Lemma 34. The translations of proto-rules for rules of this format have the form

$$q \wedge \Box c \wedge P \wedge \bigwedge_{i=1}^n \Box A_i \rightarrow \bigvee_{j=1}^m \Box B_j \vee r ,$$

where  $P$  is a conjunction of variables and negated variables, and the  $A_i$  and  $B_j$  are propositional formulae in the variables of  $P$ , the variables  $q, r$  and the formula  $\Box c$ . To see that such axioms are not sufficient to characterise  $\mathbf{S5}$  consider the two frames  $\mathfrak{F}_1 := (\mathbb{N}, \mathbb{N}^2)$  and  $\mathfrak{F}_2 := (\mathbb{N}, \leq)$ . Then for every such formula  $A$ , world  $n$  of  $\mathfrak{F}_2$  and valuation  $\sigma : \mathbb{N} \rightarrow \mathfrak{P}(\text{Var})$  with  $\mathfrak{F}_2, n, \sigma \Vdash \neg A$  we construct a valuation  $\tau$  on  $\mathfrak{F}_1$  by setting  $\tau(m) := \sigma(m + n)$ . Now it is not too difficult to check that  $\mathfrak{F}_1, 0, \tau \Vdash \neg A$  as well. Similarly every such formula which is satisfied in  $\mathfrak{F}_1$  can also be satisfied in  $\mathfrak{F}_2$ , and thus the same formulae of this format are valid in the two frames. But  $\mathfrak{F}_1$  is an  $\mathbf{S5}$ -frame, while  $\mathfrak{F}_2$  is not. As the class of  $\mathbf{S5}$ -frames is modally definable [2] the result follows as in Theorem 31.  $\square$

## 7 Discussion

Our goal in this paper was to establish a taxonomy of rule formats, together with a methodology for obtaining limitative results on the existence of sequent calculi for particular modal logics. The main application was a formal proof of non-existence of cut-free sequent calculus for  $\mathbf{S5}$  in the most general format

considered satisfying a (reasonably mild) additional permutation property. Since as well as being instrumental for syntactic proofs this property is present in virtually all cut-free systems, this strongly suggests that a cut-free calculus for  $S5$  will require additional machinery in the rule format or a very different, possibly semantic proof of cut admissibility. But even in the latter case, the rule set would need to violate permutability which we consider highly unlikely.

Clearly, these early results offer much potential for refinement, as e.g. rules with restrictions might be considered too restrictive. Other formats, e.g. using context relations [1], capture more logics but often permit trivial cut-free systems with all theorems of a logic as axioms so that further work is needed to map out this hitherto uncharted landscape. Presently, we are considering a more relaxed rule format where context formulae are permitted to change sides, and investigate applications to provability logic.

## References

1. Avron, A., Lahav, O.: Kripke semantics for basic sequent systems. In: Brünnler, K., Metcalfe, G. (eds.) *TABLEAUX 2011*. LNAI, vol. 6793, pp. 43 – 57. Springer-Verlag Berlin Heidelberg (2011)
2. Blackburn, P., de Rijke, M., Venema, Y.: *Modal Logic*. Cambridge University Press (2001)
3. Brünnler, K.: Deep sequent systems for modal logic. *Arch. Math. Log.* 48, 551–577 (2009)
4. Chellas, B.F.: *Modal Logic*. Cambridge University Press (1980)
5. Ciabattoni, A., Galatos, N., Terui, K.: Algebraic proof theory for substructural logics: Cut-elimination and completions. *Ann. Pure Appl. Logic* 163, 266–290 (2012)
6. Ciabattoni, A., Lahav, O., Spendier, L., Zamansky, A.: Automated support for the investigation of paraconsistent and other logics. In: Artemov, S., Nerode, A. (eds.) *LFCS 2013*. LNCS, vol. 7734, pp. 119–133. Springer-Verlag Berlin Heidelberg (2013)
7. Gentzen, G.: Untersuchungen über das logische Schließen. I. *Math. Z.* 39(2), 176 – 210 (1934)
8. Ghilardi, S.: Unification in intuitionistic logic. *J. Symb. Log.* 64(2), 859–880 (1999)
9. Goré, R.: Cut-free sequent and tableau systems for propositional diodean modal logics. *Studia Logica* 53, 433–457 (1994)
10. Kracht, M.: Power and weakness of the modal display calculus. In: Wansing, H. (ed.) *Proof Theory of Modal Logic*, pp. 93–121. Kluwer (1996)
11. Lellmann, B., Pattinson, D.: Cut elimination for shallow modal logics. In: Brünnler, K., Metcalfe, G. (eds.) *TABLEAUX 2011*. LNAI, vol. 6793, pp. 211–225. Springer-Verlag Berlin Heidelberg (2011)
12. Lellmann, B., Pattinson, D.: Constructing cut free sequent systems with context restrictions based on classical or intuitionistic logic. In: Lodaya, K. (ed.) *ICLA 2013*. LNAI, vol. 7750, pp. 148–160. Springer-Verlag Berlin Heidelberg (2013)
13. Negri, S.: Proof analysis in modal logic. *J. Philos. Logic* 34, 507–544 (2005)
14. Negri, S., von Plato, J.: *Structural proof theory*. Cambridge University Press (2001)
15. Poggiolesi, F.: *Gentzen Calculi for Modal Propositional Logic*, Trends In Logic, vol. 32. Springer-Verlag Berlin Heidelberg (2011)
16. Troelstra, A.S., Schwichtenberg, H.: *Basic Proof Theory*, Cambridge Tracts In Theoretical Computer Science, vol. 43. Cambridge University Press, 2 edn. (2000)

## A Additional Details for the Proofs of Theorem 12 and Lemma 21

These proofs makes use of the following notions and results from [12] which we include for convenience:

**Definition 36.** If  $\mathcal{P} \subseteq \mathcal{S}(\text{Var}) \times \mathfrak{C}(\mathcal{F})$  is a set of premisses with context restrictions, then for  $p \in \text{Var}$  the *p-elimination* of  $\mathcal{P}$  is the set

$$\mathcal{P} \ominus p := \{ (\Gamma, \Sigma \Rightarrow \Delta, \Pi; \mathcal{C}_1 \cup \mathcal{C}_2) \mid (\Gamma \Rightarrow \Delta, p; \mathcal{C}_1) \in \mathcal{P}, (p, \Sigma \Rightarrow \Pi; \mathcal{C}_2) \in \mathcal{P} \} \\ \cup \{ (\Gamma \Rightarrow \Delta; \mathcal{C}) \mid (\Gamma \Rightarrow \Delta; \mathcal{C}) \in \mathcal{P}, p \notin \Gamma, \Delta \},$$

where for restrictions  $\mathcal{C}_1 = \langle F_1, F_2 \rangle$  and  $\mathcal{C}_2 = \langle G_1, G_2 \rangle$  we write  $\mathcal{C}_1 \cup \mathcal{C}_2$  for  $\langle F_1 \cup G_1, F_2 \cup G_2 \rangle$ . Iterated elimination of variables  $\mathbf{p} = p_1, \dots, p_n$  is denoted by  $\mathcal{P} \ominus \mathbf{p}$ . For rules  $R = (\mathcal{P}_R; \Gamma \Rightarrow \Delta, \heartsuit \mathbf{p})$  and  $Q = (\mathcal{P}_Q; \heartsuit \mathbf{p}, \Sigma \Rightarrow \Pi)$  the *cut between R and Q on  $\heartsuit \mathbf{p}$*  is the rule  $\text{cut}(R, Q, \heartsuit \mathbf{p}) := (\mathcal{P}_R \cup \mathcal{P}_Q) \ominus \mathbf{p} / \Gamma, \Sigma \Rightarrow \Delta, \Pi$ .

In the presence of the propositional rules and Cut we get that performing a cut on two rules preserves soundness:

**Lemma 37.** *For every two rules  $R_1, R_2$  the rule  $\text{cut}(R_1, R_2, \heartsuit p)$  is derivable in  $\text{GCutCon}R_1R_2$ .*

*Proof.* For every set  $S$  of sequents and variable  $p$  let

$$S \ominus p := \{ \Gamma, \Sigma \Rightarrow \Delta, \Pi \mid \Gamma \Rightarrow \Delta, p \in S, p, \Sigma \Rightarrow \Pi \in S \} \cup \{ \Gamma \Rightarrow \Delta \mid p \notin \Gamma, \Delta \}$$

The proof hinges on the fact that for every set  $S$  of sequents and variable  $p$  there is a formula  $A_{S,p}$  such that

1. if  $\Gamma \Rightarrow \Delta, p \in S$ , then  $\Gamma \Rightarrow \Delta, A_{S,p}$  is derivable in  $\text{GCutCon}$  from assumptions  $S \ominus p$
2. if  $p, \Gamma \Rightarrow \Delta \in S$ , then  $A_{S,p}, \Gamma \Rightarrow \Delta$  is derivable in  $\text{GCutCon}$  from assumptions  $S \ominus p$ .

This can be seen by e.g. setting  $A_{S,p} := \bigvee_{\Gamma \Rightarrow \Delta, p \in S} (\bigvee \Gamma \vee \neg \bigwedge \Delta)$ . Then for every  $\Gamma \Rightarrow \Delta, p \in S$  the sequent  $\Gamma \Rightarrow \Delta, A_{S,p}$  is derivable using  $\text{GCutCon}$ , and for every  $p, \Sigma \Rightarrow \Pi \in S$  the sequent  $A_{S,p}, \Sigma \Rightarrow \Pi$  is derivable using  $\text{G}$  from  $S \ominus p$ .

Now if we have two rules  $R_1 = \mathcal{P}_1 / \Sigma_1 \Rightarrow \Pi_1, \heartsuit p$  and  $R_2 = \mathcal{P}_2 / \heartsuit p, \Sigma_2 \Rightarrow \Pi_2$  an instance of  $\text{cut}(R_1, R_2, \heartsuit p)$  given by a substitution  $\sigma$  and a context  $\Theta \Rightarrow \Xi$ , then the premisses of this instance are of the form  $S \ominus p$  for a suitable  $S \subseteq \mathcal{S}(\mathcal{F})$ . Thus using the fact above from  $S \ominus p$  we can derive the premisses  $S_1$  and  $S_2$  of instances of rules  $R_2$  resp.  $R_1$  given by the substitution  $\sigma'$  with  $\sigma'(p) = A_{S,p}$  and  $\sigma'(x) = x$  for  $x \neq p$  and suitable contexts  $\Omega_1 \Rightarrow \Xi_1$  and  $\Omega_2 \Rightarrow \Xi_2$  with  $\Omega_1, \Omega_2 \Rightarrow \Xi_1, \Xi_2 = \Omega \Rightarrow \Xi$ . But then using the rules  $R_1, R_2$  we can derive the sequents  $\Omega_1, \Sigma_1 \sigma' \Rightarrow \Pi_1 \sigma', \Xi_1, \heartsuit A_{S,p}$  and  $\heartsuit A_{S,p}, \Omega_2, \Sigma_2 \sigma' \Rightarrow \Pi_2 \sigma', \Xi_2$  and a cut on  $\heartsuit A_{S,p}$  yields the conclusion of the instance of  $\text{cut}(R_1, R_2, \heartsuit p)$ .  $\square$



The same technique also shows that in the proof of Theorem 12 we may eliminate the variables  $\mathbf{q}_\ell, \mathbf{q}_r$  occurring in the premisses but not the conclusion by performing  $\mathbf{q}_\ell, \mathbf{q}_r$ -elimination on the set of premisses and end up with a rule equivalent over  $\mathbf{GMonCutCon}$ . Furthermore, the fact that premisses ensuring  $t_P \rightarrow P$  instead of  $t_P \leftrightarrow P$  suffice is an easy Corollary of Lemma 37, as is the last claim of Lemma 21.