























- (1) Initial rules:  $\Gamma, A \vdash \Delta, A$  — where  $A$  is atomic<sup>6</sup> — and  $\Gamma, \perp \vdash \Delta$ .
- (2)  $\rightarrow$ -left rule: 
$$\frac{\Gamma, A \rightarrow \perp \vdash \Delta \quad \Gamma, B \vdash \Delta}{\Gamma, A \rightarrow B \vdash \Delta}$$
- (3)  $\rightarrow$ -right rule: 
$$\frac{\Gamma, A \vdash \Delta, B}{\Gamma \vdash \Delta, A \rightarrow B}$$
- (4)  $\perp$ -right rule: 
$$\frac{\Gamma, A \vdash \Delta}{\Gamma \vdash \Delta, A \rightarrow \perp}$$
- (5)  $\perp$ -left rule: 
$$\frac{\Gamma \vdash \Delta, A}{\Gamma, A \rightarrow \perp \vdash \Delta}$$
- (6)  $\forall$ -right rule: 
$$\frac{\Gamma \vdash \Delta, A[a]}{\Gamma \vdash \Delta, \forall x A[x]}$$
 — as long as  $a$ , the *eigenvariable of the rule*, does not occur in the conclusion (“denominator”) of the rule.<sup>7</sup>
- (7)  $\forall$ -left rule: 
$$\frac{\Gamma, A[a] \vdash \Delta}{\Gamma, \forall x A[x] \vdash \Delta}$$
- (8) The *modified* “GLR”<sup>8</sup> modal rule: 
$$\frac{\forall \Gamma, \Box \Gamma, \Box A \vdash A}{\Phi, \Box \Gamma \vdash \Box A, \Psi} \quad \dashv$$

The  $\Gamma$  and  $\Delta$  in the rules are called the “side formulae” (s.f.); the resulting single formula in the “denominator” in rules (2)–(8) is the “principal formula” (p.f.) of the rule; rule (1) has  $A$  as principal formula. The single formulae displayed in the “numerators” of (2)–(8) are the “minor formulae” (m.f.). A numerator sequent is a *premise* while the denominator sequent is the *conclusion* of the rule. (3) is, intuitively, the “deduction theorem”. The (arbitrary) sets  $\Phi$  and  $\Psi$  in rule (8) are *weakening* and *strengthening parts* respectively (they help us easily achieve 4.4 and 4.5).

For convenience’s sake, let us restate the following definitions, theorems and corollaries from [15]. Whenever a proof is supplied for said results here, it will only address the modal case, since the other cases are identical to the corresponding parts in loc. cit.

<sup>6</sup> The intuition behind the rule is evident. The requirement for atomic  $A$  facilitates the basis step of (meta)proofs by induction on the “order” of (formal) theorems, as it will soon become evident.

<sup>7</sup> “ $x$ ” is thought of as a *metavariable* that denotes *any* bound variable.

<sup>8</sup> Cf. [8, 11, 12] — the modifications are two: a) we require the universal closure of  $\Gamma$  in the numerator, rather than  $\Gamma$  itself, which helps to simulate axiom (7) of  $ML^3$  b) we allow the rule to introduce “weakening” and “strengthening” parts to its conclusion.

DEFINITION 4.1 (Theorems). A *theorem*, or *derived sequent*, is defined inductively to be one of:

- (1) A sequent of the type of rule (1). We say it is derived with *order* 0 or that it is an *axiom*.
- (2) A sequent of the same type as in the denominator of rule (2) provided the two corresponding sequents in the numerator are also theorems. If the latter two are derived with orders  $\leq m$  and  $\leq n$ , then the former is derived with order  $\leq 1 + \max(m, n)$ .
- (3) A sequent of the same type as in the denominator of rules (3–8) provided the corresponding sequent in the numerator is also a theorem. If the latter is derived with order  $\leq m$ , then the former is derived with order  $\leq 1 + m$ .

We say that “ $\Gamma \vdash \Delta$  is (a theorem) provable (derivable) with order  $\leq m$ ”.

REMARK 4.2. The concept of order in the definition above is indebted to [14]. A sequent derivable with order  $\leq m$  is, of course, also derivable with order  $\leq m + 1$ ,  $\leq 2^{m+1}$ , etc.

In what follows we extend the notation  $F[a]$  to  $\Gamma[a]$  and  $(\Gamma \vdash \Delta)[a]$  when attention is to be drawn to the free variable  $a$  (that may or may not actually occur).

THEOREM 4.3. *If  $(\Gamma \vdash \Delta)[a]$  is provable with order  $\leq m$  and  $b$  is some other free variable, then  $(\Gamma \vdash \Delta)[b]$  is also provable with order  $\leq m$ .*

PROOF. Induction on theorems. Let  $(\Gamma \vdash \Delta)[a]$  be the result of rule (8). This sequent has the form  $\Phi[a], \Box\Xi[a] \vdash \Box A[a], \Psi[a]$ , where  $\forall\Xi[a], \Box\Xi[a], \Box A[a] \vdash A[a]$  is derivable with order  $< m$ . By the I.H. so is  $\forall\Xi[b], \Box\Xi[b], \Box A[b] \vdash A[b]$ . Thus  $\Phi[b], \Box\Xi[b] \vdash \Box A[b], \Psi[b]$  is derivable with order  $\leq m$  by an application of rule (8). It is noted that, since boxed formulae have no free variables,  $\Box A[a] = \Box A[b]$  and  $\Box\Xi[a] = \Box\Xi[b]$ ; moreover  $\forall\Xi[a] = \forall\Xi[b]$  since  $\forall\Xi$  has no free variables either.

THEOREM 4.4 (Weakening). *If  $\Gamma \vdash \Delta$  is derived with order  $\leq m$  then so is  $\Theta, \Gamma \vdash \Delta$ .*

PROOF. Induction on the derivation of  $\Gamma \vdash \Delta$ . Our sequent  $\Gamma \vdash \Delta$  was derived with order  $\leq m$ , is the result of rule (8) and has the form  $\Sigma, \Box\Phi \vdash \Box A, \Pi$  ( $\Box A$  p.f.) where  $\forall\Phi, \Box\Phi, \Box A \vdash A$  is derived with order

$< m$ . By an application of rule (8) we can obtain  $\Phi, \Sigma, \Box\Phi \vdash \Box A, \Pi$  with order  $\leq m$  (the I.H. was not needed here).

**COROLLARY 4.5** (Strengthening). *If  $\Gamma \vdash \Delta$  is derived with order  $\leq m$  then so is  $\Gamma \vdash \Delta, \Theta$ .*

**COROLLARY 4.6.**  $\Gamma, A \vdash \Delta, A$  is derivable for any  $A$ .

**Note.** In what follows,  $\implies$  and  $\iff$  are often employed inside a proof as abbreviations of the informal “if ... then” and “iff” respectively.

**PROOF.** By 4.4 and 4.5 it suffices to show the provability of  $A \vdash A$ . We do induction on the complexity of  $A$ .<sup>9</sup> We only show the case where  $A$  is  $\Box C$ . Then

$$\text{I.H. } \implies C \vdash C \xrightarrow{(4.4)} C, \Box C \vdash C \xrightarrow{\text{repeated } \forall\text{-left}} \forall C, \Box C \vdash C \xrightarrow{\text{GLR-rule}} \Box C \vdash \Box C \quad \dashv$$

The following are important *inversion* results. They say that our rules are, in some weak sense, reversible.

- THEOREM 4.7.** (a) *If  $\Gamma, A \rightarrow B \vdash \Delta$  is derivable with order  $\leq m$ , then each of  $\Gamma, A \rightarrow \perp \vdash \Delta$  and  $\Gamma, B \vdash \Delta$  are derivable with order  $\leq m$ .*  
 (b) *If  $\Gamma \vdash \Delta, A \rightarrow B$  is derivable with order  $\leq m$ , then  $\Gamma, A \vdash \Delta, B$  is derivable with order  $\leq m$ .*  
 (c) *If  $\Gamma \vdash \Delta, A \rightarrow \perp$  is derivable with order  $\leq m$ , then  $\Gamma, A \vdash \Delta$  is derivable with order  $\leq m$ .*  
 (d) *If  $\Gamma, A \rightarrow \perp \vdash \Delta$  is derivable with order  $\leq m$ , then  $\Gamma \vdash \Delta, A$  is derivable with order  $\leq m$ .*  
 (e) *If  $\Gamma \vdash \Delta, \forall x A[x]$  is derivable with order  $\leq m$ , then  $\Gamma \vdash \Delta, A[a]$  is derivable with order  $\leq m$  (for any choice of  $a$ ).*

### Gentzen’s Hauptsatz

**DEFINITION 4.8.** A proof gives rise to a *directed graph* in a natural way. Its sequents are the graph nodes. *Edges* are introduced as follows: For every application of inference (2) – a “V-type” inference – we introduce two edges, one connecting the left premise, the other connecting the right premise, to the conclusion. For all other inferences applied in the proof – each an “I-type” inference – we introduce one edge that connects the premise to the conclusion.

<sup>9</sup> Number of  $\rightarrow, \forall, \Box$  in the formula.

- a. The *end-sequent* of a proof is the final (i.e. the bottommost) sequent of the proof.
- b. A sequence of sequents  $(S_i)_{0 \leq i \leq n}$  in a proof  $P$  is called a *path* iff it is so in the graph-theoretic sense: For  $i = 0, 1, 2, \dots, n - 1$ , there is an edge connecting  $S_i$  to  $S_{i+1}$ . If  $i < j$ , we say that  $S_i$  is *above*  $S_j$  in the proof and, correspondingly,  $S_j$  is *below*  $S_i$ . We say that the path  $(S_i)_{0 \leq i \leq n}$  has length  $n$  – that is, the path length equals the number of participating edges. We write  $l(S_0, S_n) = n$ .
- c. A *thread* (of  $P$ ) is a path that connects an axiom to the end-sequent.
- d. An inference  $I$  that is applied in a proof  $P$  *belongs to a path*  $(S_i)_{0 \leq i \leq n}$  iff it contributes an edge to the path.
- e. Let  $I$  be an inference rule<sup>10</sup> and  $S$  be a sequent. We say that  $I$  is *above*  $S$  if  $S$  is either the conclusion of  $I$  or is *below* said conclusion. If  $I$  is a GLR inference and  $S$  is a sequent, then we say that  $I$  is *directly above*  $S$  if  $I$  is above  $S$  and no GLR inferences belong to the path that connects the conclusion of  $I$  to  $S$ .<sup>11</sup>
- f. Let  $I_1$  and  $I_2$  be two *distinct* inferences that belong to the same path  $(S_i)_{0 \leq i \leq n}$  in a proof  $P$ . We say that  $I_1$  is *above*  $I_2$  ( $I_2$  is *below*  $I_1$ ) in  $P$  iff  $I_1$  contributes nodes  $S_i$  and  $S_{i+1}$  and  $I_2$  contributes nodes  $S_j$  and  $S_{j+1}$  and  $i + 1 \leq j$ .  $\dashv$

As is usual, a Gentzen-style proof leads to a tree rather than to an undirected acyclic graph since rather than reusing nodes we use new copies (e.g., several copies of the same axiom).

Given the nomenclature introduced by the above definition we may now freely use colloquialisms such as “ $I$  is the *second* GLR rule above sequent  $S$  on some thread”, meaning that there is a GLR rule  $I'$  on the thread that is directly above  $S$ , while  $I$  is directly above the premise of  $I'$ .

DEFINITION 4.9 (Width [22]). Let  $S = \Gamma \vdash A, \Delta$  be a sequent in a proof  $P$ . We denote by  $\mathcal{W}(S, A)$  the set of all rules  $I$  above  $S$  (along some thread through  $S$ ) that satisfy all the following

- (1)  $I$  is the second GLR rule above  $S$  on the relevant thread
- (2) the GLR rule  $I'$  that is between  $I$  and  $S$  on that thread has  $A$  as its diagonal formula

---

<sup>10</sup> A “rule” means an *application* of a rule in this context.

<sup>11</sup> In particular, if  $S$  is the conclusion of a GLR inference,  $I$ , then  $I$  is directly above  $S$ .

- (3) No rule on the path  $I \rightsquigarrow I' \rightsquigarrow S$  introduces  $A$  by weakening.<sup>12</sup>  
 (4)  $A$  is not the p.f. of  $I$ .

For a given  $S$ , the cardinality of  $\mathcal{W}(S, A)$  is called the *width* of  $A$ —denoted by  $w(S, A)$  with lower case  $w$ . ⊖

Clearly, if  $A$  is not boxed then (2) in 4.9 fails, and thus  $w(S, A) = 0$ . On the other hand, if  $w(S, A) > 0$ , then first off  $A = \Box C$  for some  $C$  and, secondly, each  $I$  in  $\mathcal{W}(S, A)$  has  $\Box C$  occur in the antecedent of its conclusion. Indeed,

- $\Box C$  occurs in the antecedent of the  $I'$  premise by (2).
- By (3) in 4.9,  $\Box C$  cannot be introduced by weakening, thus  $I$  must have the form  $\frac{\forall\Phi, \Box\Phi, \forall C, \Box C, \Box B \vdash B}{\Psi, \Box\Phi, \Box C \vdash \Box B, \Xi}$  where, by (4),  $B$  is not the same as  $C$ .

DEFINITION 4.10. If  $S$  is a sequent in a proof  $P$ , then we denote by  $\text{Sub}_S(P)$  the sub-proof of  $P$  that derives  $S$ . We then define:

- (1) If  $P$  is a proof and  $S$  is a sequent in  $P$ , then a *cover* of  $S$  in  $P$  is a set of sequents in  $\text{Sub}_S(P)$  such that each thread in  $\text{Sub}_S(P)$  contains exactly one member of the set.
- (2) A cover of  $S$  with the property that there are no GLR inferences in the paths between any of its members and  $S$ , is called a *classical cover*.
- (3) If  $\mathcal{U}$  is a cover of  $S$  then we define  $m(\mathcal{U}) = \sum_{S' \in \mathcal{U}} l(S', S)$ .
- (4) A cover  $\mathcal{U}$  of  $S$  is called a *maximal classical cover* if  $\mathcal{U}$  is classical and  $m(\mathcal{U}) = \max\{m(\mathcal{U}') \mid \mathcal{U}' \text{ is a classical cover of } S\}$ .<sup>13</sup>

Note that for any  $S$ ,  $\{S\}$  is a (classical) cover of  $S$ , and therefore a maximal classical cover of  $S$  always exists.

LEMMA 4.11. *Let  $P$  be a proof and  $S = \Gamma, \Box A \vdash \Delta$  a sequent in  $P$ , and let  $\mathcal{U} = \{S_1, \dots, S_n\}$  be a maximal classical cover of  $S$  in  $P$ , then:*

- (a) *Each  $S'$  on the path from  $S_i$  to  $S$  has the form  $\Gamma', \Box A \vdash \Delta'$  for  $i = 1 \dots n$ . (In particular  $S_i = \Gamma_i, \Box A \vdash \Delta_i$  for  $i = 1 \dots n$ ).*
- (b) *If, for each  $i = 1 \dots n$ , there is a proof of  $\bar{S}_i = \Gamma_i, \forall\Sigma, \Box\Sigma \vdash \Delta_i$ , then there is a proof of  $\bar{S} = \Gamma, \forall\Sigma, \Box\Sigma \vdash \Delta$ .*

<sup>12</sup> The only rule that can introduce weakening formulae is the GLR.

<sup>13</sup> Note that if we remove the word *classical* from this definition, then a maximal cover of  $S$  is always the set of initial sequents of  $\text{Sub}_S(P)$ .

PROOF. (a) Fix an  $i$ . We proceed by induction on  $l = l(S_i, S)$ . If  $l = 0$  we are done. If  $l > 0$ , then, since boxed formulae cannot be introduced by a non-GLR inference – and  $S$  is the conclusion of such an inference,  $I - \Box A$  must be present in the antecedent(s) of the premise(s) of  $I$ . Now,  $S_i$  must either be a premise of  $I$  and we are done, or  $S_i$  is above (one of) the premise(s)  $S'$  of  $I$ . Since  $l(S_i, S') < l$  we are done by the I.H.

(b) By induction on  $m = m(\mathcal{U})$ . If  $m = 0$ , then it must be that  $\mathcal{U} = \{S\}$ , and the conclusion is identical to the hypothesis.

Let next  $m > 0$ . We examine the case where  $S$  is the conclusion of rule (2). Thus, let the premises be  $S' = \Gamma', \Box A, C \vdash \Delta$  and  $S'' = \Gamma', \Box A, B \rightarrow \perp \vdash \Delta$  (so  $\Gamma = \Gamma', C \rightarrow B$ ). Let us focus on  $S'$ : We have either  $S' \in \mathcal{U}$ , or there is an  $S_k \in \mathcal{U}$ , which is above  $S'$  along a path  $S_k \rightsquigarrow S' \rightsquigarrow S$ . Let us take the latter case first. Then – since we omit members of  $\mathcal{U}$  on threads that pass through  $S''$  – we have a *proper* subset of  $\mathcal{U}$ , say  $\mathcal{U}'$ , that is a maximal classical cover of  $S'$ . Thus  $m(\mathcal{U}') < m(\mathcal{U})$  and so, by the I.H., we have a proof of  $\Gamma', \forall \Sigma, \Box \Sigma, C \vdash \Delta$  if  $C \neq \Box A$ , or of  $\Gamma', \forall \Sigma, \Box \Sigma \vdash \Delta$  if  $C = \Box A$ ; but we can always use 4.4 to turn the second sequent to the first. If  $S' \in \mathcal{U}$  then we can use the hypothesis (and possibly 4.4) to reach the same conclusion. Similarly, we have a proof of  $\Gamma', \forall \Sigma, \Box \Sigma, B \rightarrow \perp \vdash \Delta$ , and now applying rule (2) we get  $\bar{S}$ .

The other cases are done in a similar manner: if  $\frac{S'}{S}$  was the rule responsible for  $S$ , and it is that  $S' \notin \mathcal{U}$ , then  $\mathcal{U}$  is a maximal classical cover of  $S'$  as well, but its  $m$ -value is smaller and hence the I.H. applies. If  $S' \in \mathcal{U}$ , then the I.H. is not needed. Moreover note that, since  $\forall \Sigma$  and  $\Box \Sigma$  are closed, they cannot interfere with the application of rule (6).  $\dashv$

**THEOREM 4.12** (Gentzen’s *Hauptsatz*). *If  $\Gamma \vdash \Delta, A$  and  $\Theta, A \vdash \Phi$  are derivable, then so is  $\Gamma, \Theta \vdash \Delta, \Phi$ .*

**Note.** The formula  $A$  above is called the *cut formula*.

PROOF. This proof is a slightly more involved version of the proof of the Hauptsatz given in [15], the complication arising from the presence of the diagonal formula in GLR.

Now, the proof is by induction on the ordinal

$$\alpha = \omega^3 \cdot c + \omega^2 \cdot w + \omega \cdot m + n, \quad (\star)$$

where  $c$  is the *modified complexity* of the “cut formula”  $A$ . By “modified complexity” we mean the ordinal  $\omega \cdot k + r$  where  $k$  counts  $\Box$  occurrences



and  $r$  counts the total of all  $\rightarrow, \forall$  occurrences in  $A$ . Thus  $(k, r) < (k+1, r')$  and  $(k, r) < (k, r+1)$  for all  $k, r, r'$ .<sup>14</sup> Thus the induction on  $\alpha$  can be visualised as a quadruple induction: We do a *primary induction* on  $c$  as in the analogous proof in [15]. Induction on  $w = w(\Gamma \vdash \Delta, A)$  is a *secondary induction* – S.I. – on the *width* of the cut formula  $A$  that occurs to the right of  $\vdash$ . A *tertiary* and even a *quaternary induction* – T.I. and Q.I. – on derivation orders for  $\Gamma \vdash A, \Delta$  (order  $\leq m$ ) and  $\Theta, A \vdash \Phi$  (order  $\leq n$ ) will be called upon at appropriate points. Because of the “weights” attached to  $c, w, m$  and  $n$ , a reduction of  $c$  by 1 reduces  $\alpha$  even if the other parameters increase. A reduction of  $w$  also reduces  $\alpha$  as long as  $c$  does not increase.

In all inductions – P.I., S.I., T.I. and Q.I. –  $\Gamma, \Delta, \Theta$  and  $\Phi$  are parameters (i.e., *the P.I.H., S.I.H., T.I.H. and Q.I.H. hold for all  $\Gamma, \Delta, \Theta, \Phi$* ).

First, we deal with the cases where  $A$  is not a boxed formula. The only thing we need to show is that cuts in the proof in [15] that rely on induction on the order of the derivation, i.e. tertiary and quaternary induction, do not involve a cut formula with a higher width, and thus the corresponding argument in [15] is also valid here. This task is relatively easy since a quick examination of the proof in [15] reveals that in all the cuts whose feasibility relied on induction on the order of derivation the cut formula was the original  $A$ ; and since  $A$  is not boxed then the width remained at 0.

Assume now that  $A = \Box C$ . In this case too we start with a tertiary induction, this time on the derivation order  $\leq m$  of  $\Gamma \vdash \Box C, \Delta$ .

1. Say  $\Gamma \vdash \Box C, \Delta$  is an axiom. Since  $\Box C$  is not atomic, so is  $\Gamma, \Theta \vdash \Phi, \Delta$ .

2. We assume first that  $\Box C$  is not the p.f. in the last rule applied to derive  $\mathbf{S} = \Gamma \vdash \Box C, \Delta$ .

Suppose that it was obtained via rule (2) from  $S_3 = \Gamma' \vdash \Box C, \Delta$  and  $S_4 = \Gamma'' \vdash \Box C, \Delta$  (orders  $< m$ ). It is immediate that  $w(S_3, \Box C)$ ,  $w(S_4, \Box C) \leq w(S, \Box C)$  thus, using the T.I.H., both of  $\Theta, \Gamma' \vdash \Delta, \Phi$  and  $\Theta, \Gamma'' \vdash \Delta, \Phi$  are derivable. Applying to these two the rule (2) we derive  $\Theta, \Gamma \vdash \Delta, \Phi$ .

For applicable rules among (3)–(7) the argument is precisely like that for (2), except that it involves one premise.

---

<sup>14</sup> This definition of “complexity” is needed at step IV(d). The whole fuss in the definition is to have “more complexity” in  $\Box A$  than in  $\forall A$ .

3. Case where GLR was applied — still  $\Box C$  is not the p.f. thus it is a strengthening formula. The premise, obtained with order  $< m$ , is of the form  $\forall \Gamma', \Box \Gamma', \Box D \vdash D$ , where  $\Gamma = \Gamma_1, \Box \Gamma_2$  and  $\Gamma' \subseteq \Gamma_2$  while  $\Phi = \Box D, \Phi'$ . Reapplying GLR to the last sequent, with a judicious choice of weakening/strengthening, yields

$$\Theta, \overbrace{\Gamma_1, \Box \Gamma_2}^{\Gamma} - \overbrace{\Box \Gamma', \Box \Gamma'}^{\Phi} \vdash \Box D, \Phi', \Delta$$

4. Let now  $\Box C$  be the p.f. in the last step of the derivation of  $\Gamma \vdash \Box C, \Delta$ : Then  $\Gamma = \Xi, \Box \Psi_1$  while the premise has the following form, with  $\Psi \subseteq \Psi_1$ :

$$\forall \Psi, \Box \Psi, \Box C \vdash C, \text{ derived with order } < m \quad (*)$$

Our key objective is to show that

$$\forall \Psi, \Box \Psi \vdash C \quad (*_1)$$

is derivable. Once we have  $(*_1)$  we can conclude the proof exactly as in [15].

Now, an application of GLR to  $(*)$ , without weakening/strengthening parts, yields

$$\Box \Psi \vdash \Box C, \text{ derived with order } \leq m. \quad (**)$$

This has not increased the induction variable  $(*)$ .

We now turn to the analysis of the derivation of  $\Theta, \Box C \vdash \Phi$  by a quaternary induction (Q.I.) on its derivation order  $\leq n$ .<sup>15</sup> For the basis, if this sequent is an axiom, then so is  $\Gamma, \Theta \vdash \Delta, \Phi$  since  $\Box C$  is not atomic.

For the Q.I. step we only have the case that  $\Box C$  is not the p.f. in the rule that derived  $\Theta, \Box C \vdash \Phi$ .

(i)  $\Theta, \Box C \vdash \Phi$  was obtained via an applicable rule, let us call it “**R**”, among (2)–(7); say, it was (2). The premises have the forms  $S_3 = \Theta', \Box C \vdash \Phi$  and  $S_4 = \Theta'', \Box C \vdash \Phi$ , each derived with order  $< n$ . By the Q.I.H.  $\Theta', \Gamma \vdash \Phi, \Delta$  and  $\Theta'', \Gamma \vdash \Phi, \Delta$  are derivable. We can now apply rule **R** to them to derive  $\Theta, \Gamma \vdash \Phi, \Delta$ .

The cases for the other rules are similar, but with one premise. We note that for rule (6) we can employ 4.3, if necessary, to guarantee that the eigenvariable used does not appear free in  $\Gamma$ .

<sup>15</sup> Note that since the induction is on the width of the formula in the *left* upper sequent, this induction parameter does not come into play here.

(ii) Rule (8) was used to derive  $\Theta, \Box C \vdash \Phi$ . The only possible scenarios are that the premise was either  $\forall\Omega, \Box\Omega, \Box D \vdash D$  (subcase where  $\Box C$  is a weakening formula), or  $\forall\Omega, \Box\Omega, \forall C, \Box C, \Box D \vdash D$ , for some  $D$ ; either sequent being derived with order  $< n$ . Since weakening does not add to derivation order, without loss of generality we accept that *the* premise was

$$\forall\Omega, \Box\Omega, \forall C, \Box C, \Box D \vdash D, \text{ derived with order } < n. \quad (***)$$

Thus  $\Theta = \Theta_1, \Box\Theta_2$  and  $\Omega \subseteq \Theta_2$ , while  $\Phi = \Box D, \Phi'$ .

**We now embark on establishing  $(*_1)$ .** Notice that by remarks following Definition 4.9, the set  $\mathcal{W}(\Box\Psi \vdash \Box C, \Box C)$  – if not empty – will contain rules such as

$$J = \frac{\forall\Lambda, \Box\Lambda, \forall C, \Box C, \Box B \vdash B}{\Pi, \Box\Lambda, \Box C \vdash \Box B, \Xi}$$

We thus have two cases:

**Case**  $w(\Box\Psi \vdash \Box C, \Box C) = 0$ . Let us set  $S = \forall\Psi, \Box\Psi, \Box C \vdash C$ , and let  $\mathcal{U}$  be a maximal classical cover of  $S$  in the proof. Let  $S' \in \mathcal{U}$ . By Lemma 4.11(a),  $S'$  has the form  $\Gamma', \Box C \vdash \Delta'$ . We have three subcases:

1.  $S' = \Gamma', \Box C \vdash \Delta'$  is an axiom. Then so is  $S'' = \forall\Psi, \Box\Psi, \Gamma' \vdash \Delta'$ .
2.  $S'$  is the conclusion of a GLR inference of the form

$$\frac{\forall\Gamma', \Box\Gamma', \Box D \vdash D}{\Box C, \Xi', \Box\Gamma' \vdash \Box D, \Theta'}$$

where  $\Box C$  is a weakening formula; cf. 4.9.

Reusing the GLR differently, we obtain a proof for  $S'' = \forall\Psi, \Box\Psi, \Xi', \Box\Gamma' \vdash \Box D, \Theta'$ .

3.  $S'$  is the conclusion of a GLR inference that has  $\Box C$  as its p.f. (diagonal):

$$\frac{\forall\Lambda, \Box\Lambda, \forall C, \Box C \vdash C}{\Pi, \Box\Lambda, \Box C \vdash \Box C, \Xi}$$

By  $(**)$  via 4.4, we also have a proof of  $S'' = \forall\Psi, \Box\Psi, \Pi, \Box\Lambda \vdash \Box C, \Xi$ .

In all cases we succeeded to go from “the provable  $S'$ ” to “the also provable  $S''$ ”, replacing  $\Box C$  in the antecedent of  $S'$  by “ $\forall\Psi, \Box\Psi$ ”. By Lemma 4.11(b) we have a proof of  $(*_1)$ .

**Case**  $w(\Box\Psi \vdash \Box C, \Box C) > 0$ . We will derive  $(*_1)$  once more. By the condition governing this case, we have at least one GLR inference above

(\*) that has the form  $J$  above. Let us concentrate on one such actual member of  $\mathcal{W}(\Box\Psi \vdash \Box C, \Box C)$  and *eliminate* it while showing that we can still derive (\*). Below we have a partial view of the path where  $J$  and  $\Box\Psi \vdash \Box C$  belong in the proof. By Definition 4.9(4),  $C$  is not the same as  $B$  – the former cannot be the p.f.

$$\begin{array}{c}
 \vdots \pi_0 \\
 \vdots \\
 J \frac{\forall C, \Box C, \forall \Lambda, \Box \Lambda, \Box B \vdash B}{\Pi, \Box C, \Box \Lambda \vdash \Box B, \Xi} \\
 \vdots \pi \\
 \vdots \\
 \frac{\forall \Psi, \Box \Psi, \Box C \vdash C}{\Box \Psi \vdash \Box C}
 \end{array}$$

Examine the following diagram where although we employ cuts we will explain that they can be simulated by the rules (1)–(8) alone. To reduce notational clutter we use the abbreviation  $\hat{\Gamma}$  for “ $\forall \Gamma, \Box \Gamma$ ”. The  $\pi$ -labels are used to keep track of paths that are “copied-and-pasted”.

$$\begin{array}{c}
 \begin{array}{c} \vdots \\ \vdots \text{cf. 4.6} \\ \vdots \end{array} \\
 \text{GLR}_1 \frac{\forall B, \Box B \vdash B}{\forall B, \Box B, \Pi, \underbrace{\Box C}_{\text{weaken.}}, \Box \Lambda \vdash \Box B, \Xi} \\
 \begin{array}{c} \vdots \pi; \hat{B} \text{ is s.f.} \\ \vdots \end{array} \\
 \text{GLR}'' \frac{\hat{B}, \hat{\Psi}, \Box C \vdash C}{\Box \Psi, \Box B \vdash \hat{\mathcal{C}}} \\
 \begin{array}{c} \vdots \pi \\ \hat{\Psi}, \Box C \vdash C \\ \vdots \forall\text{-right} \\ \vdots \pi_0 \end{array} \\
 \text{GLR}''' \frac{\hat{\Psi}, \hat{\Lambda}, \Box B \vdash B}{\underbrace{\Box C, \Pi, \Box \Psi, \Box \Lambda \vdash \Box B, \Xi}_{\text{Weaken.}}} \\
 \begin{array}{c} \vdots \pi; \Box \Psi \text{ is s.f.} \\ \vdots \end{array} \\
 \text{Cut}_3 \frac{\hat{\Psi}, \Box C \vdash C}{\Box \Psi \vdash \hat{\mathcal{C}}} \\
 \begin{array}{c} \vdots \pi \\ \hat{\Psi}, \hat{\mathcal{C}} \vdash C \end{array} \\
 \text{Cut}_2 \frac{\hat{\Psi}, \Box C \vdash \hat{\mathcal{C}} \quad \hat{\mathcal{C}}, \Box C, \hat{\Lambda}, \Box B \vdash B}{\hat{\mathcal{C}}, \hat{\Psi}, \hat{\Lambda}, \Box B \vdash B} \text{Cut}_1
 \end{array}$$

- The conclusion of  $\text{Cut}_1$  is derivable by the P.I.H.
- The conclusion of  $\text{Cut}_2$  is derivable because  $\text{GLR}_1$  does not contribute to the width (cf. 4.9(3)) since  $\Box C$  is weakening, and thus  $w(\Box \Psi, \Box B \vdash \Box C, \Box C) < \text{original } w(\Box \Psi \vdash \Box C, \Box C)$ , allowing the use of the S.I.H.

- The conclusion of  $\text{Cut}_3$  is derivable, again utilising the S.I.H. Notice that we removed the GLR  $J$ —depicted above at the end of the  $\pi_0$ -path—from  $\mathcal{W}(\Box\Psi \vdash \Box C, \Box C)$  and  $\text{GLR}_1$  was *not* added. Thus we have reduced the *original*  $w(\Box\Psi \vdash \Box C, \Box C)$  by 1.

By reference to the proof in [15], this concludes the present proof.  $\dashv$

## 5. Comparing $\text{ML}^3$ and GLTS

In this section and later we use the notational convention

- (i)  $\Gamma \vdash_{\text{GLTS}} \Delta$  is short for “the sequent  $\Gamma \vdash \Delta$  is derivable in GLTS”.
- (ii)  $\Gamma \vdash_{\text{GTKS}} \Delta$  is short for “the sequent  $\Gamma \vdash \Delta$  is derivable in GTKS of [15]”.

The main theorem of this section is:

**THEOREM 5.1.** *If  $\Gamma \vdash_{\text{GLTS}} \Delta$  then  $\Gamma \vdash_{\text{ML}^3} \bigvee \Delta$ , and if  $\Gamma \vdash_{\text{ML}^3} A$ , then  $\bigvee \Gamma \vdash_{\text{GLTS}} A$ .*

As in [15],  $\bigvee \Delta$  means  $\bigvee_{A \in \Delta} A$ . Furthermore,  $\Gamma \vdash_X \Delta$  means that the sequent  $\Gamma \vdash \Delta$  is derivable in  $X$ , if  $X$  is a Gentzen-style system. If  $Y$  is a Hilbert-style system, then  $\Gamma \vdash_Y A$  means that  $A$  is derivable in  $Y$  from hypotheses  $\Gamma$ .

**LEMMA 5.2.** 1.  $\Gamma \vdash_{\text{GTKS}} \Delta \implies \Gamma \vdash_{\text{GLTS}} \Delta$ .  
 2.  $\Gamma \vdash_{\text{GLTS}} \Delta \iff \exists A_1 \dots \exists A_n$  such that  $L_1(A_1, \dots, A_n), \Gamma \vdash_{\text{GTKS}} \Delta$ .

**PROOF.** 1. Since both share the same non-modal inference rules, we only need to show that GLTS can simulate the rule TR of GTKS.

Indeed, if  $\forall \Gamma, \Box \Gamma \vdash_{\text{GLTS}} A$  then by 4.4,  $\forall \Gamma, \Box \Gamma, \Box A \vdash_{\text{GLTS}} A$  and thus (8) derives  $\Sigma, \Box \Gamma \vdash_{\text{GLTS}} \Box A, \Xi$ .

2. First assume the left hand side. The proof is by induction on the order of derivation.

- a. If  $\Gamma \vdash_{\text{GLTS}} \Delta$  is an axiom then so is  $L_1(A), \Gamma \vdash_{\text{GTKS}} \Delta$  for any  $A$ .
- b. If  $\Gamma \vdash_{\text{GLTS}} \Delta', \forall x A(x)$  is the conclusion of rule (6) applied to the premise  $\Gamma \vdash_{\text{GLTS}} \Delta', A(a)$ , then by the I.H. there are  $A_1, \dots, A_n$  such that  $L_1(A_1, \dots, A_n), \Gamma \vdash_{\text{GTKS}} \Delta', A(a)$ . Now apply rule (6),<sup>16</sup> bearing in mind that  $L_1(A_1, \dots, A_n)$  is a set of closed formulae, to obtain  $L_1(A_1, \dots, A_n), \Gamma \vdash_{\text{GTKS}} \Delta', \forall x A(x)$ .

<sup>16</sup> Which both systems share.

(The other non-modal rules are done similarly.)

c. If  $\Sigma, \Box\Gamma \vdash_{\text{GLTS}} \Box A, \Xi$  is the conclusion of rule (8) applied to the premise  $\forall\Gamma, \Box\Gamma, \Box A \vdash_{\text{GLTS}} A$ , then, by the I.H.,  $L_1(A_1, \dots, A_n), \forall\Gamma, \Box\Gamma, \Box A \vdash_{\text{GTKS}} A$  for some  $A_1, \dots, A_n$ .

Using rule (3) we get  $L_1(A_1, \dots, A_n), \forall\Gamma, \Box\Gamma \vdash_{\text{GTKS}} \Box A \rightarrow A$  and, via 4.4,  $\forall\Psi, \Box\Gamma \vdash_{\text{GTKS}} \Box A \rightarrow A$  with the appropriate  $\Psi$  to make TR applicable. The latter then yields  $L_1(A_1, \dots, A_n), \Sigma, \Box\Gamma \vdash_{\text{GTKS}} \Box(\Box A \rightarrow A), \Xi$ . Also, it is readily seen that  $\Box L(A), \Box(\Box A \rightarrow A) \vdash_{\text{GTKS}} \Box A$ . Now, by cutting the formula  $\Box(\Box A \rightarrow A)$  we derive  $L_1(A, A_1, \dots, A_n), \Sigma, \Box\Gamma \vdash_{\text{GTKS}} \Box A, \Xi$ .

Next, assume the right hand side. Thus, we have  $A_1, \dots, A_n$  such that  $L_1(A_1, \dots, A_n), \Gamma \vdash_{\text{GTKS}} \Delta$ . By part 1,  $L_1(A_1, \dots, A_n), \Gamma \vdash_{\text{GLTS}} \Delta$ .

Now, 3.8 and the results in [15] yield  $\Box L(A_i) \vdash_{\text{GTKS}} L(A_i)$  and, by part 1,  $\Box L(A_i) \vdash_{\text{GLTS}} L(A_i)$ . Thus, applying GLR, we obtain  $\vdash_{\text{GLTS}} \Box L(A_i)$ , for all  $i = 1 \dots n$ . It follows that we can remove  $L_1(A_1, \dots, A_n)$  from  $L_1(A_1, \dots, A_n), \Gamma \vdash_{\text{GLTS}} \Delta$  using repeated cuts.  $\dashv$

*Proof of 5.1.*  $\Gamma \vdash_{\text{GLTS}} \Delta \xleftrightarrow{5.2} \exists A_1 \dots \exists A_n : L_1(A_1, \dots, A_n), \Gamma \vdash_{\text{GTKS}} \Delta \xrightarrow{[15]} \exists A_1 \dots \exists A_n : L_1(A_1, \dots, A_n), \Gamma \vdash_{\text{M}^3} \bigvee \Delta \xleftrightarrow{3.9} \Gamma \vdash_{\text{ML}^3} \bigvee \Delta$ .

Conversely,

$\Gamma \vdash_{\text{ML}^3} A \xleftrightarrow{3.9} \exists A_1 \dots \exists A_n : L_1(A_1, \dots, A_n), \Gamma \vdash_{\text{M}^3} A \xrightarrow{[15]} \exists A_1 \dots \exists A_n : L_1(A_1, \dots, A_n), \forall\Gamma \vdash_{\text{GTKS}} A \xleftrightarrow{5.2} \forall\Gamma \vdash_{\text{GLTS}} A$ .  $\dashv$

## 6. Weak reflection

Toward a proof of 6.4 below, we state without proof the following two results (proved in [15]):

**PROPOSITION 6.1** (GTKS vs  $\text{M}^3$ ). *If  $\Gamma \vdash_{\text{GTKS}} \Delta$ , then  $\Gamma \vdash_{\text{M}^3} \bigvee \Delta$ . In the other direction, if  $\Gamma \vdash_{\text{M}^3} A$ , then  $\forall\Gamma \vdash_{\text{GTKS}} A$ .*

**PROPOSITION 6.2.** *If  $\Sigma, \Delta$  are classical, the GTKS-derivability of  $\Sigma, \Box\Gamma \vdash \Delta$  implies that of  $\Sigma \vdash \Delta$  with a proof that does not use the TR rule and employs only classical formulae.*

**Note.** In loc. cit. the additional assumption that  $\Gamma$  is classical was never used in the proof of 6.2, thus it is not made here.

**PROPOSITION 6.3.** *If  $\Gamma, \Delta$  are classical, then the derivability of  $\Gamma, \Box\Theta \vdash \Box A, \Delta$  in GTKS implies that of  $\Gamma, \forall\Theta, \Box\Theta \vdash A, \Delta$ .*

**Note.** Again, there was a restriction in the statement of 6.3 as given in loc. cit. that  $\Theta$  is classical. Given that the said restriction was not used anywhere in the proof, we have also lifted it in the proposition statement here. We emphasise, as it was also emphasised in loc. cit., that we do not require  $A$  to be classical either.

**PROPOSITION 6.4** (Weak Reflection for  $\text{ML}^3$ ). *If  $\Gamma, \Box\Gamma \vdash_{\text{ML}^3} \Box A$ , where  $\Gamma$  is classical, then  $\Gamma, \Box\Gamma \vdash_{\text{ML}^3} A$  as well.*

**Note.**  $A$  is not assumed to be classical.

**PROOF.**  $\Gamma, \Box\Gamma \vdash_{\text{ML}^3} \Box A \xrightarrow{3.9} \exists A_1, \dots, A_n : L_1(A_1, \dots, A_n), \Gamma, \Box\Gamma \vdash_{\text{M}^3} \Box A \xrightarrow{6.1} \exists A_1, \dots, A_n : L_1(A_1, \dots, A_n), \forall\Gamma, \Box\Gamma \vdash_{\text{GTKS}} \Box A \xrightarrow{6.3} \exists A_1, \dots, A_n : \forall L_0(A_1, \dots, A_n), L_1(A_1, \dots, A_n), \forall\Gamma, \Box\Gamma \vdash_{\text{GTKS}} A \xrightarrow{3.8 \& 4.12} \exists A_1, \dots, A_n : L_1(A_1, \dots, A_n), \forall\Gamma, \Box\Gamma \vdash_{\text{GTKS}} A \xrightarrow{5.2} \forall\Gamma, \Box\Gamma \vdash_{\text{GLTS}} A \xrightarrow{5.1} \Gamma, \Box\Gamma \vdash_{\text{ML}^3} A$ . A short explanation for step  $\xrightarrow{6.3}$ : The premise  $L_1(A_j)$  – i.e.,  $\Box L_0(A_j)$  – is a “ $\Box\Theta$ ”, thus, by 6.3, contributes a  $\forall L_0(A_j)$  to the premise. But  $L_0(A_j)$  is closed.  $\dashv$

**COROLLARY 6.5** (Conservation Theorem for  $\text{ML}^3$ ). *If  $\Gamma$  and  $A$  are classical, then  $\Gamma, \Box\Gamma \vdash_{\text{ML}^3} \Box A$  implies that we have a classical proof of  $A$  from  $\Gamma$ .*

**PROOF.** As in the proof of the previous proposition, we can deduce that  $\exists A_1, \dots, A_n : L_1(A_1, \dots, A_n), \forall\Gamma, \Box\Gamma \vdash_{\text{GTKS}} A$ . By 6.2 we obtain a classical proof of  $\forall\Gamma \vdash_{\text{GTKS}} A$  and hence (using 6.1) of  $\Gamma \vdash_{\text{M}^3} A$  and, by 3.9, of  $\Gamma \vdash_{\text{ML}^3} A$ .<sup>17</sup>  $\dashv$

**Note.** The converse is trivially true, as the assumption leads to  $\Gamma, \Box\Gamma \vdash_{\text{ML}^3} A$  and then we apply weak necessitation.

### What about strong reflection and strong necessitation?

Neither  $\Box p \rightarrow p$  (strong reflection) nor  $p \rightarrow \Box p$  (strong necessitation) are provable within  $\text{ML}^3$ , where  $p$  is a propositional variable: Note that if  $\vdash_{\text{ML}^3} \Box p \rightarrow p$  then  $\vdash_{\text{GLTS}} \Box p \rightarrow p$ , hence  $\Box p \vdash_{\text{GLTS}} p$ . By rule (8) we have  $\vdash_{\text{GLTS}} \Box p$  and cutting it with the first sequent we obtain  $\vdash_{\text{GLTS}} p$  which is untenable:  $\vdash p$  is neither a GLTS axiom, nor the consequence (denominator) of any rule.

<sup>17</sup> It is quite simple to establish that any classical proof in  $\text{M}^3$  is a classical proof in  $\text{ML}^3$ .

Similarly, if  $\vdash_{\text{ML}^3} p \rightarrow \Box p$ , then  $\vdash_{\text{GLTS}} p \rightarrow \Box p$  and hence  $p \vdash_{\text{GLTS}} \Box p$ . Now,  $p \vdash \Box p$  is not a (GLTS) axiom and  $p$  cannot be the p.f. of any rule. Can  $\Box p$  be such a p.f.? Well, if so, the premise must be of the form  $\forall \Gamma, \Box \Gamma, \Box p \vdash p$ . Now  $\Gamma \neq \emptyset$  since  $\Box p \vdash p$  is not derivable as we have just seen. But then the result  $\neg p \vdash \Box p$ , must contain at least one boxed formula to the left of  $\vdash$ , which is not the case.

## 7. Semantic completeness of $\text{ML}^3_+$ and $\text{ML}^3$

### 7.1. Preliminaries

The standard semantic approach to modal logic is due to Kripke, who in 1959 introduced his “possible worlds” semantics [7]. This approach involves a *frame*, which is a pair,  $\langle W, R \rangle$ , consisting of a non-empty set of *possible worlds*,  $W$ , and a binary *accessibility* relation,  $R$ , on  $W$ . Thus, for example, if  $w_1$  and  $w_2$  are possible worlds in  $W$  then we say that  $w_2$  is *accessible from*  $w_1$  iff  $w_1 R w_2$ .

We will use here a particular type of a Kripke frame, called a *pointed Kripke frame* (cf. [17]), which is a triplet  $\langle W, R, w_0 \rangle$  such that  $\langle W, R \rangle$  is a frame and  $w_0 \in W$  is an  $R$ -minimum, that is, for all  $w \in W$ ,  $w_0 = w$  or  $w_0 R w$ .

A *Kripke structure*,  $\mathcal{M}$ , is a pair  $(\langle W, R, w_0 \rangle, \{(M_w, I_w) : w \in W\})$  where  $\langle W, R, w_0 \rangle$  is a pointed frame, and  $I_w$  is an *interpretation* (on each  $w \in W$  – cf. [21]) where  $M_w$  is the *domain* of individuals associated with  $w$ . Basically, an interpretation (or sometimes, a *valuation*),  $I_w$ , effects a truth-assignment, for each world  $w \in W$ , from closed formulae to the set  $\{f, t\}$ , so that for any closed formula  $A$  in  $w$ ,  $I_w(A) = t$  indicates that  $A$  is true in  $w$ . For all modal formulae  $B$ , not necessarily closed,  $B$  is true in  $w$  iff  $I_w(\forall B) = t$ .<sup>18</sup> We write,  $w \models B$  ([17] writes  $w \Vdash B$  and says “ $w$  forces  $B$ ”, terminology that we often use here).

For the record (cf. [21]), each interpretation  $I_w$  acts on a countable (finite or enumerable) domain  $M_w$  associated with  $w$ , and is defined by induction on the complexity of *closed formulae*. The interesting points to note are two: one, for each  $w$ ,  $I_w(\forall x B[x])$  is defined “locally” – that

---

<sup>18</sup> The  $W_w$  are just a sets. The functions  $I_w$  effect a first-order modal interpretation of formulae. Apart from the slight departure mentioned in the last paragraph of the Introduction, the definition of the various  $I_w$  is unremarkable, and details are given in [20, 21].



is, it is true iff all of  $B[x := i]$  ( $i \in M_w$ ) are forced by  $w$ ; two,  $I_w(\Box B)$  is defined “globally” — that is, it is true iff all  $w'$  satisfying  $wRw'$  force  $\forall B$ .

We say that  $B$  is *true in a structure*  $\mathcal{M}$  iff  $w_0 \models B$  — i.e., is forced by the “start-world” of the frame. We also say that the structure is a *model* of  $B$  and write  $\models_{\mathcal{M}} B$ . If, for some  $\mathcal{S}$ , the foregoing holds for all  $B \in \mathcal{S}$ , we call  $\mathcal{M}$  a model of  $\mathcal{S}$  and write  $\models_{\mathcal{M}} \mathcal{S}$ . We write  $\models \mathcal{S}$  — no subscript — iff every structure is a model of  $\mathcal{S}$ . Semantic implication,  $\mathcal{S} \models B$ , means that every model of  $\mathcal{S}$  is a model of  $B$  as well.

The parallel notion (to truth) of *validity* of  $C$  in  $\mathcal{M}$  — i.e., *truth in all worlds* of  $\mathcal{M}$  — will not concern us since it is tantamount to the truth of  $\Box C \wedge C$  at  $w_0$ .

In what follows, and in analogy with the term *well-founded*, we will call a (binary) relation  $R$  *converse well-founded* iff its converse,  $R^{-1}$ , is well-founded. The next result is easily verifiable and well-known:

**PROPOSITION 7.1.** *Let  $\langle W, R \rangle$  be a frame such that  $W$  is finite and  $R$  is transitive, then  $R$  is converse well-founded if and only if  $R$  is irreflexive.*

Now, returning to our development, it is known that GL is sound with respect to Kripke models equipped with converse well-founded accessibility relations.

It is also known that the converse to this property is true for K4 and GL: that is, if  $A$  is true in all (pointed) Kripke models whose accessibility relation is transitive (resp. transitive and converse well-founded), then  $A$  is a theorem of K4 (resp. of GL).

Reference [21] introduced the system (which we here call)  $M_+^3$  whose language and axioms are those of  $M^3$  but is also equipped with the equality predicate, constant and function symbols, and also the axioms:

1.  $a = a$
2.  $\Box(a = a)$
3.  $s = t \rightarrow A[a := s] \leftrightarrow A[a := t]$ , for all terms  $s, t$ .
4.  $\Box(s = t \rightarrow A[a := s] \leftrightarrow A[a := t])$ , for all terms  $s, t$ .

The cited paper showed the completeness of  $M_+^3$  with respect to all pointed Kripke frames whose accessibility relation is transitive.

In the next subsection, we build on techniques from loc. cit. and show that the system  $ML_+^3$  (which is obtained from  $ML^3$  in the identical manner that  $M_+^3$  was from  $M^3$ ) is complete with respect to the class of pointed Kripke frames whose accessibility relation is transitive and converse well-founded. In what follows we view  $ML_+^3$  as  $M_+^3$  with the

Löb axiom added, thus the logical axioms are the schemata of Definition 3.1 plus  $\Box A \rightarrow \Box\Box A$  (explicitly listed), along with their (singly, *not* singly *and* doubly) boxed versions: in what follows, we denote all these axioms by  $\tilde{\Lambda} \cup \Box\Lambda$ . The reason we opt to add  $\Box A \rightarrow \Box\Box A$  explicitly is convenience: Makes the inclusion of the  $\Box\Box$  versions of the schemata in 3.1 redundant, as the only technical purpose for including  $\Box\Box\Lambda$  was toward deriving the schema  $\Box A \rightarrow \Box\Box A$  from the remaining schemata.

Since all the preliminary definitions and major parts of the proof in [21] apply without modification here, we will only elaborate on the parts where the two proofs diverge. Therefore, it is recommended that the readers familiarise themselves with the said paper.

### 7.2. Semantic completeness of $ML_+^3$

The following is borrowed from [21].

LEMMA 7.2 (Main Semantic Lemma for  $M_+^3$ ). *Let  $\mathcal{T}$  be a consistent set of wfmf over the language  $L$  of  $M_+^3$ , and let  $N$  be an arbitrary enumerable set. Then there is a finite or enumerable subset  $M$  of  $N$  and a consistent Henkin completion  $\Gamma$  of  $\mathcal{T}$  over  $L(M)$  — the language  $L$  extended by adding all the members of  $M$  as new constants.*

*That is,  $\Gamma$  is a set of wfmf over  $L(M)$  such that*

- i.  $\Gamma$  is consistent and the set-difference  $\Gamma - \mathcal{T}$  is a set of sentences.
- ii. (Maximality) For any sentence  $A$  over  $L(M)$ , either  $A \in \Gamma$  or  $\neg A \in \Gamma$ , where we henceforth write “ $\neg A$ ” for “ $A \rightarrow \perp$ ”.
- iii. (Henkin property) If  $\Gamma$  proves  $\exists xA$  over  $L(M)$  then it also proves  $A[c]$  for some  $c \in M$ ;  $c$  is called a Henkin (witnessing) constant.
- iv. (Distinguishing Constants) If  $a \neq b$  in  $M$  (metamathematically), then  $\Gamma \vdash \neg a = b$ .

Note that the maximality of  $\Gamma$  implies that  $\Gamma$  is deductively closed: for all  $A$ , if  $\Gamma \vdash A$ , then  $A \in \Gamma$ . The converse is, of course, trivial.

As in [21], *strong completeness* — “for  $M_+^3$ : if  $\mathcal{T} \models A$ , then  $\mathcal{T} \vdash_{M_+^3} A$ ” — is proved by arguing the contrapositive: “If, for a closed  $A$ ,  $\mathcal{T} \not\models A$ , then there is a model  $\mathcal{M}$  of  $\mathcal{T}$  that falsifies  $A$ ;  $I_{w_0}(A) = f$ ”.

This is taken as is from [21] for  $M_+^3$  and using the “standard trick” (cf. [17]) of cutting down a K4 model into a *finite* GL model — adapted here to the first-order case — we will cut down a  $M_+^3$  model into a finite  $ML_+^3$  model.

First let us define the *reduced set of subformulae* of a given closed formula  $A$ .

DEFINITION 7.3. For any closed formula  $A$  over the language  $L$ , the *reduced set of closed subformulae* of  $A$ ,  $S_r(A)$ , is defined as follows:

$$S_r(A) = \{A\} \cup \begin{cases} \emptyset & A \text{ is atomic} \\ S_r(B) \cup S_r(C) & A \text{ is } B \rightarrow C \\ S_r(B) & A \text{ is } \neg B \\ S_r(\forall B) & A \text{ is } \Box B \\ \{B[x := t] : t \text{ is a closed term over } L\} & A \text{ is } \forall x B[x] \end{cases}$$

**Note.** The aim in the last clause is not to miss any subformulae of  $A$  of the form  $\Box E$  by virtue of them “hiding” in the scope of a  $\forall$ . It should be obvious that as all such subformulae are closed, we will get multiple identical copies of boxed subformulae, one for each  $t$ . Thus, any one, “canonically chosen”,  $t$  will do as well, rather than taking all the (infinitely many)  $t$ .

It is immediate from 7.3 that:

PROPOSITION 7.4. For any closed formula  $A$ , the set  $S_{\Box, \forall}(A)$ , defined as  $\{\Box B, \forall B \mid \Box B \in S_r(A)\}$ , is finite.

Let now  $A$  be closed such that  $\mathcal{S} \not\vdash_{\text{ML}^3_+} A$ . Then it is also the case that  $\mathcal{S} \not\vdash_{\text{M}^3_+} A$ . Thus, by [21], let  $\mathcal{M} = (\langle W, R, w_0 \rangle, \{(M_w, I_w) : w \in W\})$  be a model of  $\mathcal{S}$  such that  $I_{w_0}(A) = f$ . Here each  $w$  is a consistent Henkin completion of  $\tilde{\Lambda} \cup \Box \tilde{\Lambda}$  – taking  $\mathcal{T} = \emptyset$  in 7.2 – while  $w_0$  is a consistent Henkin completion of  $\mathcal{S} \cup \{\neg A\}$ . The  $I_w$  are defined as in loc. cit. starting with

$$\text{for each Boolean variable } q, \mathcal{I}_w(q) = \mathbf{t} \text{ iff } q \in w \quad (\diamond)$$

and for each  $n$ -ary predicate  $\phi$ , and  $i_1, \dots, i_n$  in  $M_w$ ,

$$\mathcal{I}_w(\phi(\vec{i}_n)) = \mathbf{t} \text{ iff } \phi(\vec{i}_n) \in w, \quad (\heartsuit)$$

where  $M_w$  is the “domain  $M$ ” associated with the Henkin completion  $w$  as in 7.2.

We now *identify* any two worlds of  $\mathcal{M}$  that contain precisely the *same subset* of  $S_{\Box, \forall}(A)$  – that is, if  $w', w'' \in W$  then  $w' \sim w''$  iff  $w' \cap S_{\Box, \forall}(A) = w'' \cap S_{\Box, \forall}(A)$ . Next, we first pick  $w_0$  as the *representative* of the (equivalence) class of which it is a member, and then we arbitrarily

pick a *representative* from each one of the remaining classes. We will call these world-representatives  $\alpha, \beta, \gamma$ , etc. rather than  $w, w', w_2$ , etc. In particular, we will let  $\alpha_0 = w_0$ . Thus,  $A \notin \alpha_0$ . By 7.4, we obtain a finite set of worlds  $W_r = \{\alpha_0, \alpha_1, \dots, \alpha_m\}$ . The corresponding domains are naturally called  $M_{\alpha_j}$ .

Now that we have cut down the set of worlds into a finite subset, our task is to redefine the accessibility relation  $R$ , to get an irreflexive and transitive relation,  $\widehat{R}$  on the set  $W_r$ .

**DEFINITION 7.5** (The Accessibility Relation modulo a fixed  $A$ ). Let  $\beta, \gamma \in W_r$ . Then  $\beta \widehat{R} \gamma$  iff

- i. For every  $\Box B \in \beta$  that is a reduced subformula of  $A$  we have that  $\forall B$  and  $\Box B$  are in  $\gamma$ .
- ii. There is some reduced subformula  $\Box E$  of  $A$  such that  $\Box E \in \gamma$  but  $\Box E \notin \beta$ .

**Note.** With 7.5 recorded, we now further reduce  $W_r$ , if applicable, by dropping all the  $\alpha_j$  that are not  $\widehat{R}$ -accessible from  $\alpha_0$ . We continue calling it “ $W_r$ ”.

**PROPOSITION 7.6.**  $\widehat{R}$  is irreflexive and transitive.

**PROOF.** If  $\beta \widehat{R} \beta$  for some  $\beta \in W_r$  then, by 7.5, there must be some  $\Box E \in \beta$  such that  $\Box E \notin \beta$  which is impossible, therefore  $\widehat{R}$  is irreflexive.

Now, assume that  $\beta \widehat{R} \gamma \widehat{R} \delta$  for  $\beta, \gamma, \delta$  in  $W_r$ . Let  $\Box B$  be a subformula of  $A$  such that  $\Box B \in \beta$ . Then, by 7.5,  $\Box B \in \gamma$  and thus  $\{\forall B, \Box B\} \subseteq \delta$ . Now, since  $\beta \widehat{R} \gamma$  there is some  $\Box E \in \gamma$  such that  $\Box E \notin \beta$ , but, because  $\gamma \widehat{R} \delta$ , we know that  $\Box E \in \delta$ . Thus, by definition,  $\beta \widehat{R} \delta$ .  $\dashv$

We now state:

**THEOREM 7.7.** For the aforementioned pair  $\mathcal{S}$  and  $A$  such that  $\mathcal{S} \not\vdash_{\text{ML}_+^3} A$ , there is a Kripke model  $\widehat{\mathcal{M}} = (\langle W_r, \widehat{R}, \alpha_0 \rangle, \{(M_\beta, I_\beta) : \beta \in W_r\})$  of  $\mathcal{S}$  such that  $W_r$  is finite,  $\widehat{R}$  is transitive and irreflexive (hence also converse well-founded) and  $I_{\alpha_0}(A) = f$ .

The proof hinges on the following lemma that is the counterpart of Lemma 6.3 in [21]:

**LEMMA 7.8.** We have fixed a sentence  $A$  as in the foregoing, as well as an  $\widehat{\mathcal{M}}$  obtained by cutting down a model  $\mathcal{M}$  of  $\mathcal{S}$  for  $\text{M}_+^3$  such that  $w_0$  does not force  $A$ .

For each  $\alpha \in W_r$  and any given reduced subformula  $B$  of  $A$ , we have  $\mathcal{I}_\alpha(B) = t$  iff  $B \in \alpha$ .

PROOF. The proof will be done using induction on the *modified complexity* of a formula (cf. the proof of 4.12), a natural choice since we took  $\forall B$  to be a “subformula” of  $\Box B$ .

The basis is the definition of  $I_\alpha$  for  $\alpha \in W_r$  as in ( $\diamond$ ) and ( $\heartsuit$ ) on p. 355. For the induction step we will only develop the two interesting cases, where  $B = \forall xC$  or  $B = \Box C$ . The other cases are handled similarly to those in Lemma 6.3 of [21].

1.  $B = \forall xC$ . This is argued precisely as in [21]:

Let first  $I_\alpha(\forall xC) = t$ . Then (Kripke semantics) it must be that  $I_\alpha(C[x := i]) = t$  for all  $i \in W_\alpha$ , and therefore, by the I.H.,  $C[x := i] \in \alpha$  for all such  $i$ . Now, if  $\forall xC \notin \alpha$ , then, by maximality,  $\neg \forall xC$  is in  $\alpha$ ; that is,  $\exists x\neg C$  is. It follows that for some Henkin constant  $c \in M_\alpha$ ,  $\neg C[x := c]$  is in  $\alpha$  contradicting the latter’s consistency.

Conversely, assume that  $\forall xC \in \alpha$ . By Axiom (2) in 3.1, modus ponens, and deductive closure,  $C[x := i]$  is in  $\alpha$ , for all  $i \in M_\alpha$ . By the I.H.,  $I_\alpha(C[x := i]) = t$  for all such  $i$ , and by the definition of  $I_\alpha$  we get that  $I_\alpha(\forall xC) = t$ .

2.  $B = \Box C$ . Let first  $\Box C \in \alpha$ . To conclude that  $B$  is forced by  $\alpha$ , examine the arbitrary  $\beta$  such that  $\alpha \hat{R} \beta$  – of course, if none such exists, then there is nothing to prove. As  $\Box C$  is a subformula of our “reference”,  $A$ , we have (cf. 7.5) that  $\Box C$  and  $\forall C$  are in  $\beta$ . The I.H. applies to  $\forall C$  (lower modified complexity than  $B$ ), and thus it is forced by (is true in)  $\beta$ . By Kripke semantics,  $I_\alpha(\Box C) = t$ .

Conversely, let (arguing contrapositively; cf. [21])  $\Box C \notin \alpha$ . Consider the set  $X = \{\Box C, \neg \forall C\} \cup \{\forall D, \Box D \mid \Box D \text{ is a reduced subformula of } A \text{ and } \Box D \in \alpha\}$ . We write  $X$  as

$$\{\forall D_1, \Box D_1, \dots, \forall D_r, \Box D_r, \Box C, \neg \forall C\}$$

and claim that it is consistent. If not, using propositional logic, we get:

$$\vdash \forall D_1 \wedge \Box D_1 \wedge \dots \wedge \forall D_r \wedge \Box D_r \rightarrow (\Box C \rightarrow \forall C)$$

By  $\vdash \forall C \rightarrow C$  (Axiom 2), the above yields:

$$\vdash \forall D_1 \wedge \Box D_1 \wedge \dots \wedge \forall D_r \wedge \Box D_r \rightarrow (\Box C \rightarrow C)$$

Using weak necessitation (3.7), Axiom (5) and 3.5 we get:

$$\vdash \Box \forall D_1 \wedge \Box \Box D_1 \wedge \dots \wedge \Box \forall D_r \wedge \Box \Box D_r \rightarrow \Box(\Box C \rightarrow C)$$

Noting that  $\vdash \Box D_i \rightarrow \Box \forall D_i \wedge \Box \Box D_i$ <sup>19</sup> and that  $\vdash \Box(\Box C \rightarrow C) \rightarrow \Box C$ , we obtain:

$$\vdash (\Box D_1 \wedge \dots \wedge \Box D_r) \rightarrow \Box C$$

However, by construction, the (deductively closed) world  $\alpha$  contains all of  $\Box D_1, \dots, \Box D_r$  but does not contain  $\Box C$  – a contradiction. Therefore, the set  $X$  is *consistent*.

This fact and Lemma 7.2 imply that there must be a world, say  $w$ , in  $W$  that contains  $X$ . Now, since every world in  $W$  that is equivalent to  $w$  must contain the same subset of  $S_{\Box, \forall}(A)$  as  $w$  does (and, by maximality, they also must contain the same subset of negations of members of  $S_{\Box, \forall}(A)$ ), we immediately conclude that all the members of the equivalence class of  $w$  contain  $X$ . Let us call  $\beta$  the representative picked from this equivalence class for membership in  $W_r$ . In particular, notice that  $\forall C \notin \beta$ .

By construction of  $X$  and  $\beta$ , if a subformula  $\Box D$  of  $A$  is in  $\alpha$ , then it is in  $\beta$ . Moreover, all  $\forall D$  are in  $\beta$ . Finally,  $\Box C \in \beta$  but  $\Box C \notin \alpha$ , the latter by this case’s main assumption. Thus,  $\alpha \widehat{R} \beta$ . But,  $\forall C$  is not in  $\beta$  because  $\neg \forall C$  is. Thus  $\beta$  forces the falsehood of  $\forall C$ , and we conclude that  $I_\alpha(\Box C) = f$ . ⊥

PROOF OF 7.7. Since  $A \notin \alpha_0 = w_0$ , then by the previous lemma (recall that we assume that  $A$  is closed),  $\mathcal{M}$  is a Kripke model of  $\mathcal{S}$  for  $M_+^3$  such that  $W_r$  is finite,  $\widehat{R}$  is transitive and converse well-founded, and  $I_{\alpha_0}(A) = f$ . But this model of  $\mathcal{S}$  is also a model for  $ML_+^3$  since the properties of  $\widehat{R}$  make the Löb axiom (schema) true at  $\alpha_0$ , as is well-known (e.g., [17]). ⊥

REMARK 7.9. 1. That  $ML_+^3$  is sound with respect to finite transitive-irreflexive Kripke structures is true with an unremarkably standard proof (adapted from that for GL) that we omit.

2. In the case of  $ML^3$ , it is easier to obtain a maximal Henkin set as in 7.2 that only satisfies the first four requirements.<sup>20</sup> Having that in mind, the proof of the semantic completeness of  $ML^3$  is almost identical to the one above. ⊥

---

<sup>19</sup> Boolean logic, Axiom 7 and  $\vdash \Box D \rightarrow \Box \Box D$

<sup>20</sup> Recall that neither the equality symbol,  $=$ , nor constants or functions are part of the language of  $ML^3$ .

## 8. Interpolation

This section considers the reverse of cut for GTKS and GLTS: If we can prove  $\vdash A \rightarrow B$ , then we can find a formula  $C$  — the *interpolant* of  $A$  and  $B$  — that bears some relationship to  $A$  and  $B$ , so that both  $\vdash A \rightarrow C$  and  $\vdash C \rightarrow B$  are provable. To make “bears some relationship” precise, let us indicate by the symbol  $\langle A \rangle$  (cf. [14]), the *support of  $A$* , that is, the set of all free variables, Boolean variables and predicate letters appearing in  $A$ . For example,  $\langle p \rangle = \{p\}$ ,  $\langle \phi a \rangle = \{\phi, a\}$ ,  $\langle \Box \phi a \rangle = \{\phi\}$ , and  $\langle \perp \rangle = \emptyset$ . We extend the symbol so that  $\langle \Sigma \rangle = \bigcup_{A \in \Sigma} \langle A \rangle$ .

We now state a general version of Craig’s interpolation theorem. A partition of a sequent  $\Gamma \vdash \Delta$  is an expression  $(\Gamma_1, \Delta_1; \Gamma_2, \Delta_2)$ , where  $\Gamma = \Gamma_1 \cup \Gamma_2$  and  $\Delta = \Delta_1 \cup \Delta_2$ . A  $\Gamma_i$  and a  $\Delta_j$  can be empty, consistently however with the previous two equalities. Moreover,  $\Gamma_1 \cap \Gamma_2 = \emptyset = \Delta_1 \cap \Delta_2$ .

**THEOREM 8.1** (Craig’s interpolation theorem for GTKS and GLTS). *If  $\Gamma \vdash_{\mathcal{X}} \Delta$  where  $\mathcal{X} \in \{\text{GTKS}, \text{GLTS}\}$  then, for any chosen partition  $(\Gamma_1, \Delta_1; \Gamma_2, \Delta_2)$  of  $\Gamma \vdash_{\mathcal{X}} \Delta$ , an  $A$  (called an interpolant of  $\{\Gamma_1, \Delta_1; \Gamma_2, \Delta_2\}$ ) can be found satisfying:*

- (1)  $\Gamma_1 \vdash_{\mathcal{X}} A, \Delta_1$  and  $\Gamma_2, A \vdash_{\mathcal{X}} \Delta_2$ .
- (2)  $\langle A \rangle \subseteq \langle \Gamma_1 \cup \Delta_1 \rangle \cap \langle \Gamma_2 \cup \Delta_2 \rangle$ .

**PROOF.** The proof is done by induction on the order of derivation of  $\Gamma \vdash \Delta$ ; however, since it follows very similar lines to the well-known Maehara-Takeuti proof [18], it will be omitted. ◻

## 9. Arithmetical soundness of $\text{ML}^3$

First, let us suppose that the (free) variables  $v_0, v_1, \dots$  and the (bound) variables  $x_0, x_1, \dots$  are common to the languages of  $\text{ML}^3$  and of arithmetic. Second, let  $Pr(\Gamma \cdot \ulcorner \cdot \urcorner)$  denote the provability predicate in Peano Arithmetic. Following [2] we define:

**DEFINITION 9.1** (Realisation). 1. A *realistion* is a function  $*$  from a set of predicate letters to formulas (not necessarily atomic) in the language of arithmetic such that for all  $n$ , if  $\pi$  is an  $n$ -place predicate letter in the domain of  $*$ , then  $\pi^*$  is a formula in which exactly the first  $n$  variables (i.e.  $v_0, \dots, v_{n-1}$ ) occur free.

2. We say that  $F$  is a *formulaic expression* if  $F = F'(q_1, \dots, q_n)$  where  $F'$  is a formula and  $q_1, \dots, q_n$  are variables (free or bound). For example, if  $\pi$  is a predicate letter then  $\pi(x_2, v_4)$  is a formulaic expression.
3. A *realisation of a formulaic expression*  $F$  of  $\text{ML}^3$  is a realisation whose domain contains all predicate letters occurring in  $F$ .  $\dashv$

DEFINITION 9.2 (Arithmetical Interpretation —  $\text{ML}^3$ ). For every *formulaic expression*  $F$  of  $\text{ML}^3$  and realisation  $*$  of  $F$ , we define the arithmetical interpretation (or *translation*)  $F^*$  of  $F$  under  $*$  as follows:

1.  $\perp^* = \perp$ .
2. If  $F$  is the *formulaic expression*  $\pi(q_1, \dots, q_n)$  where  $\pi$  is a sentence letter, then  $F^*$  is the result  $\pi^*(q_1, \dots, q_n)$  of respectively substituting  $q_1, \dots, q_n$  for  $v_0, \dots, v_{n-1}$  in  $\pi^*$  (while systematically changing, if necessary, the bound variables of  $\pi^*$  in order to avoid any bound variable among  $q_1, \dots, q_n$  from being captured by a quantifier).<sup>21</sup>
3. If  $F = A \rightarrow B$  then  $F^* = A^* \rightarrow B^*$ .
4. If  $F = \forall x_i A[v_k := x_i]$  then  $F^* = \forall x_j (A[v_k := x_j])^*$ . Where, in the case that  $x_i$  occurs in  $(A[v_k])^*$ , then  $x_j$  is the first bound variable that does not. Otherwise  $i = j$ .<sup>22</sup>
5. If  $F = \Box A$  then  $F^* = Pr(\ulcorner \forall A^* \urcorner)$ .  $\dashv$

We only state the following lemma since its proof is quite standard.

LEMMA 9.3. *Let  $A$  and  $B$  be formulae. Then:*

1. *If  $\vdash_{\mathcal{P}\mathcal{A}} A$  then  $\vdash_{\mathcal{P}\mathcal{A}} Pr(\ulcorner \forall A \urcorner)$ .*
2.  *$\vdash_{\mathcal{P}\mathcal{A}} Pr(\ulcorner \forall (A \rightarrow B) \urcorner) \rightarrow (Pr(\ulcorner \forall A \urcorner) \rightarrow Pr(\ulcorner \forall B \urcorner))$ .*
3.  *$\vdash_{\mathcal{P}\mathcal{A}} Pr(\ulcorner \forall (Pr(\ulcorner \forall A \urcorner) \rightarrow A) \urcorner) \rightarrow Pr(\ulcorner \forall A \urcorner)$ .*
4.  *$\vdash_{\mathcal{P}\mathcal{A}} Pr(\ulcorner \forall A \urcorner) \rightarrow Pr(\ulcorner \forall Pr(\ulcorner \forall A \urcorner) \urcorner)$ .*

One can easily derive that  $\vdash_{\mathcal{P}\mathcal{A}} (\forall A)^* \leftrightarrow \forall A^*$  and that if  $x_i$  does not occur in  $(A[v_j])^*$  then  $(A[v_j := x_i])^* = (A[v_j])^*[v_j := x_i]$ . These two facts are helpful in the proof of the main result of our section (which we omit):

---

<sup>21</sup> Thus, if  $\pi$  is a 1-place sentence letter, and  $\pi^* = \forall x_1(v_0 + v_0 = x_1)$  then  $(\pi(x_1))^* = \forall x_2(x_1 + x_1 = x_2)$ .

<sup>22</sup> Thus, if  $\pi$  is a 2-place sentence letter and  $\pi^* = \forall x_1(v_0 + x_1 = v_1)$ , then in order to find  $(\forall x_1 \pi(v_5 := x_1, v_9))^*$  we first note that  $\pi(v_5, v_9)$  contains  $x_1$  and thus we need to examine  $\forall x_2(\pi(x_2, v_9))^*$  which is equal to  $\forall x_2 \forall x_1(x_2 + x_1 = v_9)$ .



LEMMA 9.4 (Arithmetical Soundness of  $\text{ML}^3$ ). *If  $\vdash_{\text{ML}^3} A$  then for every realisation  $*$ ,  $\vdash_{\mathcal{PA}} A^*$ .*

PROOF. The proof can be done by induction on proofs in  $\text{ML}^3$  (3.1): (Sketch) The interpretation,  $*$ , preserves tautologies since it commutes with Boolean connectives. Furthermore, the definition of  $*$  (and the fact the  $\mathcal{PA}$  is a first order system itself) will ensure that the interpretation of any instance of (2)–(4) in 3.1 is provable in  $\mathcal{PA}$ . Similarly, 9.3 implies that the interpretation of any instance of (5)–(7) is provable in  $\mathcal{PA}$ , and 9.3(1) implies that interpretations of boxed instances of (1)–(7) are also provable in  $\mathcal{PA}$ . Finally,  $*$  preserves implications by MP and Gen since it commutes with Boolean connectives and (practically) commutes with  $\forall$ .  $\dashv$

## 10. Conclusions

We have introduced a first-order extension  $\text{ML}^3$  of GL and an equivalent to it Gentzen-style system GLTS.

Unlike QGL, the “natural” first-order extension of GL,  $\text{ML}^3$  supports cut-elimination (its Gentzen-style proxy is cut free). The precise technical demonstration that cut elimination fails in QGL is found in [1] and it has its root in the fact that the underlying first-order language, unlike ours, allows  $\Box A$  to have free variables.

Moreover, we proved that  $\text{ML}^3$  supports Craig Interpolation, its modal box simulates the classical provability  $\vdash$  on classical formulae, is sound with respect to arithmetical interpretations, and is semantically complete with respect to converse well-founded transitive finite Kripke structures. Remains to be seen whether the latter property may prove to be a stepping stone toward proving  $\text{ML}^3$  to be arithmetically complete — as it was in Solovay’s proof for the case of GL — thus providing the first example of a usable first-order provability logic.

**Acknowledgement.** This research was partially supported by NSERC grant No. 8250.

## References

- [1] Avron, A., “On modal systems having arithmetical interpretations”, *Journal of Symbolic Logic*, 49 (1984), 3: 935–942. DOI: [10.2307/2274147](https://doi.org/10.2307/2274147)
- [2] Boolos, G., *The Logic of Provability*, Cambridge University Press, Cambridge, 1993.
- [3] Fine, K., “Failures of the interpolation lemma in quantified modal logic”, *Journal of Symbolic Logic*, 44 (1979), 2: 201–206. DOI: [10.2307/2273727](https://doi.org/10.2307/2273727)
- [4] Fitting, M., and R. L. Mendelsohn, *First-Order Modal Logic*, Kluwer Academic Publishers, Dordrecht, 1998.
- [5] Gries, D., and F. B. Schneider, “Adding the everywhere operator to propositional logic”, *Journal Logic Computat.*, 8 (1998), 1: 119–129. DOI: [10.1093/logcom/8.1.119](https://doi.org/10.1093/logcom/8.1.119)
- [6] Kibedi, F., and G. Tourlakis, “A modal extension of weak generalisation predicate logic”, *Logic Journal of the IGPL*, 14 (2006), 4: 591–621. DOI: [10.1093/jigpal/jzl025](https://doi.org/10.1093/jigpal/jzl025)
- [7] Kripke, S. A., “A completeness theorem in modal logic”, *Journal of Symbolic Logic*, 24 (1959): 1–14. DOI: [10.2307/2964568](https://doi.org/10.2307/2964568)
- [8] Leivant, D., “On the proof theory of the modal logic for arithmetic provability”, *Journal of Symbolic Logic*, 46 (1981), 3: 531–538. DOI: [10.2307/2273755](https://doi.org/10.2307/2273755)
- [9] Mendelson, E., *Introduction to Mathematical Logic*, 3rd edition, Wadsworth & Brooks, Monterey, California, 1987.
- [10] Montagna, F., “The predicate modal logic of provability”, *Notre Dame Journal of Formal Logic*, 25 (1984): 179–189. DOI: [10.1305/ndjfl/1093870577](https://doi.org/10.1305/ndjfl/1093870577)
- [11] Sambin, G., and S. Valentini, “A modal sequent calculus for a fragment of arithmetic”, *Studia Logica*, 39 (1980), 2/3: 245–256. DOI: [10.1007/BF00370323](https://doi.org/10.1007/BF00370323)
- [12] ———, “The modal logic of provability. The sequential approach”, *Journal of Philosophical Logic*, 11 (1982), 3: 311–342. DOI: [10.1007/BF00293433](https://doi.org/10.1007/BF00293433)
- [13] Sasaki, K., “Löb’s axiom and cut-elimination theorem”, *Academia Mathematical Sciences and Information Engineering Nanzan University*, 1 (2001): 91–98.
- [14] Schütte, K., *Proof Theory*, Springer-Verlag, New York, 1977.
- [15] Schwartz, Y., and G. Tourlakis, “On the proof-theory of two formalisations of modal first-order logic”, *Studia Logica*, 96 (2010), 3: 349–373. DOI: [10.1007/s11225-010-9294-y](https://doi.org/10.1007/s11225-010-9294-y)
- [16] Shoenfield, J. R., *Mathematical Logic*, Addison-Wesley, Reading, Massachusetts, 1967.
- [17] Smoryński, C., *Self-Reference and Modal Logic*, Springer-Verlag, New York, 1985.

- [18] Takeuti, G., *Proof Theory*, North-Holland, Amsterdam, 1975.
- [19] Tournakis, G., *Lectures in Logic and Set Theory; Volume 1: Mathematical Logic*, Cambridge University Press, Cambridge, 2003.
- [20] Tournakis, G., and F. Kibedi, “A modal extension of first order classical logic. Part I”, *Bulletin of the Section of Logic*, 32 (2003), 4: 165–178. [http://www.filozof.uni.lodz.pl/bulletin/pdf/32\\_4\\_1.pdf](http://www.filozof.uni.lodz.pl/bulletin/pdf/32_4_1.pdf)
- [21] \_\_\_\_\_, “A modal extension of first order classical logic. Part II”, *Bulletin of the Section of Logic*, 33 (2004), 1: 1–10. [http://www.filozof.uni.lodz.pl/bulletin/pdf/33\\_1\\_1.pdf](http://www.filozof.uni.lodz.pl/bulletin/pdf/33_1_1.pdf)
- [22] Valentini, S., “The modal logic of provability. Cut-elimination”, *Journal of Philosophical Logic*, 12 (1983), 4: 471–476. DOI: 10.1007/BF00249262
- [23] Vardanyan, V. A., “On the predicate logic of provability”, preprint of the Scientific Council on the Complexity Problem “Cybernetics”, Academy of Sciences of the USSR (1985, in Russian).

YEHUDA SCHWARTZ  
Toronto Stn Q  
P.O. Box 1204, 27 St Clair Ave E  
Toronto M4T 1L0, ON, Canada  
[jude\\_math@yahoo.com](mailto:jude_math@yahoo.com)

GEORGE TOURLAKIS  
Dept. of EECS, Lassonde School of Engineering  
York University  
4700 Keele St.  
Toronto M3J 1P3, ON, Canada  
[gt@cse.yorku.ca](mailto:gt@cse.yorku.ca)