**B**ulletin of the **S**ection of **L**ogic
Volume 44:3/4 (2015), pp. 131–147

http://dx.doi.org/10.18778/0138-0680.44.3.4.03

Feng Gao
George Tourlakis

# A SHORT AND READABLE PROOF OF CUT ELIMINATION FOR TWO FIRST-ORDER MODAL LOGICS*

**Abstract**

A well established technique toward developing the proof theory of a Hilbert-style modal logic is to introduce a Gentzen-style equivalent (a Gentzenisation), then develop the proof theory of the latter, and finally transfer the metatheoretical results to the original logic (e.g., [1, 6, 8, 18, 10, 12]). In the first-order modal case, on one hand we know that the Gentzenisation of the straightforward first-order extension of GL, the logic QGL, admits no cut elimination (if the rule is included as primitive; or, if not included, then the rule is not admissible [1]). On the other hand the (cut-free) Gentzenisations of the first-order modal logics $M^3$ and $ML^3$ of [10, 12] do have cut as an admissible rule. The syntactic cut admissibility proof given in [18] for the Gentzenisation of the propositional provability logic GL is extremely complex, and it was the basis of the proofs of cut admissibility of the Gentzenisations of $M^3$ and $ML^3$, where the presence of quantifiers and quantifier rules added to the complexity and length of the proof.

A recent proof of cut admissibility in a cut-free Gentzenisation of GL is given in [5] and is quite short and easy to read. We adapt it here to revisit the proofs for the cases of $M^3$ and $ML^3$, resulting to similarly short and easy to read proofs, only slightly complicated by the presence of quantification and its relevant rules.

*Keywords*: Modal logic, GL, QGL, first-order logic, proof theory, cut elimination, cut admissibility, provability logic.

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

# 1. Introduction

The propositional *provability logic* GL expresses provability within Peano Arithmetic (PA) as was established by Solovay ([14]). The proof-theory of GL has been studied effectively by proxy, introducing a Gentzen style logic equivalent to GL – a Gentzenisation of GL, cf. [8, 18] – and developing the proof theory of the latter, such as *cut elimination*, *Craig interpolation*, *disjunction property*, etc.

Logicians have turned their attention to first-order modal logics in the search for a predicate provability logic. Results of Vardanyan (in [3]) and Montagna [7] showed that the "natural" first-order extension of GL, known as *Quantified GL* (QGL), is not as "nice": Its Gentzenisation *provably* is not a provability logic (loc. cit.), but it even fails other nice properties such as cut elimination ([1]) and Craig interpolation; however it satisfies the disjunction property ([1]).

QGL's inability to support cut elimination must be attributed to its *language*: $\Box A$ has as free variables all those free in $A$. In fact, an almost identical recently introduced first-order extension of GL (the $ML^3$ of [12]) differs from QGL in that its language requires that $\Box A$ is a *sentence* for all $A$.[1] In loc. cit. a proof of cut elimination of its Gentzenisation (the GLTS defined in Section 2) is given in full detail (as well as a proof of Craig interpolation and a proof of a special case of the disjunction property). It must be stated that the genesis of $ML^3$ was *not* aimed at tweaking QGL to restore cut elimination. Rather, $ML^3$ is an evolution of $M^3$ introduced much earlier in [16, 17] in order to formalise the classical first-order "$\vdash$" within a first-order modal logic. Loc. cit. proved that such formalisation was achieved in $M^3$ (*main conservation result*), by proving semantically through Kripke models of $M^3$ that for classical $\Gamma, A$ and $B$,[2] we have $\Gamma, A \vdash B$ classically iff $\Gamma, \Box\Gamma \vdash \Box A \to \Box B$ – the latter proof carried in $M^3$. In [10] the (syntactic) proof theory of $M^3$ was developed by defining a cut-free Gentzenisation for it, called GTKS, and proving that cut is an admissible rule. Logic $ML^3$ is a common first-order extension of GL and

---

[1] $ML^3$ has an additional axiom schema, absent from QGL: $\Box A \to \Box\forall x A$. This is an essential independent axiom – as we prove in [15] – needed for the "conservation property" to hold in $ML^3$.

[2] Capital Greek letters such as $\Gamma, \Delta, \Psi, \Phi, \Omega$ denote *sets* of formulae; $\Box\Gamma$ is by definition, $\{\Box A : A \in \Gamma\}$. For a formula $A$, $\forall A$ is its universal closure, obtained by prefixing $A$ by $\forall x_1 \forall x_2 \ldots \forall x_i \ldots \forall x_n$, for each $x_i$ that is free in $A$. $\forall\Gamma$ means $\{\forall A : A \in \Gamma\}$.

$M^3$ introduced in [12], essentially obtained from the latter by adding Löb's axiom without changing the language. The cut rule was proved to be admissible in its cut-free Gentzenisation, and once again the conservation result was syntactically proved for $ML^3$. This variant of QGL not only supports cut elimination (by simulating cut in its cut-free Gentzenisation) but also is complete with respect to finite, transitive reverse well-founded Kripke structures. By contrast, QGL is not complete with respect to any set of Kripke structures ([7]).

This paper revisits and significantly simplifies the proofs given in [10, 12] that cut is an admissible rule. The proofs in loc. cit., especially the one in the 2nd reference, which is based on Valentini's ([18]) proof for the Gentzenisation of GL, are extremely complex. The ones given here are based on the recent proof in [5] (for GL) and are as simple as one would hope for cut elimination/emulation to be.

## 2. Two Gentzen-style modal first-order logics

[10, 12] defined two first-order modal Hilbert-style logics, $M^3$ and $ML^3$, and begun building their proof theory. To this end two Gentzen-style cut-free logics were introduced, GTKS and GLTS, and were proved to be equivalent to $M^3$ and $ML^3$ respectively. We revisit here the two cut admissibility results proved in loc. cit. offering greatly simplified proofs.

The rules for the Gentzenisations of $M^3$ and $ML^3$ are given in the next two definitions (cf. [10, 12]). Upper case Latin letters stand for formulae while upper case Greek letters $\Gamma, \Delta, \Psi, \Sigma$ (and other choices that are *not* also Latin capital letters) stand for *finite sets* of formulae; so do primed such letters. The expression $\Gamma \vdash \Delta$ is called a *sequent* and intuitively says that the set of hypotheses (formulae) in $\Gamma$ proves the *disjunction* of the formulae in $\Delta$. $\Gamma$ is the *antecedent* part of the sequent, while $\Delta$ is the *succedent*. "$\Gamma, A$" and "$A, \Gamma$" mean $\Gamma \cup \{A\}$.

We will not repeat the description of the common language of all four logics ($M^3$, $ML^3$, GTKS and GLTS) in detail, but we will revisit the less standard points here. In fact, we will not define $M^3$ or $ML^3$, since the sole purpose of this paper is to offer simplified cut admissibility proofs for GTKS and GLTS; the latter two logics we define here in detail.

The primary connectives are $\bot, \rightarrow, \forall, \Box$. There are two types of (object) variables, *free* ($a, b, c, a', c'', a_0, b_{12}$, etc.), and *bound* ($x, y, z, x', y'', z_0$, $x_{12}$, etc.). The syntax of formulae ensures that $\Box A$ is a *sentence*, for all formulae $A$.[3] The expression, $\Box A$ is *metanotation* for the expression obtained from $A$ as follows: (1) Replace all free variables that occur in $A$ by the lexicographically smallest[4] *unused* (in $A$) bound variables $x_{j_1}, \ldots, x_{j_k}$; this results to an expression we will call $A'$. (2) Let $\alpha$ represent the string formed by arranging the used in (1) bound variables in their lexicographic order. (3) Then "$\Box A$" names the string $\Box\alpha A'$. Note that if $A$ has no free variables, then the meta name $\Box A$ names itself (that is, $A$ is the same string as $A'$ and $\alpha$ is empty).[5]

For any expression $F$,[6] $F[a]$ or $F[x]$ indicates that we want to pay attention to the free variable $a$ or bound variable $x$ that possibly occur in $F$. In the context of the notation $F[a]$, $F[t]$ denotes the result of replacing $a$ by $t$, everywhere in $F$ – an operation on the expression $F$ that we will on occasion denote more explicitly by "$F[a := t]$". The $[a := t]$ operation has the highest priority, so, for example, $A \to B[a := x]$ stands for $A \to (B[a := x])$.

$\forall x A[x]$ (or just $\forall x A$) are metanotation for our familiar "for all values of $x$, $A[x]$ holds". Thus, *provided that $x$ does not occur in $A$*, $\forall x A[x]$ names $\forall x A[a := x]$, for some $a$ known from the context. Note that in the last expression $[a := x]$ applies to $A$ before $\forall x$ does.

Definition 2.1 (GTKS Rules [10]).

(1) Initial rules: $\Gamma, A \vdash \Delta, A$ and $\Gamma, \bot \vdash \Delta$, where $A$ is atomic.

(2) $\to$-*left* rule: $\dfrac{\Gamma, A \to \bot \vdash \Delta \qquad \Gamma, B \vdash \Delta}{\Gamma, A \to B \vdash \Delta}$, where $B$ is not $\bot$.

(3) $\to$-*right* rule: $\dfrac{\Gamma, A \vdash \Delta, B}{\Gamma \vdash \Delta, A \to B}$

(4) $\bot$-*right* rule: $\dfrac{\Gamma, A \vdash \Delta}{\Gamma \vdash \Delta, A \to \bot}$

---

[3]The motivation and rationale for this choice of an "opaque" $\Box$ vs. the "transparent" one in the case of QGL has been explained elsewhere ([16, 17, 10, 12, 11]) and will not be repeated here.

[4]The infinite set of bound variables is finitely generated as suggested above from the alphabet $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, ', x, y, z\}$, whose members we list here in the intended increasing order.

[5]This description and use of $\Box A$ as metanotation parallels the one in Bourbaki ([4]) for the meta-expression $\tau_x A$.

[6]This expression could be a formula $A$, a set of formulae $\Sigma$, or a sequent $\Psi \vdash \Omega$.

(5) $\bot$-*left* rule: $\dfrac{\Gamma \vdash \Delta, A}{\Gamma, A \to \bot \vdash \Delta}$

(6) $\forall$-*right* rule: $\dfrac{\Gamma \vdash \Delta, A\,[a]}{\Gamma \vdash \Delta, \forall x A\,[x]}$ – as long as $a$, the *eigenvariable* of the rule, does not occur in the conclusion ("denominator") of the rule.

(7) $\forall$-*left* rule: $\dfrac{\Gamma, A\,[a] \vdash \Delta}{\Gamma, \forall x A\,[x] \vdash \Delta}$

(8) The *modified* "TR" modal rule: $\dfrac{\forall\Gamma, \Box\Gamma \vdash A}{\Phi, \Box\Gamma \vdash \Box A, \Psi}$ ■

Definition 2.2 (GLTS Rules [12]).

(1)–(7) As for GTKS, but instead of TR see GLR below:

(8) The *modified* "GLR" modal rule: $\dfrac{\forall\Gamma, \Box\Gamma, \Box A \vdash A}{\Phi, \Box\Gamma \vdash \Box A, \Psi}$

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 (for example, formula $A \to B$ is the p.f. of rule (3)); rule (1) has $A$ as principal formula. The single formulae displayed in the "numerators" of (2)–(8) are the "minor formulae" (m.f.) of the rule (for example, formula $A$ and $B$ are the m.f. of rule (3)). A numerator sequent is a *premise* while the denominator sequent is the *conclusion* of the rule. $\Phi$ and $\Psi$ in rule (8) of both definitions are *weakening* and *strengthening* parts respectively. $\Box A$ in 2.2(8) is the *diagonal formula*. We call the rule (2) "Y-type" (as adjective) because of its shape. All the other rules are "I-type". ■

Remark 2.3. The departure from [8, 18] in using here $\forall\Gamma$ in the premise of TR and GLR, rather than using $\Gamma$, permits a central part of the proof (given in [10, 12]) – that GTKS and GLTS are equivalent to $M^3$ and $ML^3$ respectively – to conclude successfully. That is the part of the proof that derives in GTKS (and GLTS) the common axiom schema $\Box A \to \Box\forall x A$ of $M^3$ and $ML^3$. ■

Definition 2.4 (Theorems). A *theorem*, or *derived sequent*, is defined recursively to be one of:

(1) A sequent of one of the two types in 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 $m$ and $n$, then the former is derived with order $1 + \max(m, n)$.
(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 $m$, then the former is derived with order $1 + m$. ■

REMARK 2.5. The above recursive definition of theorems implicitly defines a *tree* – a *proof tree* – with root labeled by the theorem. This root has *one* (case where an I-rule was the last one applied) or *two subtrees* (case where the Y-rule was the last one applied), which have root(s) labeled by the *premise(s)* of the last rule used to derive the theorem. The leafs of the proof tree are labeled by the axioms.

A derivable sequent may be derived with many different proof trees, and therefore with many different orders as the latter depend on the particular proof we have in mind. Thus the sentence "$\Gamma \vdash \Delta$ is (a theorem) provable (or derivable) with order $m$" simply means that it *is* possible to derive said sequent with order $m$.

We note the absence of weakening/strengthening rules, unlike the original formulation of Gentzen's in the case of classical logic. This is so because it is desirable to introduce weakening and strengthening as *admissible* rather than as *primary* rules in a Gentzen logic, of which we aim to develop the proof theory. For example, proofs by induction on the *height of proof trees* are much simpler in the absence of such primary rules. This approach was earlier followed in [9],[7] where his weakening/strengthening "structural rules" are admissible, and was also present in [13, 8]. The second of the last two references incorporates weakening and strengthening parts in TR and GLR and in the rules under (1), just as we do. See also the proof of 2.7 below. ■

---

[7] Schütte, loc. cit. uses a generalisation of sequent calculus for first-order classical and Intuitionistic logic, where his "negative" and "positive" parts generalise Gentzen's "antecedent" and "succedent" formulae. Nevertheless, his techniques adapted to Gentzen's setting make it a straightforward matter to not require weakening/strengthening as *primary* rules.

The following theorems and corollaries hold in both GTKS and GLTS. The proofs are indebted to [9] and were adapted in the sequent setting in [10, 12]. We include the proofs for 2.6, 2.7 and 2.9 and omit the others as being similar.

THEOREM 2.6 (cf. [10, 12, 9]). *If* $(\Gamma \vdash \Delta)\,[a]$ *is provable with order* $m$ *and* $b$ *is some other free variable, then* $(\Gamma \vdash \Delta)\,[b]$ *is provable with order* $\leq m$.

PROOF. By induction on the order of derivation, $m$, of $(\Gamma \vdash \Delta)[a]$. For $m = 0$, $(\Gamma \vdash \Delta)[a]$ is an axiom. Then so is $(\Gamma \vdash \Delta)[b]$. For the induction step we prove the case for $m > 0$. Cases (2)–(8) are numbered by the rule number (Definition 2.1) of the *last rule applied in deriving* $(\Gamma \vdash \Delta)[a]$.

(2) $(\Gamma \vdash \Delta)[a] = \Gamma[a], A[a] \to B[a] \vdash \Delta[a]$. Thus the premises of the rule,[8] $\Gamma[a], A[a] \to \bot \vdash \Delta[a]$ and $\Gamma[a], B[a] \vdash \Delta[a]$, are each derived with orders $< m$. By the I.H., $\Gamma[b], A[b] \to \bot \vdash \Delta[b]$ and $\Gamma[b], B[b] \vdash \Delta[b]$ are derived with orders $< m$, thus $\Gamma[b], A[b] \to B[b] \vdash \Delta[b]$ is derived with order $\leq m$.

(3) $(\Gamma \vdash \Delta)[a] = \Gamma[a] \vdash \Delta[a], A[a] \to B[a]$. Then the premise $\Gamma[a], A[a] \vdash \Delta[a], B[a]$ is derived with order $< m$. By I.H. so is $\Gamma[b], A[b] \vdash \Delta[b], B[b]$, from which $\Gamma[b] \vdash \Delta[b], A[b] \to B[b]$ is derived with order $\leq m$.

(4), (5) We omit the similar cases for these rules.

(6) $(\Gamma \vdash \Delta)[a] = \Gamma[a] \vdash \Theta[a], \forall x A[a, x]$. The premise $\Gamma[a] \vdash \Theta[a], A[a, a_0]$ is derivable with order $< m$, where $a_0$ is the eigenvariable used. Let $a_1$ be a new variable that does not occur in $(\Gamma \vdash \Delta)[a]$ and is distinct from $b$. Applying the I.H. twice – first changing $a_0$ into $a_1$ and then $a$ into $b$ – we get $\Gamma[b] \vdash \Theta[b], A[b, a_1]$, which is derivable with order $< m$.[9] Thus, $\Gamma[b] \vdash \Theta[b], \forall x A[b, x]$ – i.e., $(\Gamma \vdash \Delta)[b]$ – is derivable with order $\leq m$, with eigenvariable $a_1$.

(7) $(\Gamma \vdash \Delta)[a] = \Theta[a], \forall x A[a, x] \vdash \Delta[a]$. The premise $\Theta[a], A[a, c] \vdash \Delta[a]$ is derivable with order $< m$ and so is $\Theta[b], A[b, c] \vdash \Delta[b]$ by I.H. Applying the rule to the latter we derive $\Theta[b], \forall x A[b, x] \vdash \Delta[b]$ with order $\leq m$.

---

[8] In each case of propagating the claim from order $< m$ to order $m$, we *indicate without comment* the p.f. for each rule considered, for example, $A[a] \to B[a]$ here.

[9] Since $a_0$ does not occur in $\Gamma[a] \vdash \Theta[a], \forall x A[a, x]$, $(\Gamma[a] \vdash \Theta[a], A[a, a_0])[a_0 := a_1] = \Gamma[a] \vdash \Theta[a], A[a, a_1]$.

(8) $(\Gamma \vdash \Delta)[a] = \Phi[a], \Box\Omega[a] \vdash \Box A[a], \Psi[a]$. The premise $\forall\Omega[a], \Box\Omega[a] \vdash A[a]$ is derivable with order $< m$. By the I.H. so is $\forall\Omega[b], \Box\Omega[b] \vdash A[b]$. Thus $\Phi[b], \Box\Omega[b] \vdash \Box A[b], \Psi[b]$ is derivable with order $\leq m$ by an application of the same rule. It is noted that, since boxed formulae have no free variables, $\Box A[a] = \Box A[b]$ and $\Box\Omega[a] = \Box\Omega[b]$; moreover $\forall\Omega[a] = \forall\Omega[b]$ since $\forall\Omega$ has no free variables either.
The last case was argued based on rule TR, but the proof based on rule GLR is entirely similar, the presence of the diagonal formula $\Box A$ in the antecedent of the premise not adding any complexity. ■

Theorem 2.7 (Weakening; cf. [10, 12, 13, 9]). *For either GTKS or GLTS, if $\Gamma \vdash \Delta$ is derived with order $m$ then $\Phi, \Gamma \vdash \Delta$ is derivable with order $\leq m$.*

Proof. By induction on the order of derivation, $m$, of $\Gamma \vdash \Delta$. For $m = 0$, $\Gamma \vdash \Delta$ is an axiom. Then so is $\Phi, \Gamma \vdash \Delta$. For the induction step we prove the case for $m > 0$. Cases (2)–(8) are numbered by the rule number of the last rule applied in deriving $\Gamma \vdash \Delta$.

(2) Suppose $\Gamma, A \to B \vdash \Delta$ is derived with order $m$. Thus the premises of the rule,[10] $\Gamma, A \to \bot \vdash \Delta$ and $\Gamma, B \vdash \Delta$, are each derived with orders $< m$. By the I.H., $\Phi, \Gamma, A \to \bot \vdash \Delta$ and $\Phi, \Gamma, B \vdash \Delta$ are derived with orders $< m$, thus $\Phi, \Gamma, A \to B \vdash \Delta$ is derived with order $\leq m$ using the same rule.

(3), (4), (5) We omit the similar cases for these rules.

(6) Let $\Gamma \vdash \Delta = \Gamma \vdash \Theta, \forall x A[x]$. The premise $\Gamma \vdash \Theta, A[a_0]$ is derivable with order $< m$, where $a_0$ is the eigenvariable used. By 2.6, $\Gamma \vdash \Theta, A[a_1]$ is derivable with order $< m$, where $a_1$ is a new variable that does not occur in $\Phi, \Gamma \vdash \Delta$. By the I.H. $\Phi, \Gamma \vdash \Theta, A[a_1]$ is derivable with order $< m$ and thus so is $\Phi, \Gamma \vdash \Theta, \forall x A[x]$ with order $\leq m$ and eigenvariable $a_1$.

(7) We omit this case as it is similar to the previous.

(8) $\Gamma \vdash \Delta = \Theta, \Box\Omega \vdash \Box A, \Psi$. The premise $\forall\Omega, \Box\Omega \vdash A$ is derivable with order $< m$. Thus $\Phi, \Theta, \Box\Omega \vdash \Box A, \Psi$ is derivable with order $\leq m$ by an application of TR.
This case was argued about rule TR (and the I.H. was not used), but the proof for rule GLR is entirely similar, the presence of the

[10] Once again, we implicitly indicate the p.f. in each case considered, for example, $A \to B$ here.

diagonal formula $\Box A$ in the antecedent of the premise not adding any complexity. ■

Theorem 2.8 (Strengthening; cf. [10, 12]). *For either GTKS or GLTS, if* $\Gamma \vdash \Delta$ *is derived with order* $m$ *then* $\Gamma \vdash \Delta, \Theta$ *is derivable with order* $\leq m$.

Proof. Similar to the proof of 2.7. ■

Theorem 2.9 (Inversion rules; cf. [10, 12, 13, 9]). *For either GTKS or GLTS, we have*

(1) *If* $\Gamma, A \to B \vdash \Delta$ *is derivable with order* $m$*, then each of* $\Gamma, A \to \bot \vdash \Delta$ *and* $\Gamma, B \vdash \Delta$ *are derivable with order* $\leq m$.

(2) *If* $\Gamma \vdash \Delta, A \to B$ *is derivable with order* $m$*, then* $\Gamma, A \vdash \Delta, B$ *is derivable with order* $\leq m$.

(3) *If* $\Gamma \vdash \Delta, A \to \bot$ *is derivable with order* $m$*, then* $\Gamma, A \vdash \Delta$ *is derivable with order* $\leq m$.

(4) *If* $\Gamma, A \to \bot \vdash \Delta$ *is derivable with order* $m$*, then* $\Gamma \vdash \Delta, A$ *is derivable with order* $\leq m$.

(5) *If* $\Gamma \vdash \Delta, \forall x A[x]$ *is derivable with order* $m$*, then* $\Gamma \vdash \Delta, A[a]$ *is derivable with order* $\leq m$ *(for any choice of* $a$*).*

Proof. By induction on the order of derivation $m$. We include the standard proof for a few cases and refer the reader to the literature for the ones we omit.

(1) If $\Gamma, A \to B \vdash \Delta$ is an axiom then so is $\Gamma \vdash \Delta$ since $A \to B$ is not atomic, and hence so are $\Gamma, A \to \bot \vdash \Delta$ and $\Gamma, B \vdash \Delta$.
For the induction step we have two cases:
   - Case where $A \to B$ is the p.f. of rule (2) that derived $\Gamma, A \to B \vdash \Delta$. Then the rule premises, $\Gamma, A \to \bot \vdash \Delta$ and $\Gamma, B \vdash \Delta$ are each derived with order $< m$ by 2.4.
   - Case where $A \to B$ is *not* the p.f. of the rule $(k)$ (for $k = 2, 3, 4, 5, 6, 7, 8$) that derived $\Gamma, A \to B \vdash \Delta$.
     Consider the subcase where the Y-rule was the last applied with p.f. $X \to Y$ other than $A \to B$, that is, $\Gamma = \Gamma', X \to Y$. The premises $\Gamma', A \to B, X \to \bot \vdash \Delta$ and $\Gamma', A \to B, Y \vdash \Delta$ are derivable with order $< m$ each.

By the I.H., the sequents $\Gamma', A \to \bot, X \to \bot \vdash \Delta$ and $\Gamma', B, X \to \bot \vdash \Delta$, as well as $\Gamma', Y, A \to \bot \vdash \Delta$ and $\Gamma', B, Y \vdash \Delta$ are also derivable with orders $< m$, Therefore, applying rule (2) to the first and third, and then to the second and fourth, we derive (with order $\leq m$) $\Gamma, A \to \bot \vdash \Delta$ and $\Gamma, B \vdash \Delta$, respectively.
Similar argument for the I-rules 3–7.
Finally, consider the subcase where the TR or GLR was used to derive $\Gamma, A \to B \vdash \Delta$. Here the subcase that $A \to B$ is a *side formula* cannot apply, since the s.f. are of the form $\forall X$ or $\Box X$. If on the other hand $A \to B$ is a weakening formula, then $\Gamma, A \to B \vdash \Delta$ was obtained with order $m$ from a sequent of the form $\Gamma' \vdash C$ that was derived with order $< m$. Applying TR (or GLR) to the latter, but changing the weakening part $A \to B$ to $B$, we obtain $\Gamma, B \vdash \Delta$ with order $\leq m$. Then we invoke again the rule on the same premise, this time applying the weakening part $A \to \bot$ to obtain $\Gamma, A \to \bot \vdash \Delta$ also with order $\leq m$.

(5) If $\Gamma \vdash \Delta, \forall x A[x]$ is an axiom then so is $\Gamma \vdash \Delta, A[a]$ for any choice of $a$ since $\forall x A[x]$ is not atomic.
For the induction step we consider first the case where $\forall x A[x]$ is the p.f. of rule (6) that derived $\Gamma \vdash \Delta, \forall x A[x]$. Then $\Gamma \vdash \Delta, A[a_0]$ is derived with order $< m$ by 2.4, where $a_0$ is the eigenvariable used. By 2.6, for any $a$, $\Gamma \vdash \Delta, A[a]$ is derived with order $< m$ as well.[11]
Say, on the other hand, that $\forall x A[x]$ is not the p.f. in the rule $(k)$ $(k = 2, 3, 4, 5, 6, 7)$ that derived $\Gamma \vdash \Delta, \forall x A[x]$.
If the Y-rule was used to derive the previous sequent, then $\Gamma = \Gamma', X \to Y$ and the premises used were $\Gamma', X \to \bot \vdash \Delta, \forall x A[x]$ and $\Gamma', Y \vdash \Delta, \forall x A[x]$, each being derivable with order $< m$. By the I.H. each of $\Gamma', X \to \bot \vdash \Delta, A[b]$ and $\Gamma', Y \vdash \Delta, A[b]$ is derivable with order $< m$ – using a $b$ that does not occur in $\Gamma' \cup \Delta \cup \{\forall x A[x], X, Y\}$. By rule (2), $\Gamma \vdash \Delta, A[b]$ is derivable with order $\leq m$, and an application of 2.6 allows us to use any free variable $a$ in the place of $b$.
The case of I-rules 3–7 is argued similarly.
Finally, let TR (or GLR) be the rule applied last to derive $\Gamma \vdash \Delta, \forall x A[x]$, from premise $\Gamma' \vdash C$, itself derived with order $< m$. Thus

[11] Recall that $a_0$ occurs nowhere in $\Gamma \vdash \Delta$, so the substitution $(\Gamma \vdash \Delta, A[a_0])[a_0 := a]$ will be localised to $A[a_0]$.

$\forall x A[x]$ must be a strengthening formula and reapplying TR (or GLR) with a different strengthening, $A[a]$ for any $a$ we may choose, derives $\Gamma \vdash \Delta, A[a]$ with order $\leq m$. ■

### 2.1. Reducibility

Definition 2.10. In either GTKS or GLTS we define that a sequent $\Gamma \vdash \Delta$ is *irreducible* if one of the following applies:

(1) $\bot \in \Gamma$.
(2) There exists an atomic formula, $A$, such that $A \in \Gamma \cap \Delta$.
(3) The members of $\Gamma$ are atomic or boxed and $\Delta$ is atomic, $\bot \notin \Gamma$ and $\Gamma \cap \Delta = \emptyset$.

We say that a sequent $\Gamma \vdash \Delta$ is *reducible* if that sequent is *not irreducible.* ■

Definition 2.11. In either GTKS or GLTS, and in the case of a reducible sequent, at least one rule from Definitions 2.1 and 2.2 applies backwards to yield a *predecessor* sequent. The predecessor relation between so related sequents, $\Gamma' \vdash \Delta'$ (predecessor) and $\Gamma \vdash \Delta$ we will denote by $\prec$, that is, $\Gamma' \vdash \Delta' \prec \Gamma \vdash \Delta$. ■

Remark 2.12. In either GTKS or GLTS the relation $\prec$ is well-founded – that is, there can be no infinite "descending" $\prec$-paths because each rule, (2)–(8), when applied "backwards" from $\Gamma \vdash \Delta$, *reduces* the number of occurrences of one of the connectives $\rightarrow, \forall, \Box$ in $\Gamma \vdash \Delta$. [12]

The case of TR/GLR calls for some more elaboration: Each backwards application *reduces the number of occurrences of* $\Box$ *in the succedent* and so after a finite number of (backwards) steps neither of the two will be applicable. Now, each of these two rules introduces a $\forall\Omega$ in the antecedent, which will be eventually depleted by reverse applications of rule (7). This latter rule does not introduce any *new* TR/GLR-specific p.f. to the *right* of $\vdash$ that were not already subformulae of $\Gamma \vdash \Delta$. Finally, we note that a reverse application of GLR introduces a $\Box A$ in the antecedent (diagonal formula), but this is not a p.f. for *any* rule, and causes no thread backwards.

[12] The Y-rule, applied backwards, still has the $\rightarrow$ connective in one of the predecessor sequents. However rule (5), applied backwards, will remove it.

Thus one can do induction along $\prec$ or on the *reducibility rank* – $RR(\Gamma \vdash \Delta)$ – of $\Gamma \vdash \Delta$, that is, the *path length* upwards from this sequent to an irreducible sequent. The minimal elements of this order are the irreducible sequents.

We note that, for any $\Phi$ and $\Psi$, we have $RR(\Phi, \Gamma \vdash \Delta, \Psi) \leq RR(\Gamma \vdash \Delta)$. If $\Gamma \vdash \Delta$ is derivable, then this is what 2.7 and 2.8 say. If not, then adding weakening (strengthening) to $\Gamma \vdash \Delta$ is effected by introducing it via applications of TR/GLR along a reverse path along $\prec$, from this sequent to an irreducible (but not an axiom) $\Gamma' \vdash \Delta'$; or by modifying the side formulae of $\Gamma' \vdash \Delta'$. Neither of these actions lengthen the path. ■

## 3. Cut Derivability in GTKS and GLTS

Proposition 3.1. *The following two statements are equivalent for every formula $A$ and any $\Gamma, \Delta, \Theta, \Omega, \Phi, \Psi$:*

*a. If $\Gamma \vdash \Delta, A$ and $A, \Theta \vdash \Omega$ then $\Gamma, \Theta \vdash \Delta, \Omega$ (cut admissibility).*
*b. If $A \to A, \Phi \vdash \Psi$ then $\Phi \vdash \Psi$.*

Proof. *a.* $\to$ *b.* Derivability of $A \to A, \Phi \vdash \Psi$ entails that of $A, \Phi \vdash \Psi$ and $\Phi \vdash \Psi, A$ (2.9, cases 1 and 4) and we are done by a.

*b.* $\to$ *a.* The assumption in 3.1. entails (by weakening/strengthening) the derivability of $\Gamma, \Theta \vdash \Delta, \Omega, A$ and $A, \Gamma, \Theta \vdash \Delta, \Omega$. By rule (2.1) we get $A \to A, \Gamma, \Theta \vdash \Delta, \Omega$. We are done by b. ■

Lemma 3.2 (Cut admissibility Lemma for GTKS). *For any formula $A$, if $A \to A, \Gamma \vdash \Delta$ is derivable, then so is $\Gamma \vdash \Delta$.*

Proof. The proof is by induction on the ordinal

$$\alpha = \omega^2 \cdot C + \omega \cdot RR + m \tag{1}$$

where $C$ is the *modified complexity* of $A$.[13] – this is the *primary* (P.I.) or main induction. A secondary induction (S.I.) is done along the $\prec$ relation on the $\Gamma \vdash \Delta$ "companion" of $A \to A$ – more accurately on $RR(\Gamma \vdash \Delta)$ –

[13] By *modified complexity* we mean the ordinal $\omega \cdot k + r$ where $k$ counts $\Box$ occurrences and $r$ counts the total of all $\to, \forall$ occurrences in $A$. Thus $(k, r) < \left(k+1, r'\right)$ and $(k, r) < (k, r+1)$ for all $k, r, r'$.

and on occasion we do a "local induction" (L.I.) on the order of derivation of $A, \Gamma \vdash \Delta$, which we typically call $m$ in this proof. Thus we will embark on a triple induction, where $C$ is not allowed to increase during the induction step of either the S.I. or L.I., and neither $C$ nor $RR$ are allowed to increase during the induction step of L.I.

Case 1. $A$ is atomic.

By invertibility (Theorem 2.9, case (1) followed by case (4)), both $A, \Gamma \vdash \Delta$ *and* $\Gamma \vdash \Delta, A$ are derivable. By a L.I. on the order of derivation $m$ of $A, \Gamma \vdash \Delta$ we prove the derivability of $\Gamma \vdash \Delta$.

(i) Basis. If $m = 0$, what if $\Gamma \vdash \Delta$ is not itself an axiom? Then $A \in \Delta$, so $\Gamma \vdash \Delta = \Gamma \vdash \Delta, A$ is derivable, contradicting our "what if".

(ii) Let us now take a L.I.H. and assume that $A, \Gamma \vdash \Delta$ is obtained by one of the rules (2)–(7) with order $m$. *Note that $A$ cannot be the p.f. in the application of such rules.*

- Case where the "Y-rule" derived $A, \Gamma \vdash \Delta$: Then some $A, \Gamma' \vdash \Delta$ and $A, \Gamma'' \vdash \Delta$ (the rule's premises) are derivable each with order $< m$, and the same is true, by weakening 2.7, for $A, \Gamma, \Gamma' \vdash \Delta$ and $A, \Gamma, \Gamma'' \vdash \Delta$.
  Since $\Gamma, \Gamma' \vdash \Delta, A$ and $\Gamma, \Gamma'' \vdash \Delta, A$ are also derivable by weakening, the local I.H. yields the derivability of each of $\Gamma, \Gamma' \vdash \Delta$ and $\Gamma, \Gamma'' \vdash \Delta$, and an application of the Y-rule derives $\Gamma, \Gamma \vdash \Delta = \Gamma \vdash \Delta$ as needed.[14]
- Case where one of the "I-rules" (3)–(7) derived $A, \Gamma \vdash \Delta$. This is similar to and slightly simpler than the Y-case.
  *Note that* $\Gamma' \vdash \Delta \prec \Gamma \vdash \Delta$ *and* $\Gamma'' \vdash \Delta \prec \Gamma \vdash \Delta$, *hence RR did not increase during this induction step (cf. also concluding part of Remark 2.12).*

(iii) $A, \Gamma \vdash \Delta$ is obtained by rule (8). Thus, $A$ is a weakening formula, but then $\Gamma \vdash \Delta$ is also derivable by omitting the weakening $A$ (the L.I.H. was not needed in this case).

Case 2. $A = B \to C$.

[14] Recall that what we are proving via this "local" induction is that if *both* $A, \Gamma \vdash \Delta$ *and* $\Gamma \vdash \Delta, A$, are provable *then* so is $\Gamma \vdash \Delta$. Thus, the acrobatics involving weakening are needed to ensure that the "and" holds: Even though, e.g., $A, \Gamma' \vdash \Delta$ is provable, we cannot necessarily expect that so is $\Gamma' \vdash \Delta, A$. But $\Gamma, \Gamma' \vdash \Delta, A$ *is* provable!

By 2.9, cases (1) and (4), we can also derive $B \to C, \Gamma \vdash \Delta$ and $\Gamma \vdash \Delta, B \to C$; and, again by invertibility, we can derive $S_1 = \Gamma \vdash \Delta, B$ – thus also $S'_1 = \Gamma \vdash \Delta, C, B$ – and $S_2 = C, \Gamma \vdash \Delta$ and also $S_3 = B, \Gamma \vdash \Delta, C$ by case (2). Now, we can derive $S_4 = B \to B, \Gamma \vdash \Delta, C$ from $S'_1$ and $S_3$ and rule (2); similarly, we can derive $C \to C, B \to B, \Gamma \vdash \Delta$ from $S_4$ and the obvious weakening of $S_2$. We can finally apply the P.I.H. twice to get $\Gamma \vdash \Delta$.

Case 3. $A = \forall x B$. Now, $S = \forall x B, \Gamma \vdash \Delta$ *and* $\Gamma \vdash \Delta, \forall x B$ are derivable by 2.9. We do a L.I. on the order of derivation $m$ of $S$, as before, to show that $\Gamma \vdash \Delta$ is derivable.

(i) Now, if $S$ is an axiom, then so is $\Gamma \vdash \Delta$ since $\forall x B$ is not atomic.

(ii) Otherwise, let first $A, \Gamma \vdash \Delta$ be obtained by one of the rules (2)–(7) with order $m$. We have the cases,

  (a) $A$ is the p.f. in the derivation of $A, \Gamma \vdash \Delta$. Then $B[a], \Gamma \vdash \Delta$ is derivable for some $a$ and so is $\Gamma \vdash \Delta, B[a]$ by 2.9, last case. Hence $B[a] \to B[a], \Gamma \vdash \Delta$ is, by rule (2), and we are done by the P.I.H. (the L.I.H was not needed here).

  (b) $A$ is *not* p.f. in the derivation of $A, \Gamma \vdash \Delta$, and $A, \Gamma \vdash \Delta$ is obtained by one of the rules (2)–(7) with order $m$.
  - Case where the Y-rule derived $A, \Gamma \vdash \Delta$: Exactly as in the corresponding case under (ii) of Case 1.
  - Case where some I-rule among 3–7 derived $A, \Gamma \vdash \Delta$. Again as in (ii) of Case 1.

  The same note regarding *RR* as in Case 1(ii) applies here as well.

(iii) $A, \Gamma \vdash \Delta$ is obtained by rule (8). Exactly as (iii) under Case 1.

Case 4. $A = \Box B$.

(I) $\Gamma \vdash \Delta$ is irreducible. Thus, $\Box B, \Gamma \vdash \Delta$ is derivable as an initial sequent, which means that $\Gamma \vdash \Delta$ is also an initial sequent.

(II) $\Gamma \vdash \Delta$ is reducible. We have two subcases:

  (i) $A, \Gamma \vdash \Delta$ is obtained by one of the rules (2)–(7) with order $m$. As $A$ cannot be p.f. in any of rules (2)–(7), sub-subcase iib of Case 3 applies, and we have nothing further to add here. The note inserted at the end of Case 1(ii) applies here as well: *RR* did not increase.

(ii) (Adapting Brighton's approach ([5]) in this case.) Case where the only applicable rule to $\Gamma \vdash \Delta$ is (8). By invertibility, $S = \Box B, \Gamma \vdash \Delta$ and $S^{'} = \Gamma \vdash \Delta, \Box B$ are derivable. Now, if $\Box B \in \Gamma$ or $\Box B \in \Delta$, then $S = \Gamma \vdash \Delta$ or $S' = \Gamma \vdash \Delta$ respectively, and we are done. So let $\Box B \notin \Gamma \cup \Delta$, and let us also pay no attention to the possibility that $\Box B$ is weakening/strengthening introduced by the TR rule, as then we are done immediately.
Thus, $S$ and $S'$ were obtained by proofs ending as:

$$\frac{S1}{S} = \frac{\forall B, \Box B, \forall \Gamma^{'}, \Box \Gamma^{'} \vdash D}{\Box B, \underbrace{\Phi, \Box \Gamma^{'}}_{\Gamma} \vdash \underbrace{\Box D, \Psi}_{\Delta}}$$

and

$$\frac{S2}{S'} = \frac{\forall \Gamma^{'}, \Box \Gamma^{'} \vdash B}{\Phi, \Box \Gamma' \vdash \Box D, \Psi, \Box B}$$

Now, the derivable $S1, S2$ above can also derive

$$S3 = \Box B, \Box \Gamma^{'} \vdash \Box D$$

$$S4 = \Box \Gamma^{'} \vdash \Box B$$

and

$$S2' = \Box B, \forall \Gamma^{'}, \Box \Gamma^{'} \vdash B, \text{ this by weakening,}$$

respectively. Now, we can obtain the derivable sequent $S5 = \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'} \vdash B$ from $S2'$ and $S4$ (via (2) and (5)), and thus also $S5' = \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'} \vdash \forall B$ by repeated application of rule (6) – note that the left hand side of $\vdash$ in $S5$ is closed. We can also obtain $S6 = \forall B, \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'} \vdash D$ from S1 and S4. Thus we can next obtain $S7 = \forall B \to \forall B, \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'} \vdash D$ from $S5'$ and $S6$. We can now apply the P.I.H. to obtain $S8 = \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'} \vdash D$ from S7 (recall that $\forall B$ has lower (modified) complexity than $\Box B$).
But $\forall \Gamma^{'}, \Box \Gamma^{'} \vdash D \prec \Phi, \Box \Gamma^{'} \vdash \Box D, \Psi = \Gamma \vdash \Delta$, thus, by S.I.H., $\forall \Gamma^{'}, \Box \Gamma^{'} \vdash D$ is derivable. This, via TR, derives $\Phi, \Box \Gamma^{'} \vdash \Box D, \Psi = \Gamma \vdash \Delta$. ■

Lemma 3.3 (Cut admissibility Lemma for GLTS). *For any formula $A$, if $A \to A, \Gamma \vdash \Delta$ is derivable, then so is $\Gamma \vdash \Delta$.*

Proof. As in the proof of 3.2 except for Case 4(IIii): The only applicable rule to $\Gamma \vdash \Delta$ is GLR. By invertibility, $S = \Box B, \Gamma \vdash \Delta$ and $S^{'} = \Gamma \vdash \Delta, \Box B$ are derivable. Now, if $\Box B \in \Gamma$ or $\Box B \in \Delta$, then $S = \Gamma \vdash \Delta$ or $S' = \Gamma \vdash \Delta$ respectively, and we are done. So let $\Box B \notin \Gamma \cup \Delta$, and let us also pay no attention to the possibility that $\Box B$ is weakening/strengthening introduced by the GLR rule, as then we are done immediately.

Thus, $S$ and $S'$ were obtained by proofs ending as:

$$\frac{S1}{S} = \frac{\forall B, \Box B, \forall \Gamma^{'}, \Box \Gamma^{'}, \Box D \vdash D}{\Box B, \underbrace{\Phi, \Box \Gamma^{'}}_{\Gamma} \vdash \underbrace{\Box D, \Psi}_{\Delta}}$$

and

$$\frac{S2}{S'} = \frac{\forall \Gamma^{'}, \Box \Gamma^{'}, \Box B \vdash B}{\Phi, \Box \Gamma' \vdash \Box D, \Psi, \Box B}$$

Now, the derivable $S1, S2$ above can also derive

$$S3 = \Box B, \Box \Gamma^{'} \vdash \Box D$$

$$S4 = \Box \Gamma^{'} \vdash \Box B$$

respectively. Now, we can obtain the derivable sequent $S5 = \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'} \vdash B$ from $S2$ and $S4$ (via (2) and (5)), and thus also $S5' = \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'} \vdash \forall B$ by repeated application of rule (6) – note that the left hand side of $\vdash$ in $S5$ is closed. We can also obtain $S6 = \forall B, \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'}, \Box D \vdash D$ from S1 and S4. Thus we can next obtain $S7 = \forall B \to \forall B, \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'}, \Box D \vdash D$ from $S5'$ and $S6$. We can now apply the P.I.H. to obtain $S8 = \Box B \to \Box B, \forall \Gamma^{'}, \Box \Gamma^{'}, \Box D \vdash D$ from S7 (recall that $\forall B$ has lower (modified) complexity than $\Box B$). But $\forall \Gamma^{'}, \Box \Gamma^{'}, \Box D \vdash D \prec \Phi, \Box \Gamma^{'} \vdash \Box D, \Psi = \Gamma \vdash \Delta$, thus, by S.I.H., $\forall \Gamma^{'}, \Box \Gamma^{'}, \Box D \vdash D$ is derivable. This, via GLR, derives $\Phi, \Box \Gamma^{'} \vdash \Box D, \Psi = \Gamma \vdash \Delta$. ■

Theorem 3.4 (Cut admissibility for GTKS and GLTS). *In each of GTKS and GLTS, if $\Gamma \vdash \Delta, A$ and $A, \Theta \vdash \Omega$, then $\Gamma, \Theta \vdash \Delta, \Omega$.*

Proof. By 3.2 and 3.3 via 3.1. ■

## References


[1] A. Avron, *On modal systems having arithmetical interpretations*, **J. of Symb. Logic** 49, no. 3 (1984), pp. 935–942.

[2] Jon Barwise (ed.), **Handbook of Mathematical Logic**, North-Holland, Amsterdam, 1978.

[3] George Boolos, **The Logic of Provability**, Cambridge University Press, Cambridge, 1993.

[4] N. Bourbaki, **Éléments de Mathématique; Théorie des Ensembles**, Hermann, Paris, 1966.

[5] Jude Brighton, **Cut Elimination for GLS Using the Terminability of its Regress Process**, `DOI 10.1007/s10992-015-9368-4`, March 2015.

[6] Daniel Leivant, *On the proof theory of the modal logic for arithmetic provability*, **Journal of Symbolic Logic** 46, no. 3 (1981), pp. 531–538.

[7] Franco Montagna, *The predicate modal logic of provability*, **Notre Dame J. of Formal Logic** 25 (1984), pp. 179–189.

[8] G. Sambin and S. Valentini, *The Modal Logic of Provability. The Sequential Approach*, **Journal of Philosophical Logic** 11, no. 3 (1982), pp. 311–342.

[9] K. Schütte, **Proof Theory**, Springer-Verlag, New York, 1977.

[10] Y. Schwartz and G. Tourlakis, *On the proof-theory of two formalisations of modal first-order logic*, **Studia Logica** 96, no. 3 (2010), pp. 349–373.

[11] Y. Schwartz and G. Tourlakis, *A Proof Theoretic Tool for First-Order Modal Logic*, **Bulletin of the Section of Logic** 42:3/4 (2013), pp. 93–110.

[12] Y. Schwartz and G. Tourlakis, *On the Proof-Theory of a First-Order Version of GL*, **Logic and Logical Philosophy (LLP)** 23, no. 3 (2014), pp. 329–363.

[13] Helmut Schwichtenberg, *Proof theory: Some applications of cut-elimination*, in Barwise [2], pp. 867–895.

[14] R. Solovay, *Provability Interpretations of Modal Logics*, **Israel Journal of Mathematics** 25 (1976), pp. 287–304.

[15] G. Tourlakis, *Some metatheoretical results about first-order modal logics*, Submitted (In refereeing), October 2015.

[16] G. Tourlakis and F. Kibedi, *A modal extension of first order classical logic – Part I*, **Bulletin of the Section of Logic** 32:4 (2003), pp. 165–178.

[17] G. Tourlakis and F. Kibedi, *A modal extension of first order classical logic – Part II*, **Bulletin of the Section of Logic** 33:1 (2004), pp. 1–10.

[18] S. Valentini, *The Modal Logic of Provability: Cut-Elimination*, **Journal of Philosophical Logic** 12, no. 4 (1983), pp. 471–476.