The Elimination of Atomic Cuts and the Semishortening Property for Gentzen's Sequent Calculus with Equality*

F.Parlamento, F.Previale
Department of Mathematics, Computer Science and Physics
University of Udine, via Delle Scienze 206, 33100 Udine, Italy.
Department of Mathematics
University of Turin, via Carlo Alberto 10, 10123 Torino, Italy
e-mail: *franco.parlamento@uniud.it*, *flavio.previale@unito.it*

14/6/2019

Mathematical Subject Classification: 03F05

Abstract

We study various extensions of Gentzen's sequent calculus obtained by adding rules for equality. One of them is singled out as particularly natural and shown to satisfy full cut elimination, namely also atomic cuts can be eliminated. Furthermore we tell apart the extensions that satisfy full cut elimination from those that do not and establish a strengthened form of the nonlenghtening property of Lifschitz and Orevkov.

1 Introduction

The most common way of treating equality in sequent calculus is to add to Gentzen's system in [4] appropriate sequents with which derivations can start, beside the logical axioms of the form $F \Rightarrow F$ (see for example [2], [13], [14] and [5]). In this way equality is considered and treated as a mathematical relation subject to specific axioms. For the standard choice of axioms, Gentzen's cut elimination theorem holds only in a weakened form: every derivation can be transformed into one which contains only cuts whose cut formula is atomic (see [10] pg 138 for example). That does not allow to obtain directly the wealth of applications that full cut elimination has, such as the conservativity of first order

*Work partially supported by funds PRIN-MIUR of Italy, Grant "Logica, Modelli, Insiemi" and the department PRID funding HiWei and presented to LC2016(Leeds))

logic with equality over first order logic without equality, or the disjunction and existence property for intuitionistic logic with equality. However, the calculus LK'_e in [13] shows how to overcome that limitation by choosing as axioms an appropriate set of sequents generated by means of the cut rule applied to atomic formulae. For, every sequent that is derivable from such axioms is actually cut-free derivable from them. As we will prove, provided very natural introductions rules for equality are adopted, also the applications of the cut rule needed to generated the LK'_e axioms can be dispensed with, so that sequent calculi for which full cut elimination holds are obtained. In the framework of Gentzen's systems, in which the structural rules of exchange, weakening, contraction and cut are separated from the logical ones, equality is dealt with in [7] by means of two left introduction rules replacing terms on both sides of a sequent. Although [7] announces the validity of cut elimination for the system, it actually deals only with its cut free part. On the other hand, in the framework of the sequent calculi free of structural rules, [9] shows how the initial sequents that concern equality can be turned into nonlogical rules in order to obtain calculi for which all the structural rules, including the cut rule, are admissible. Yet the application of such rules can eliminate equalities, so that obtaining the mentioned applications of cut elimination is not entirely straightforward. Earlier sequent systems for classical logic, free of structural rules, with rules for equality similar to some that we will study, were introduced in [16] and [6] (see also [3], that presents a system of that type also for intuitionistic logic). However, in the present paper we will deal only with extensions of the Gentzen's systems LJ and LK, leaving a treatment of the extensions of the sequent calculi free of structural rules to a further work. Our main purpose is to determine the introduction rules for equality, considered as a logical constant in the light of Leibnitz's indiscernibility principle, and show how this approach, developed in the framework of second order logic, leads very naturally to calculi for which full cut elimination holds. Among such calculi, we propose, as particularly appropriate, the systems, to be denoted by $LJ^=$ and $LK^=$, that are obtained by adding to LJ and LK the *Reflexivity Axiom* $\Rightarrow t = t$ and the following two left introduction rules for $=$:

$$\frac{\Gamma \Rightarrow \Delta, F[v/r]}{r = s, \Gamma \Rightarrow \Delta, F[v/s]} =_1 \qquad \text{and} \qquad \frac{\Gamma \Rightarrow \Delta, F[v/r]}{s = r, \Gamma \Rightarrow \Delta, F[v/s]} =_2$$

where F is a formula; $F[v/r]$ and $F[v/s]$, as in [15], denote the result of the replacement in F of all the occurrences of the variable v by the terms r or s and Γ, Δ are finite sequences of formulae, with $\Delta = \emptyset$ in the intuitionistic case. In fact we will prove that full cut elimination holds for $LJ^=$ and $LK^=$.

Looking at such systems in the light of Leibnitz's indescernibility principle confirms their adequacy and suggests other rules that could be added to LJ and LK to obtain equivalent calculi. Among such rules we have the following:

$$\frac{F[v/r], \Gamma \Rightarrow \Delta}{r = s, F[v/s], \Gamma \Rightarrow \Delta} =^l_1 \qquad \text{and} \qquad \frac{F[v/r], \Gamma \Rightarrow \Delta}{s = r, F[v/s], \Gamma \Rightarrow \Delta} =^l_2$$

The four *equality rules* $=_1$, $=_2$, $=_1^l$ and $=_2^l$ turn out to be all equivalent over the structural rules and the Reflexivity Axiom, but it is only for some of the possible choices that we obtain systems for which cut elimination holds. The proof that that is the case for $LK^=$ and $LJ^=$ can be given by extending Gentzen's original proof, but we will rather take advantage of the full cut elimination for LK'_e by showing that its axioms have cut free derivations in the systems that contain the structural rules, the Reflexivity Axiom and the two rules $=_1$ and $=_2$. As we will show, cut elimination holds also for the systems, to be denoted by $LJ_1^=$ and $LJ_2^=$, that are obtained from LJ by adding the Reflexivity Axiom and $=_1$ and $=_1^l$ or $=_2$ and $=_2^l$. Despite the similarity of the pair of rules $=_1$ and $=_2$ and the pair $=_1^l$ and $=_2^l$, the system obtained from LJ by adding the Reflexivity Axiom and both $=_1^l$ and $=_2^l$ does not satisfy cut elimination. That turns out to be the case also for the systems that are obtained from LJ by adding the Reflexivity Axiom together with $=_1$ and $=_2^l$ or $=_1^l$ and $=_2$.

Furthermore we will show that if all the four equality rules above are adopted, then we obtain a system $LJ_{12}^=$ for which cut elimination holds also if their application is required to be $\prec$-nonlengthening, with respect to any binary antisymmetric relation on terms $\prec$. We recall from [7], that an equality-inference as represented above is said to be $\prec$- nonlengthening if $s \not\prec r$. Actually, we will show that cut elimination holds for the system in which all the equality-inferences are required to be $\prec$-nonlengthening and all the $=_1$ and $=_1^l$-inferences are required to be also $\prec$-shorthening, namely to satisfy the stronger condition $r \prec s$. We call *semishortening* the derivations whose equality inferences satisfy such restrictions. Alternatively we can require that all the equality-inferences be $\prec$-nonlengthening and all the $=_2$ and $=_2^l$-inferences be also $\prec$-shorthening.

All the above results hold without any essential change for the classical version of the calculi considered, in particular for the classical version $LK^=$ of $LJ^=$ and $LK_{12}^=$ of $LJ_{12}^=$. The system $LK_{12}^=$ is equivalent to the already mentioned system G^e in [7], so that we have a proof that G^e satisfies cut elimination and furthermore, improving the result stated in [8], that any derivation in G^e can be transformed into a cut-free derivation in the same system of its endsequent that is $\prec$-semishortening, in the sense explained above for the intuitionistic case, with respect to any antisymmetric relation $\prec$.

1.1 Equality rules

On the ground of Leibnitz's indiscernibility principle, equality can be defined by letting $a = b$ to mean $\forall X(X(a) \leftrightarrow X(b))$. Thanks to the rules for $\forall$ and $\rightarrow$, $\forall X(X(a) \leftrightarrow X(b))$ is equivalent to $\forall X(X(a) \rightarrow X(b))$. The details of the derivation involved can be found in 7.1 in Appendix 1 ($A1.7.1$). Similarly for the details of the other derivations this section refers to. Thus as the definition of $a = b$ we can simply take $\forall X(X(a) \rightarrow X(b))$, so that equality stands on a par with the definition of $\wedge, \vee, \neg, \exists$ in terms of universal quantification and implication, spelled out, for example, in [11] pg. 67. From that definition of

$=$ and the rules of Gentzen's sequent calculus for $\forall$ and $\rightarrow$, we can derive the following left and right introduction rules for $=$:

$$\frac{\Lambda \Rightarrow F[v/r] \qquad F[v/s], \Gamma \Rightarrow \Delta}{r = s, \Lambda, \Gamma \Rightarrow \Delta} \; =^{(2)}\Rightarrow \qquad\qquad \frac{Z(r), \Gamma \Rightarrow Z(s)}{\Gamma \Rightarrow r = s} \; \Rightarrow=^{(2)}$$

where Z is a free predicate variable that does not occur in Γ, and $|\Delta| \leq 1$ ($A1.7.2$).

Conversely, the sequents $r = s \Rightarrow \forall X(X(r) \rightarrow X(s))$ and $\forall X(X(r) \rightarrow X(s)) \Rightarrow r = s$ are derivable by using the rules $=^{(2)}\Rightarrow$ and $\Rightarrow=^{(2)}$ ($A1.7.3$). Therefore the second order version of LJ supplemented by the rules $=^{(2)}\Rightarrow$ and $\Rightarrow=^{(2)}$, that we denote by $LJ^{(2)=}$, is an adequate sequent calculus to deal with equality in second order logic. Furthermore, the right introduction rule $\Rightarrow=^{(2)}$ turns out to be equivalent to the Reflexivity Axiom ($A1.7.4$). Thus we are led to consider the sequent calculus for first order logic with equality that is obtained from $LJ^{(2)=}$ by replacing $\Rightarrow=^{(2)}$ by the Reflexivity Axiom and requiring that all the formulae and terms involved be first order formulae and terms. We will denote by $==\Rightarrow$ and $\Rightarrow=$ the rule and axiom obtained in that way, and by $LJ^{(1)=}$the sequent calculus that is obtained by adding them to Gentzen's LJ. As we will see, full cut elimination holds for $LJ^{(1)=}$. However $LJ^{(1)=}$ is far from being a satisfactory sequent calculus for first order logic with equality, since the application of the rule $==\Rightarrow$ eliminates all the logical constants (including $=$) occurring in F. On that respect replacing in $LJ^{(1)=}$ the rule $==\Rightarrow$ by the equivalent rules $=_1$ and $=_2$ yields the much more appropriate equivalent system $LJ^{=}$ ($A1.7.5$).

Thanks to the cut rule, $==\Rightarrow$ is readily seen to be equivalent to the axiom $r = s, F[v/r] \Rightarrow F[v/s]$, so that the "standard" system $LJ + (\Rightarrow r = r) + (r = s, F[v/r] \Rightarrow F[v/s])$ of intuitionistic predicate logic with equality is equivalent to $LJ^{(1)=}$ and $LJ^{=}$ ($A1.7.6$). The simplest rule equivalent ($A1.7.9$) over the structural rules, to $==\Rightarrow$, that does not introduce any formula in the antecedent, namely that satisfies the characteristic feature of any sequent rule codifying the natural deduction rules, is the following *Congruence Rule*

$$\frac{\Gamma \Rightarrow r = s \qquad \Lambda \Rightarrow F[v/r]}{\Gamma, \Lambda \Rightarrow F[v/s]} \; CNG$$

Although the rule CNG might seem of scarce interest for the sequent calculus, since it eliminates equalities, it will play a crucial role in the proof of cut elimination for $LJ^{=}$. The reason is that, while it is admissible over the cut free part of $LJ^{=}$, the form of cut on equalities that incorporates is strong enough to make the full cut admissible.

2 Transformation of derivations into separated form

In the following LJ and LK will denote the sequent calculi introduced by Gentzen in [4], except that, as in [13], in the left introduction rule $\forall \Rightarrow$ for $\forall$ and in the right introduction rule $\Rightarrow \exists$ for $\exists$ the *free object variable* is replaced by an arbitrary term and, to avoid the use of the exchange rules that does not play any essential role, we will assume that the antecedent and the succedent of a sequent are finite multisets rather than finite sequences of formulae.

DEFINITION 2.1 *$LJ^=$ and $LK^=$ are obtained by adding to LJ and LK, the Reflexivity Axiom $\Rightarrow t = t$ and the* equality rules $=_1$ *and* $=_2$ *presented in the Introduction i.e.*

$$\frac{\Gamma \Rightarrow \Delta, F[v/r]}{r = s, \Gamma \Rightarrow \Delta, F[v/s]} =_1 \qquad \textit{and} \qquad \frac{\Gamma \Rightarrow \Delta, F[v/r]}{s = r, \Gamma \Rightarrow \Delta, F[v/s]} =_2$$

The principal formula $r = s$ in the presentation of $=_1$ and $s = r$ in the presentation of $=_2$, will be called the operating equality*, while $F[v/r]$ and $F[v/s]$ will be called the* input *and* output *formula respectively. An equality rule is* atomic *if its input/output formula is atomic.*

Concerning the given presentation of the equality rules it is important to note that, since $F[v/r]$ and $F[v/s]$ can always be represented as $(F[v/v'])[v'/r]$ and $(F[v/v'])[v'/s]$ for any v' that is new to F, r and s, the variable v may be assumed to lie outside any given finite set of variables, in particular to occur neither in r nor in s. Furthermore, for a a similar reason, given a representation $F[v/q[u/r]]$ of a formula, we may assume that u does not occur in F, so that $F[v/q[u/r]]$ coincides with $(F[v/q])[u/r]$.

At the purely equality level, $LJ^=$ and $LK^=$ are equivalent, namely a sequent $\Gamma \Rightarrow F$ is derivable in $LJ^=$ without applications of logical rules, if and only if it is derivable in $LK^=$, without applications of logical rules. In fact a straightforward induction on the height of derivations establishes the following:

PROPOSITION 2.1 *If a sequent $\Gamma \Rightarrow \Delta$ is derivable in $LK^=$ without applications of logical rules, then there is a formula F in Δ such that $\Gamma \Rightarrow F$ has a derivation without applications of logical rules, that contains only sequents with exactly one formula in the succedent. In particular Δ cannot be empty, so that $\Gamma \Rightarrow$ is not derivable in $LK^=$ without applications of logical rules.*

Proposition 2.1 motivates the following definition:

DEFINITION 2.2 *EQ is the calculus acting on sequents with one formula in the succedent, having the logical axioms $F \Rightarrow F$, the Reflexivity Axioms $\Rightarrow t = t$; the* weak left structural rules of weakening and contraction*:*

$$\frac{\Gamma \Rightarrow H}{F, \Gamma \Rightarrow H} \qquad \frac{F, F, \Gamma \Rightarrow H}{F, \Gamma \Rightarrow H}$$

the cut rule*:*

$$\frac{\Gamma \Rightarrow F \quad F, \Lambda \Rightarrow H}{\Gamma, \Lambda \Rightarrow H}$$

and the equality left introduction rules $=_1$ *and* $=_2$*:*

$$\frac{\Gamma \Rightarrow F[v/r]}{r = s, \Gamma \Rightarrow F[v/s]} \qquad \frac{\Gamma \Rightarrow F[v/r]}{s = r, \Gamma \Rightarrow F[v/s]}$$

.

Our proof of cut elimination for $LJ^=$ and $LK^=$ will split into two parts. First we show that every derivation can be transformed into one that consists of derivations in EQ followed by applications of weak structural rules, namely structural rules different from the cut rule, and logical rules only, and then that cut elimination holds for EQ.

DEFINITION 2.3 *A derivation in* $LJ^=$ *or* $LK^=$ *is said to be* separated *if it consists of derivations in EQ, followed by applications (possibly none) of logical and weak structural rules (both left and right). The definition applies to any system S that extends LJ or LK by the addition of some of the equality rules* $=_1$, $=_2$, $=^l_1$,$=^l_2$, $==>$ *and CNG, with EQ replaced by the system* $EQ(S)$ *that consists of the structural rules, the Reflexivity Axiom and the equality rules of S.*

PROPOSITION 2.2 *Every derivable sequent in* $LJ^=$ *or* $LK^=$ *has a separated derivation whose equality rules are* atomic*. The same holds for any system S that extends LJ or LK by the addition of the Reflexivity Axiom and at least one of the equality rules* $=_1$, $=_2$, $=^l_1$,$=^l_2$, $==>$ *and CNG.*

Proof We can base the proof on the result in [13] pg 39-40 that determines a set, let us call it E', of atomic sequents closed under the cut rule and term replacement, with the property that every sequent derivable from logical axioms and sequents in E' is derivable without using the cut rule. E' is defined as follows. A sequent is said to be *simple* if it can be obtained from the sequents expressing reflexivity, symmetry, transitivity and functional congruence of equality, i.e.

$$\Rightarrow t = t$$
$$r = s \Rightarrow s = r$$
$$r = s, s = t \Rightarrow r = t$$
$$r_1 = s_1, \ldots r_n = s_n \Rightarrow f(r_1, \ldots, r_n) = f(s_1, \ldots, s_n)$$

by means of the left structural rules and the cut rule. E' is obtained by adding to the set of simple sequents, those of the form

$$r_1 = s_1, \ldots, r_m = s_m, p(r'_1, \ldots, r'_n) \Rightarrow p(s'_1, \ldots, s'_n)$$

where, for $1 \leq i \leq n$, the sequent $r_1 = s_1, \ldots, r_m = s_m \Rightarrow r'_i = s'_i$ is simple. LK'_e and LJ'_e are the calculi obtained from LK and LJ by adding as axioms the sequents in E'. Since the sequents of the form $r = s \Rightarrow r = s$ are simple, given that they can be derived from $r = s \Rightarrow s = r$ and $s = r \Rightarrow r = s$ by means of a cut, the sequents expressing relational congruence, i.e.

$$r_1 = s_1, \ldots, r_n = s_n, p(r_1, \ldots, r_n) \Rightarrow p(s_1, \ldots, s_n)$$

belong to E'. Therefore the axioms of the standard system LK_e in [13] pg.37 are also axioms of LK'_e and LJ'_e , hence, by Proposition 7.2 in [13], the sequents of the form $r = s, F[v/r] \Rightarrow F[v/s]$ and $s = r, F[v/r] \Rightarrow F[v/s]$ are derivable in LK'_e and LJ'_e. Therefore, thanks to the cut rule, every derivation in $LK^=$ or $LJ^=$ can be transformed into a derivation in LK'_e or LJ'_e of its endsequent and, therefore, into a cut free derivation in the same system. The conclusion follows since the sequents in E' can be easily derived in EQ, as a matter of fact in EQ with $=_1$ and $=_2$ restricted to atomic input/output formulae (actually only one of these two rules is needed for that purpose). The second part follows immediately from the first since, as seen in the Introduction, all the equality rules are equivalent over the structural rules and the Reflexivity Axiom, so that, if S extends LJ or LK by the addition of some of the equality rules $=_1$, $=_2$, $=^l_1$,$=^l_2$, $==\!\!>$ and CNG, we have that S is equivalent to $LJ^=$ or $LK^=$ and $EQ(S)$ is equivalent to EQ. □

The proof of proposition 2.2 can also be given in a more direct way, that does not depend on Takeuti's result, by extending Gentzen's original proof for cut elimination. First one notes that, thanks to the cut rule, the equality rules can be derived from their atomic form, and then shows that the cut rule is admissible over its restriction to atomic cut formulae and that, if the premiss of an atomic equality inference or the premisses of an atomic cut have separated derivations, the same holds for their conclusions. That every sequent derivable in $LK^=$ or $LJ^=$ has a separated derivation in the same system follows by a straightforward induction on the height of derivations.The details can be found in Appendix 2.

3 Elimination of atomic cuts

By Proposition 2.2, to show that the cut rule is eliminable from derivations in $LJ^=$ or $LK^=$ it suffices to show that it can be eliminated from the derivations in EQ. Actually the proof of cut elimination to be given for EQ applies without

any change to EQ with atomic equality rules. Thus, as a consequence of Proposition 2.2, every derivation in $LJ^{=}$ or $LK^{=}$ can be transformed into a cut-free derivation of its endsequent with atomic equality inferences.

To establish cut elimination for EQ, we will make use of the following equality calculus EQ_N, where N stands for *natural*.

DEFINITION 3.1 *EQ_N is the calculus acting on sequents with one formula in the succedent, obtained from EQ by replacing the rules $=_1$ and $=_2$ with the rule CNG:*

$$\frac{\Gamma \Rightarrow r = s \qquad \Lambda \Rightarrow F[v/r]}{\Gamma, \Lambda \Rightarrow F[v/s]}$$

DEFINITION 3.2 *$cf.EQ$ and $cf.EQ_N$ denote the systems EQ and EQ_N deprived of the cut rule.*

PROPOSITION 3.1 *EQ and EQ_N are equivalent.*

Proof The following are derivations of $=_1$ and $=_2$ from CNG and of CNG from $=_1$:

$$\frac{r = s \Rightarrow r = s \qquad \Gamma \Rightarrow F[v/r]}{r = s, \Gamma \Rightarrow F[v/s]}\ \text{CNG}$$

$$\frac{\dfrac{s = r \Rightarrow s = r \qquad \Rightarrow s = s}{s = r \Rightarrow r = s}\ \text{CNG} \qquad \Gamma \Rightarrow F[v/r]}{s = r, \Gamma \Rightarrow F[v/s]}\ \text{CNG}$$

$$\frac{\Gamma \Rightarrow r = s \qquad \dfrac{\Lambda \Rightarrow F[v/r]}{r = s, \Lambda \Rightarrow F[v/s]}\ =_1}{\Gamma, \Lambda \Rightarrow F[v/s]}\ \text{cut}$$

□

3.1 Cut-elimination for EQ_N

Notation In the following $\Gamma \sharp F$ will denote any multiset of formulae from which Γ can be obtained by eliminating any number, possibly none, of occurrences of F.

PROPOSITION 3.2 *If $\Gamma \Rightarrow F$ and $\Lambda \sharp F \Rightarrow G$ are derivable in $cf.EQ_N$, then also $\Gamma, \Lambda \Rightarrow G$ is derivable in $cf.EQ_N$.*

Proof Let $\mathcal{D}$ and $\mathcal{E}$ be derivations in $cf.EQ$ of $\Gamma \Rightarrow F$, and $\Lambda \sharp F \Rightarrow G$ respectively. We have to show that there is a derivation $\mathcal{F}$ in $cf.EQ$ of $\Gamma, \Lambda \Rightarrow G$.

If $\Lambda \sharp F$ coincides with Λ, in particular if $\Lambda \sharp F$ is empty, or F occurs in Λ, then to obtain $\mathcal{F}$ it suffices to add to $\mathcal{E}$ the weakenings and, in the latter case, the

contractions needed to obtain $\Gamma, \Lambda \Rightarrow G$. Otherwise we proceed by induction on the height $h(\mathcal{E})$ of $\mathcal{E}$. If $h(\mathcal{E}) = 0$, then $\mathcal{E}$ reduces to $F \Rightarrow F$ and for $\mathcal{F}$ we can take $\mathcal{D}$ itself.

If $\mathcal{E}$ ends with a weak structural inference whose principal formula occurs in Λ, then the premiss of the last inference of $\mathcal{E}$ is of the form $\Lambda' \sharp F \Rightarrow G$, where Λ is obtained by adding a formula to Λ' or contracting two identical formulae of Λ'. Then we can apply the induction hypothesis, to $\mathcal{D}$ and the immediate subderivation of $\mathcal{E}$, to obtain a cut-free derivation of $\Gamma, \Lambda' \Rightarrow G$ and then the last inference of $\mathcal{E}$ (either a weakening or a contraction) to obtain the desired cut-free derivation of $\Gamma, \Lambda \Rightarrow G$.

If $\mathcal{E}$ ends with a weak structural inference whose principal formula is one of the occurrences of F in $\Lambda \sharp F$, that does not occur in Λ, then the premiss of the last inference of $\mathcal{E}$ is still of the form $\Lambda \sharp F \Rightarrow G$ and the conclusion follows directly by the induction hypothesis applied to $\mathcal{D}$ and the immediate subderivation of $\mathcal{E}$.

If $\mathcal{E}$ ends with a CNG-inference, then G has the form $H[v/s]$ and $\mathcal{E}$ can be represented as:

$$\frac{\begin{array}{cc}\mathcal{E}_0 & \mathcal{E}_1 \\ \Lambda_0 \sharp F \Rightarrow r = s & \Lambda_1 \sharp F \Rightarrow H[v/r]\end{array}}{\Lambda \sharp F \Rightarrow H[v/s]}$$

where Λ_0, Λ_1 coincides with Λ. By induction hypothesis we have cut-free derivations of $\Gamma, \Lambda_0 \Rightarrow r = s$ and $\Gamma, \Lambda_1 \Rightarrow H[v/r]$, from which $\mathcal{F}$ is obtained by applying the same CNG-inference and some contractions. □

COROLLARY 3.1 *If a sequent is derivable in EQ_N, then it is also derivable in $cf.EQ_N$.*

Proof By the previous Proposition, applied in the specific case in which $\Lambda \sharp F$ is Λ, F, it follows that the cut rule is admissible in $cf.EQ_N$ and therefore eliminable from derivations in EQ_N. □

Since the derivation of CNG in EQ given in Proposition 3.1 makes use of the cut rule we cannot conclude immediately from Corollary 3.1 that cut elimination holds for EQ as well.

3.2 Admissibility of CNG in $cf.EQ$

PROPOSITION 3.3 *The rule CNG is admissible in $cf.EQ$, namely, if $\Gamma \Rightarrow r = s$ and $\Lambda \Rightarrow F[v/r]$ are derivable in $cf.EQ$, then also $\Gamma, \Lambda \Rightarrow F[v/s]$ is derivable in $cf.EQ$.*

Proof Let $\mathcal{D}$ and $\mathcal{E}$ be derivations in $cf.EQ$ of $\Gamma \Rightarrow r = s$ and $\Lambda \Rightarrow F[v/r]$ respectively. We have to show that there is a derivation $\mathcal{F}$ of $\Gamma, \Lambda \Rightarrow F[v/s]$ in

$cf.EQ$. If r and s coincide, to obtain $\mathcal{F}$ it suffices to apply to the endsequent of $\mathcal{E}$ the appropriate weakenings needed to introduce Γ. Otherwise we proceed by induction on the height of $\mathcal{D}$, with respect to an arbitrary $\mathcal{E}$. In the base case $\mathcal{D}$ reduces to the axiom $r = s \Rightarrow r = s$. In that case as $\mathcal{F}$ we can take:

$$\dfrac{\begin{array}{c}\mathcal{E}\\ \Lambda \Rightarrow F[v/r]\end{array}}{r = s, \Lambda \Rightarrow F[v/s]}$$

which uses $=_1$. If $\mathcal{D}$ ends with a structural rule, to obtain $\mathcal{F}$ it suffices to apply the induction hypothesis to the immediate subderivation $\mathcal{D}_0$ of $\mathcal{D}$ and to $\mathcal{E}$ and then the last structural rule of $\mathcal{D}$.

If $\mathcal{D}$ ends with an $=_1$-inference, then it can be represented as:

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma' \Rightarrow r^\circ[u/p] = s^\circ[u/p]\end{array}}{p = q, \Gamma' \Rightarrow r^\circ[u/q] = s^\circ[u/q]}\ =_1$$

where u does not occur in F. Thus r and s are $r^\circ[u/q]$ and $s^\circ[u/q]$ respectively, and Γ is $p = q, \Gamma'$. Then we let $\mathcal{E}'$ be the following derivation:

$$\dfrac{\begin{array}{c}\mathcal{E}\\ \Lambda \Rightarrow F[v/r^\circ[u/q]]\end{array}}{p = q, \Lambda \Rightarrow F[v/r^\circ[u/p]]}\ =_2$$

that ends with a correct $=_2$-inference, since, given that u does not occur in F, we have that $F[v/r^\circ[u/q]]$ coincides with $(F[v/r^\circ])[u/q]$ and $F[v/r^\circ[u/p]]$ coincides with $(F[v/r^\circ])[u/p]$.

By induction hypothesis applied to $\mathcal{D}_0$ and $\mathcal{E}'$ there is a derivation $\mathcal{F}_0$ of $p = q, \Gamma', \Lambda \Rightarrow F[v/s^\circ[u/p]]$. As $\mathcal{F}$ we can then take the following derivation:

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{F}_0\\ p = q, \Gamma', \Lambda \Rightarrow F[v/s^\circ[u/p]]\end{array}}{p = q, p = q, \Gamma', \Lambda, \Rightarrow F[v/s^\circ[u/q]]}\ =_1}{p = q, \Gamma', \Lambda \Rightarrow F[v/s^\circ[u/q]]}$$

which uses a $=_1$-inference, that, as the $=_2$-inference of the previous derivation, is correct, since u does not occur in F, followed by a contraction.

If $\mathcal{D}$ ends with a $=_2$-inference the argument is entirely similar. □

3.3 Cut elimination for EQ

THEOREM 3.1 *If $\Gamma \Rightarrow F$ is derivable in EQ, then it is derivable also in $cf.EQ$. The same holds for EQ with the equality rules restricted to be atomic.*

Proof By Proposition 3.1 a derivation $\mathcal{D}$ of $\Gamma \Rightarrow F$ in EQ can be transformed into a derivation $\mathcal{D}'$ in EQ_N of $\Gamma \Rightarrow F$. By the eliminability of the cut-rule in EQ_N, $\mathcal{D}'$ can be transformed into a derivation $\mathcal{D}''$ in $cf.EQ_N$ of $\Gamma \Rightarrow F$. Finally, by the admissibility of CNG in $cf.EQ$, $\mathcal{D}''$ can be transformed into a derivation in $cf.EQ$ of $\Gamma \Rightarrow F$. That the result holds for EQ with atomic equality rules follows from the fact that F in Definition 3.1 and in the proofs of Proposition 3.1 and 3.3, as well as H in the proof of proposition 3.2 can be required to be atomic, without affecting any of the arguments. □

3.4 Cut elimination for $LJ^=$ and $LK^=$

From Proposition 2.2 and Theorem 3.1 we obtain the full cut elimination theorem for $LJ^=$ and $LK^=$.

THEOREM 3.2 *The cut rule is eliminable from derivations in $LJ^=$ and in $LK^=$. Moreover every derivation in $LJ^=$ or $LK^=$ can be transformed into a cut-free derivation of its endsequent, in which the equality inferences are atomic and precede the structural and logical ones.*

Proof Given a derivation $\mathcal{D}$ in $LJ^=$ or $LK^=$ of a sequent $\Gamma \Rightarrow \Delta$, by Proposition 2.2, $\mathcal{D}$ can be transformed into a separated derivation $\mathcal{D}'$ of $\Gamma \Rightarrow \Delta$ in the same system. $\mathcal{D}'$ consists of subderivations in EQ with atomic equality inferences followed by applications of weak structural and logical rules only. By Theorem 3.1 the applications of the cut rule to be found in such subderivations in EQ of $\mathcal{D}'$ can be eliminated, thus obtaining a cut-free derivation of $\Gamma \Rightarrow \Delta$ in $LJ^=$ or $LK^=$ with atomic equality inferences. The last part of the claim follows by the fact that every derivation $\mathcal{E}$ in $cf.EQ$ with atomic equality rules can be transformed into a derivation (of the same height) of its endsequent in which the atomic equality rules are applied before the structural rules. That can be shown by a straightforward induction on the structural depth $sd(\mathcal{E})$ of $\mathcal{E}$, defined as the sum of the structural depths of its equality inferences, where the structural depth of an equality inference is the number of structural inferences of $\mathcal{E}$ by which it is preceded. In fact, if $sd(\mathcal{E}) = 0$, then $\mathcal{E}$ has the desired property and, if $sd(\mathcal{E}) > 0$, it suffices to select any equality inference immediately preceded by a structural inference and permute the two of them to obtain a derivation of the endsequent of $\mathcal{E}$ (of the same height) and smaller (by one) structural depth. □

Remark Since the rules $=_1$ and $=_2$ are derivable in $LJ^{(1)=}$, without using the cut rule (Appendix 1, 7.6) from cut elimination for $LJ^=$ it follows immediately that cut elimination holds also for $LJ^{(1)=}$ introduced in 1.1. Clearly that holds also for the systems $LK^{(1)=}$ that differs from $LJ^{(1)=}$ only for being based on classical rather than intuitionistic logic.

Note The rule CNG is among those used in the extension of the system CERES in [1], pg.170. The idea of using the admissibility of the rule CNG

in EQ_N to prove Theorem 3.2 first appeared in [12]. However the proof of admissibility and the way of deriving the cut elimination theorem for $LJ^=$ and $LK^=$ given above are a substantial improvement of those to be found in [12]. In fact [12] established admissibility of CNG over EQ, with the rule $=_2$ replaced by the left symmetry rule (see 1.9), by means of a more complicated induction on the height of the derivation of the second, rather than the first, premiss. Then cut elimination was proved directly in a more involved way, instead of deriving it from the straightforward cut elimination for EQ_N.

4 Systems satisfying full cut elimination

4.1 Admissibility of $=_1^l$ and $=_2^l$ in $cf.EQ$

Since, as noticed in the Introduction and proved in Appendix 1, all the equality rules are equivalent over the Reflexivity Axiom and the structural rules, $=_1^l$ and $=_2^l$ are derivable in EQ. Therefore, since the cut rule is eliminable from derivations in EQ, we immediately have the following:

PROPOSITION 4.1 *The rules $=_1^l$ and $=_2^l$ are admissible in $cf.EQ$.*

DEFINITION 4.1 *Let EQ_1 be obtained from EQ by replacing $=_2$ by $=_1^l$ and EQ_2 be obtained from EQ by replacing $=_1$ by $=_2^l$. Furthermore let EQ_{12} be obtained from EQ by adding $=_1^l$ and $=_2^l$. $cf.EQ_1$, $cf.EQ_2$ and $cf.EQ_{12}$ denote EQ_1, EQ_2 and EQ_{12} deprived of the cut rule.*

4.2 Admissibility of $=_2$ in $cf.EQ_1$ and of $=_1$ in $cf.EQ_2$

In the following $E \equiv E'$ will denote syntactic equality between the terms or formulae that are denoted by E and E'. We will also make use of the simultaneous substitution of all the free occurrences of different variables with corresponding terms. For example $G[u/p, v/r]$, will stand for the result of the simultaneous replacement of all the free occurrences in G of u and v by the terms p and q respectively. As for $F[v/r]$, in using the representation $G[u/p, v/r]$ for a given formula, it is not restrictive to assume that u and v do not occur in any given finite set of variables, in particular neither in p nor in r. With the latter proviso $G[u/p, v/r]$ coincides with the result of the iterated substitution in whatever order, namely with $(G[u/p])[v/r]$ and $(G[u/r])[v/p]$.

As already noted in [7] an application of $=_1$ or $=_2$ and $=_1^l$ or $=_2^l$, with input formula $F[v/r]$, where v has $n > 1$ occurrences in F, can be replaced by n applications of the same rule, with the same operating equality, with v having exactly one occurrence in F, followed by $n - 1$ applications of the contraction rule.

We will say that an equality rule is a *singleton* equality rule if there is exactly one occurrence of the term r in the input formula that is replaced by s in order to obtain the output formula, namely if in its representations given above, there is exactly one occurrence of v in F.

Thus we have the following:

LEMMA 4.1 *Any of the equality rules* $=_1$ *and* $=_2$ *as well as* $=_1^l$ *and* $=_2^l$ *is derivable by means of the contraction rule from the corresponding singleton equality rule.*

DEFINITION 4.2 $cf.EQ_1^1$, $cf.EQ_2^1$ *and* $cf.EQ_{12}^1$ *are obtained from* $cf.EQ_1$, $cf.EQ_2$ *and* $cf.EQ_{12}^1$ *by replacing their equality rules by the corresponding singleton equality rules.*

PROPOSITION 4.2 $=_2$ *is admissible in* $cf.EQ_1$ *and* $=_1$ *is admissible in* $cf.EQ_2$.

Proof By Lemma 4.1 it suffices to prove that the singleton versions of $=_2$ or $=_1$ are admissible in the systems $cf.EQ_1^1$ or $cf.EQ_2^1$, namely that, for v having exactly one occurrence in F, the following hold:

a) if $\Gamma \Rightarrow F[v/r]$ is derivable in $cf.EQ_1^1$ then also $s = r, \Gamma \Rightarrow F[v/s]$ is derivable in $cf.EQ_1^1$ and

b) if $\Gamma \Rightarrow F[v/r]$ is derivable in $cf.EQ_2^1$, then also $r = s, \Gamma \Rightarrow F[v/s]$ is derivable in $cf.EQ_2^1$.

As for *a*), let $\mathcal{D}$ be a derivation in $cf.EQ_1^1$ of $\Gamma \Rightarrow F[v/r]$. We proceed by induction on the height $h(\mathcal{D})$ of $\mathcal{D}$ to show that in $cf.EQ_1^1$ there is a derivation $\mathcal{D}'$ of $s = r, \Gamma \Rightarrow F[v/s]$. If $h(\mathcal{D}) = 0$ then $\mathcal{D}$ reduces to $F[v/r] \Rightarrow F[v/r]$ or to $\Rightarrow t_0 = t_1[v/r]$, with $t_0 \equiv t_1[v/r]$, or to $\Rightarrow t_0[v/r] = t_1$, with $t_0[v/r] \equiv t_1$. In the former case as $\mathcal{D}'$ we can take:

$$\frac{F[v/s] \Rightarrow F[v/s]}{s = r, F[v/r] \Rightarrow F[v/s]} =_1^l$$

which obviously belongs to $cf.EQ_1^1$. If $\mathcal{D}$ reduces to $\Rightarrow t_0 = t_1[v/r]$, with $t_0 \equiv t_1[v/r]$, as $\mathcal{D}'$ we can take:

$$\frac{\Rightarrow t_1[v/s] = t_1[v/s]}{s = r \Rightarrow t_0 = t_1[v/s]} =_1$$

which is correct since $t_0 \equiv t_1[v/r]$ and belongs to $cf.EQ_1^1$. The case in which $\mathcal{D}$ reduces to $\Rightarrow t_0[v/r] = t_1$, with $t_0[v/r] \equiv t_1$, is entirely similar.

If $h(\mathcal{D}) > 0$, and $\mathcal{D}$ ends with a structural rule, then the conclusion is a straightforward consequence of the induction hypothesis. If $\mathcal{D}$ ends with an $=_1$-inference, then we distinguish the following three cases.

Case 1. The unique occurrence of r to be replaced by s is already present in the premiss of the last inference of $\mathcal{D}$. Then $\mathcal{D}$ can be represented as:

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma' \Rightarrow F^\circ[u/p, v/r]\end{array}}{p = q, \Gamma' \Rightarrow F^\circ[u/q, v/r]} =_1$$

where v does not occur in q, so that $F \equiv F^\circ[u/q]$ nor in p and has a unique occurrence in $F^\circ[u/q]$, while u does not occur in r nor in s and has a unique occurrence in $F^\circ[v/r]$. In fact, since u does not occur in r, $F^\circ[u/p, v/r]$ coincides with $(F^\circ[v/r])[u/p]$ and $F^\circ[u/q, v/r]$ coincides with $(F^\circ[v/r])[u/q]$, the above is indeed a correct representation of a derivation in $cf.EQ_1^1$. Since v does not occur in p, we have that $F^\circ[u/p, v/r]$ coincides with $(F^\circ[u/p])[v/r]$ and, furthermore, since it has a unique occurrence in $F^\circ[u/q]$, v has a unique occurrence in $F^\circ[u/p]$ as well. Therefore we can apply the induction hypothesis to $\mathcal{D}_0$ to obtain a derivation $\mathcal{D}_0'$ in $cf.EQ_1^1$ of $s = r, \Gamma' \Rightarrow (F^\circ[u/p])[v/s]$, namely of $s = r, \Gamma' \Rightarrow F^\circ[u/p, v/s]$. As $\mathcal{D}'$ we can then take

$$\dfrac{\begin{array}{c}\mathcal{D}_0'\\ s = r, \Gamma' \Rightarrow F^\circ[u/p, v/s]\end{array}}{p = q, s = r, \Gamma' \Rightarrow F^\circ[u/q, v/s]} =_1$$

which ends with a correct singleton $=_1$-inference with operating equality $p = q$, since, given that u does not occur in s, we have that $F^\circ[u/p, v/s]$ coincides with $(F^\circ[v/s])[u/p]$ and $F^\circ[u/q, v/s]$ coincides with $(F^\circ[v/s])[u/q]$ and, furthermore, since it has a unique occurrence in $F^\circ[v/r]$, u has a unique occurrence in $F^\circ[v/s]$ as well.

Case 2. The unique occurrence of r to be replaced by s is introduced as a subterm of the right-hand side of the operating equality of the last inference of $\mathcal{D}$. Then $\mathcal{D}$ can be represented as:

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma' \Rightarrow F^\circ[u/p]\end{array}}{p = q[v/r], \Gamma' \Rightarrow F^\circ[u/q[v/r]]} =_1$$

where u has a unique occurrence in F° and v has a unique occurrence in q and does not occur in p nor in F°, so that $F^\circ[u/q[v/r]]$ coincides with $(F^\circ[u/q])[v/r]$, and therefore $F \equiv F^\circ[u/q]$. As $\mathcal{D}'$ we can then take:

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma' \Rightarrow F^\circ[u/p]\end{array}}{p = q[v/s], \Gamma' \Rightarrow F^\circ[u/q[v/s]]} =_1}{s = r, p = q[v/r], \Gamma' \Rightarrow F^\circ[u/q[v/s]]} =_1^l$$

which ends with a correct singleton $=_2^l$-inference with operating equality $s = r$, since, given that v does not occur in p, we have that $p = q[v/s]$ coincides with

$(p = q)[v/s]$ and $p = q[v/r]$ coincides with $(p = q)[v/r]$ and there is a unique occurrence of v in $p = q$.

Case 3. The unique occurrence of r to be replaced by s is of the form $r^\circ[u/q]$ and is introduced by the last inference of $\mathcal{D}$ by the replacement of p by q in a term of the form $r^\circ[u/p]$ present in the premiss. Then $\mathcal{D}$ can be represented as:

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma' \Rightarrow F[v/r^\circ[u/p]]\end{array}}{p = q, \Gamma' \Rightarrow F[v/r^\circ[u/q]]} =_1$$

where u has a unique occurrence in r° and it does not occur in s nor in F. By induction hypothesis we have a derivation $\mathcal{D}'_0$ in $cf.EQ_1^1$ of $s = r^\circ[u/p], \Gamma' \Rightarrow F[v/s]$. As $\mathcal{D}'$ we can take:

$$\dfrac{\begin{array}{c}\mathcal{D}'_0\\ s = r^\circ[u/p], \Gamma' \Rightarrow F[v/s]\end{array}}{p = q, s = r^\circ[u/q], \Gamma' \Rightarrow F[v/s]} =_1^l$$

which ends with a correct singleton $=_1^l$-inference since, given that u does not occur in s, we have that $s = r^\circ[u/p]$ coincides with $(s = r^\circ)[u/p]$ and $s = r^\circ[u/q]$ coincides with $(s = r^\circ)[u/q]$ and there is a unique occurrence of u in $s = r^\circ$.

If $\mathcal{D}$ ends with a $=_1^l$-inference, then $\mathcal{D}$ has the form:

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ G[u/p], \Gamma' \Rightarrow F[v/r]\end{array}}{p = q, G[u/q], \Gamma' \Rightarrow F[v/r]} =_1^l$$

where u and v have a unique occurrence in G and F respectively. By induction hypothesis we have a derivation $\mathcal{D}'_0$ in $cf.EQ_1^1$ of $s = r, G[u/p], \Gamma' \Rightarrow F[v/s]$. As $\mathcal{D}'$ we can take:

$$\dfrac{\begin{array}{c}\mathcal{D}'_0\\ s = r, G[u/p], \Gamma' \Rightarrow F[v/s]\end{array}}{s = r, p = q, G[u/q], \Gamma' \Rightarrow F[v/s]} =_1^l$$

The proof of b) is entirely similar.□

THEOREM 4.1 *Cut elimination holds for EQ_1 and EQ_2.*

Proof Since $=_1^l$ and $=_2$ are equivalent over the structural rules and the Reflexivity Axiom (by using the cut rule), any derivation $\mathcal{D}$ in EQ_1 can be transformed into a derivation $\mathcal{D}'$ in EQ of the same endsequent. By the cut elimination theorem for EQ, $\mathcal{D}'$ can be transformed into a cut-free derivation $\mathcal{D}''$ in EQ. Since $=_2$ is admissible in $cf.EQ_1$, the applications of the $=_2$-rule in $\mathcal{D}''$ can be replaced by derivations in $cf.EQ_1$ of their conclusions thus obtaining the desired cut free derivation in EQ_1 of the endsequent of $\mathcal{D}$. The argument for EQ_2 is entirely analogous. □

Note Since $=_2^l$ is derivable in EQ_1, from the admissibility of the cut rule in $cf.EQ_1$ it follows that $=_2^l$ is also admissible in $cf.EQ_1$. Similarly also $=_1^l$ is admissible in $cf.EQ_2$.

DEFINITION 4.3 *For $i = 1, 2$, $LJ_i^=$ and $LK_i^=$ denote the systems obtained by adding the Reflexivity Axiom as well as $=_i$ and $=_i^l$ to LJ and LK respectively.*

As for Theorem 3.2, from Proposition 2.2 and Theorem 4.1, we have the following:

THEOREM 4.2 *Cut elimination holds for $LJ_1^=$, $LJ_2^=$, $LK_1^=$ and $LK_2^=$. Moreover every derivation in such systems can be transformed into a cut-free derivation of its endsequent, whose equality inferences are atomic precede the structural and logical ones .*

By providing appropriate counterexamples, to be discussed in the next section of the paper, Theorem 3.2 and Theorem 4.2 can be strengthened into the following:

THEOREM 4.3 *Any extension of LJ or LK obtained by adding the Reflexivity Axiom and some of the rules $=_1, =_2, =_1^l$ and $=_2^l$ is adequate for intuitionistic or classical first order logic with equality, but it satisfies cut elimination if and only if it contains (at least) either both $=_1$ and $=_2$, or both $=_1$ and $=_1^l$ or both $=_2$ and $=_2^l$.*

Proof For the extensions of LJ or LK obtained by adding the Reflexivity Axiom and one of the pairs $=_1$ and $=_2$, $=_1$ and $=_1^l$ and $=_2$ and $=_2^l$, the "if" part is established directly by Theorem 3.2 and 4.2. For an extension S that, besides one of such "good" pairs, contains at least one of the remaining two rules, we note that a derivation in S can be turned into one that uses only the two rules of the "good" pair that S contains. Then, either by Theorem 3.2 or 4.2, the latter derivation can be transformed into a cut-free derivation that obviously belongs to S. Clearly such extensions include all those that have at least three equality rules chosen among $=_1, =_2, =_1^l$ and $=_2^l$. The "only if" part is established by the counterexamples provided in the next section. □

4.3 Counterexamples to the validity of cut elimination

EQ, EQ_1 and EQ_2 are the only systems satisfying cut elimination that can be obtained by adding to the structural rules the Reflexivity Axiom and two equality rules chosen among $=_1$, $=_2$, $=_1^l$ and $=_2^l$. In fact $b = c, a = c \Rightarrow a = b$ has the following cut-free derivations:

$$\frac{a = c \Rightarrow a = c}{b = c,\ a = c \Rightarrow a = b}\ =_2 \qquad \text{and} \qquad \frac{a = b \Rightarrow a = b}{b = c,\ a = c \Rightarrow a = b}\ =_1^l$$

but it has no cut-free derivation, if a, b and c are distinct and only the use of $=_1$ and $=_2^l$ is allowed. More generally no sequent of the form $*)$ $\Gamma \Rightarrow a = b$, where the formulae in Γ are among $c = c$, $a = c$ and $b = c$, can have a cut free derivation using only $=_1$ and $=_2^l$. In fact, $*)$ is not the conclusion of a non trivial $=_1$-inference, since c occurs in the right-hand side of all the possible operating equalities, so that it would occur in the succedent of the conclusion of any such inference. If it is the conclusion of a $=_2^l$-inference, with operating equality $a = c$, the output formula must be necessarily another occurrence of $a = c$, obtained by replacing with a the first occurrence of c in the input formula $c = c$, to be found in the antecedent of the premiss. The same holds if the operating equality is $b = c$. Thus the premiss of the inference is still a sequent of the form $*)$. Obviously that is also the case if $*)$ is the conclusion of a weakening or contraction. Hence if $*)$ is the conclusion of an inference different from a cut, then also the premiss of the inference has the form $*)$. Now, assuming that a, b and c are distinct, no axiom has the form $*)$. Thus there are no derivations of height zero of sequents of that form. From that, by a straightforward induction argument, it follows that there are no cut-free derivations at all of sequents of the form $*)$. In particular, if a, b and c are distinct, $a = c, b = c \Rightarrow a = b$ has no cut-free derivation using only $=_1$ and $=_2^l$.

Similarly $c = b, c = a \Rightarrow a = b$ has cut-free derivations using $=_1$ or $=_2^l$, but it has no cut-free derivation, if a, b and c are distinct and only the use of $=_2$ and $=_1^l$ is allowed. Notice that, since, in LJ + Reflexivity Axiom, the equality rules are equivalent, the above sequents are derivable in the systems obtained by adding to LJ or LK the Reflexivity Axiom and only one of the rules $=_1, =_2, =_1^l$ or $=_2^l$, but they cannot have a cut free derivation in any of such systems.

Finally $a = b \Rightarrow f(a) = f(b)$ has the cut-free derivations:

$$\frac{\Rightarrow f(a) = f(a)}{a = b \Rightarrow f(a) = f(b)} =_1 \qquad \text{and} \qquad \frac{\Rightarrow f(b) = f(b)}{a = b \Rightarrow f(a) = f(b)} =_2$$

but, if a and b are distinct, it has no cut-free derivation using only $=_1^l$ and $=_2^l$. In fact, if a and b are distinct, no sequent of the form $\Gamma \Rightarrow f(a) = f(b)$, where the formulae in Γ are among $a = a$, $b = b$, $a = b$ and $b = a$, can have a cut free derivation using only $=_1^l$ and $=_2^l$.

By the above discussion also the "only if part" of Theorem 4.3 is established.

If we add to the calculus the *left symmetry rule*, that leads from $r = s, \Gamma \Rightarrow \Delta$, to $s = r, \Gamma \Rightarrow \Delta$, then the rule $=_1^l$ is immediately derivable from $=_2^l$ and conversely. As a consequence, in the presence of the left symmetry rule, cut elimination holds also for the pairs $=_1, =_2^l$ and $=_2, =_1^l$, which entails that the left symmetry rule is not admissible in the cut-free systems with only $=_1$ and $=_2^l$ or $=_2$ and $=_1^l$. On the other hand adding the left symmetry rule to the pair $=_1^l, =_2^l$ does not result into a system for which cut elimination holds, since it

has the following cut-free derivation from $=^l_1$, $=^l_2$ and the contraction rule:

$$\dfrac{\dfrac{\dfrac{r = s,\ \Gamma \Rightarrow \Delta}{s = r, r = r,\ \Gamma \Rightarrow \Delta}\ =^l_1}{s = r, s = r, s = r,\ \Gamma \Rightarrow \Delta}\ =^l_2}{s = r, \Gamma \Rightarrow \Delta}$$

Since the left symmetry rule is derivable in EQ_1, EQ_2 and EQ (by means of the cut rule) as shown by the derivations:

$$\dfrac{\dfrac{\Rightarrow s = s}{s = r \Rightarrow r = s}\ =_1 \qquad r = s,\ \Gamma \Rightarrow \Delta}{s = r,\ \Gamma \Rightarrow \Delta}$$

$$\dfrac{\dfrac{\Rightarrow r = r}{s = r \Rightarrow r = s}\ =_2 \qquad r = s,\ \Gamma \Rightarrow \Delta}{s = r,\ \Gamma \Rightarrow \Delta}$$

it is admissible in the cut-free part of any of these systems (a fact that can also be easily proved directly by induction on the height of derivations). On the other hand it is not derivable in any of $cf.EQ$, $cf.EQ_1$ and $cf.EQ_2$. For $cf.EQ$ that is obvious since $=_1$ and $=_2$ add formulae in the antecedent and modify only the formula in the succedent of a sequent. As for $cf.EQ_1$ ($cf.EQ_2$) it suffices to note that all the sequents in a derivation that starts with a sequent containing $a = b$ in the antecedent, must contain an equality of the form $a = t$ ($t = b$) in the antecedent. As a consequence, for example, there cannot be any derivation in $cf.EQ_1$ or $cf.EQ_2$ of $b = a \Rightarrow c = d$ from $a = b \Rightarrow c = d$, with a, b, c and d distinct.

5 The semishortening property

DEFINITION 5.1 *Let $LJ^=_{12}$, $LK^=_{12}$ and EQ_{12} be obtained by adding to $LJ^=$, $LK^=$ and EQ respectively, the rules $=^l_1$ and $=^l_2$.*

By Theorem 4.3, cut elimination holds for both $LJ^=_{12}$ and $LK^=_{12}$. On the ground of the contraction rule, $LK^=_{12}$ is equivalent to the system G^e in [7], which generalizes the rules $=_1$ and $=^l_1$ by permitting the substitution of r by s in more than one formula and merges them into a single rule of the form:

$$\dfrac{\Gamma[v/r] \Rightarrow \Delta[v/r]}{r = s, \Gamma[v/s] \Rightarrow \Delta[v/r]}$$

and, similarly, generalizes and merges the rules $=_2$ and $=^l_2$ into:

$$\frac{\Gamma[v/r] \Rightarrow \Delta[v/r]}{s = r, \Gamma[v/s] \Rightarrow \Delta[v/s]}$$

Thus, as an immediate consequence of Theorem 4.3, cut elimination holds for G^e. Actually [7] deals only with cut-free derivations in G^e and shows that they can be transformed into cut-free derivations in which the equality rules are applied before all the other rules and such that above the conclusion C of an equality inference there are no terms *longer* than those occurring in C, under various notions of *length* of a term. Clearly a cut-free derivation in EQ_{12} may contain terms longer than those occurring in its endsequent only if it contains some equality inference that is *lengthening* in the sense that the term r in its premiss is longer that the term s by which it is replaced in its conclusion. If we let $s \prec r$ to mean that r is longer that s, the result in [7] applies to all the binary relations $\prec$ on terms that are strict partial orders congruent with respect to substitution, namely $r \prec s$ entails $t[v/r] \prec t[v/s]$, for any term r, s and t. [8] states that it suffices to require that $\prec$ be antisymmetric. We will base our definitions on such a weaker requirement and prove a stronger result, namely that any derivation in EQ_{12} can be transformed into one whose equality inferences are all non lengthening, and those of the form $=_1$ and $=_1^l$, or, alternatively, those of the form $=_2$ and $=_2^l$, are actually shortening, namely satisfy the stronger condition $r \prec s$. It will suffice to deal with the former case, since the latter is completely symmetric. In the following $\prec$ will be a fixed, but arbitrary binary antisymmetric relation on terms, namely for any term r and s, $r \prec s$ entails $s \not\prec r$.

DEFINITION 5.2 *An application of an $=_1$-inference or of an $=_1^l$-inference with operating equality $r = s$ (or of an application of an $=_2$-inference or of an $=_2^l$-inference with operating equality $s = r$) is said* nonlengthening *if $s \not\prec r$ and* shortening *if $r \prec s$. A derivation is said to be nonlengthening if all its equality inferences are nonlengthening and* semishortening *if it is nonlengthening and, furthermore, all its $=_1$ and $=_1^l$-inferences are also shortening.*

PROPOSITION 5.1 *The equality rules $=_1$ and $=_2$ are admissible in $cf.EQ_{12}$ restricted to semishortening derivations. More precisely, there are two effective operations $\mathcal{G}_1$ and $\mathcal{G}_2$ such that:*

a) if $\mathcal{D}$ is a semishortening derivation in $cf.EQ_{12}$ of $\Gamma \Rightarrow F[v/r]$, then for any term s, $\mathcal{G}_1(\mathcal{D}, r, s)$ is a semishortening derivation in $cf.EQ_{12}$ of $r = s, \Gamma \Rightarrow F[v/s]$ and

b) if $\mathcal{D}$ is a semishortening derivation in $cf.EQ_{12}$ of $\Gamma \Rightarrow F[v/r]$, then for any term s, $\mathcal{G}_2(\mathcal{D}, r, s)$ is a semishortening derivation in $cf.EQ_{12}$ of $s = r, \Gamma \Rightarrow F[v/s]$.

The same holds if all the involved equality rules are required to be atomic.

Proof To be more accurate, $\mathcal{G}_1$ and $\mathcal{G}_2$ actually have four arguments, i.e. $\mathcal{D}$, F, $[v/r]$ and s, and their definition requires that $F[v/r]$ coincides with the succedent of the endsequent of $\mathcal{D}$. However, since it will be clear from the context what F and $[v/r]$ are, there is no harm in using the simplified notations $\mathcal{G}_1(\mathcal{D}, r, s)$ and $\mathcal{G}_2(\mathcal{D}, r, s)$. Since the derivation of an equality rule from its singleton form, discussed in connection with Lemma 4.1, uses (repeatedly) the same operating equality, it is obvious that if an equality rule is semishortening, so it is the derivation of its conclusion from its premiss by means of the corresponding singleton equality rule. As a consequence, as in the proof of Proposition 4.2, it suffices to prove that there are effective operations $\mathcal{G}_1$ and $\mathcal{G}_2$ for which a) and b) holds for $cf.EQ^1_{12}$, under the assumption that v has exactly one occurrence in F.

If $r \prec s$ then $\mathcal{G}_1(\mathcal{D}, r, s)$ is obtained by applying to $\mathcal{D}$ an $=_1$-inference with operating equality $r = s$ and if $s \not\prec r$ (in particular if $r \prec s$), $\mathcal{G}_2(\mathcal{D}, r, s)$ is obtained by applying to $\mathcal{D}$ an $=_2$-inference, with operating equality $s = r$. Hence in defining $\mathcal{G}_1$ we may assume that $r \not\prec s$, while in defining $\mathcal{G}_2$ we may assume that $s \prec r$.

$\mathcal{G}_1(\mathcal{D}, r, s)$ and $\mathcal{G}_2(\mathcal{D}, r, s)$ are defined simultaneously by recursion on the height $h(\mathcal{D})$ of $\mathcal{D}$, for arbitrary s. We will deal only with the definition of $\mathcal{G}_1$. The definition of $\mathcal{G}_2$ is similar.

If $h(\mathcal{D}) = 0$ we have the following cases.

Case 0.1 $\mathcal{D}$ reduces to $F[v/r] \Rightarrow F[v/r]$. As $\mathcal{G}_1(\mathcal{D}, r, s)$ we can take

$$\frac{F[v/s] \Rightarrow F[v/s]}{r = s, F[v/r] \Rightarrow F[v/s]} =^l_2$$

which belongs to $cf.EQ^1_{12}$ and is nonlengthening, since we are assuming that $r \not\prec s$.

Case 0.2 $\mathcal{D}$ reduces to $\quad \Rightarrow t_0 = t[v/r]$ with $t_0 \equiv t[v/r]$. As $\mathcal{G}_1(\mathcal{D}, r, s)$ we can take:

$$\frac{\Rightarrow t[v/s] = t[v/s]}{r = s \Rightarrow t[v/r] = t[v/s]} =_2$$

which belongs to $cf.EQ^1_{12}$ and is nonlengthening. Similarly if $\mathcal{D}$ reduces to $\Rightarrow t[v/r] = t_0$ with $t_0 \equiv t[v/r]$.

If $h(\mathcal{D}) > 0$ and $\mathcal{D}$ ends with a structural rule and has the form:

$$\frac{\begin{array}{c}\mathcal{D}_0 \\ \Gamma' \Rightarrow F[v/r]\end{array}}{\Gamma \Rightarrow F[v/r]}$$

where $\mathcal{D}_0$ is semishortening and belongs to $cf.EQ^1_{12}$,

$\mathcal{G}_1(\mathcal{D}, r, s)$ is obtained by applying the same structural rule to the endsequent of $\mathcal{G}_1(\mathcal{D}_0, r, s)$ that, by induction hypothesis, is a semishortening derivation in $cf.EQ^1_{12}$ of $r = s, \Gamma' \Rightarrow F[v/s]$.

Otherwise we have the following four cases depending on the ending equality inference of $\mathcal{D}$.

Case 1. $\mathcal{D}$ ends with an $=_1$-inference. Then we have three subcases. We will omit a detailed verification that the derivations we are going to display are indeed correct, since it is entirely similar to that carried through in the proof of Proposition 4.2, but point out that they belong to $cf.EQ^1_{12}$.

Case 1.1. $\mathcal{D}$ can be represented as :

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma' \Rightarrow F^\circ[u/p, v/r]\end{array}}{p = q, \Gamma' \Rightarrow F^\circ[u/q, v/r]} =_1$$

where $\mathcal{D}_0$ is semishortening and belongs to $cf.EQ^1_{12}$, u and v have exactly one occurrence in F°, v does not occur in q and $p \prec q$. By induction hypothesis $\mathcal{G}_1(\mathcal{D}_0, r, s)$ is a semishortening derivation in $cf.EQ^1_{12}$ of $r = s, \Gamma' \Rightarrow F^\circ[u/p, v/s]$. Then we can let $\mathcal{G}_1(\mathcal{D}, r, s)$ be:

$$\dfrac{\begin{array}{c}\mathcal{G}_1(\mathcal{D}_0, r, s)\\ r = s, \Gamma' \Rightarrow F^\circ[u/p, v/s]\end{array}}{p = q, r = s, \Gamma' \Rightarrow F^\circ[u/q, v/s]} =_1$$

which is semishortening, since $p \prec q$ and belongs to $cfEQ^1_{12}$.

Case 1.2 $\mathcal{D}$ can be represented as:

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma' \Rightarrow F^\circ[u/p]\end{array}}{p = q[v/r], \Gamma' \Rightarrow F^\circ[u/q[v/r]]} =_1$$

where $\mathcal{D}_0$ is semishortening and belongs to $cf.EQ^1_{12}$, v occurs in q but not in p and u and v have exactly one occurrence in F° and $F^\circ[u/q]$ respectively. By induction hypothesis there is a semishorthening derivation $\mathcal{G}_1(\mathcal{D}_0, p, q[v/s])$ in $cf.EQ^1_{12}$ of $p = q[v/s], \Gamma' \Rightarrow F^\circ[u/q[v/s]]$ and we can let $\mathcal{G}_1(\mathcal{D}, r, s)$ be:

$$\dfrac{\begin{array}{c}\mathcal{G}_1(\mathcal{D}_0, p, q[v/s])\\ p = q[v/s], \Gamma' \Rightarrow F^\circ[u/q[v/s]]\end{array}}{r = s, p = q[v/r], \Gamma' \Rightarrow F^\circ[u/q[v/s]]} =^l_2$$

which is semishortening, given that in defining $\mathcal{G}_1$ we are assuming that $r \not\prec s$, and belongs to $cf.EQ^1_{12}$.

Case 1.3 r has the form $r^\circ[u/q]$ and $\mathcal{D}$ can be represented as:

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma' \Rightarrow F[v/r^\circ[u/p]]\end{array}}{p = q, \Gamma' \Rightarrow F[v/r^\circ[u/q]]} =_1$$

where $\mathcal{D}_0$ is semishortening and belongs to $cf.EQ^1_{12}$, $p \prec q$, u has exactly one occurrence in r°, but it occurs neither in F nor in s, and v has exactly one occurrence in F.

By induction hypothesis there is a semishortening derivation $\mathcal{G}_1(\mathcal{D}_0, r^\circ[u/p], s)$ of $r^\circ[u/p] = s, \Gamma' \Rightarrow F[v/s]$ and we can let $\mathcal{G}_1(\mathcal{D}, r, s)$ be:

$$\frac{\begin{array}{c}\mathcal{G}_1(\mathcal{D}_0, r^\circ[u/p], s) \\ r^\circ[u/p] = s, \Gamma' \Rightarrow F[v/s]\end{array}}{p = q, r^\circ[u/q] = s, \Gamma' \Rightarrow F[v/s]} =^l_1$$

which is semishortening since $p \prec q$ and belongs to $cf.EQ^1_{12}$.

Case 2 $\mathcal{D}$ ends with an $=_2$ inference.

Case 2.1 $\mathcal{D}$ can be represented as:

$$\frac{\begin{array}{c}\mathcal{D}_0 \\ \Gamma' \Rightarrow F^\circ[u/p, v/r]\end{array}}{q = p, \Gamma' \Rightarrow F^\circ[u/q, v/r]} =_2$$

where $\mathcal{D}_0$ is semishortening and belongs to $cf.EQ^1_{12}$, u and v have exactly one occurrence in F°, v does not occur in q and $q \not\prec p$. Then to define $\mathcal{G}_1(\mathcal{D}, r, s)$ it suffices to replace in Case 1.1, $p = q$ by $q = p$ and replace the last $=_1$-inference of $\mathcal{G}_1(\mathcal{D}, r, s)$ by the $=_2$-inference having operating equality $q = p$, rather than $p = q$. Since $q \not\prec p$, the derivation so obtained is semishorthening. Furthermore it belongs to $cf.EQ^1_{12}$.

Case 2.2 $\mathcal{D}$ can be represented as:

$$\frac{\begin{array}{c}\mathcal{D}_0 \\ \Gamma' \Rightarrow F^\circ[u/p]\end{array}}{q[v/r] = p, \Gamma' \Rightarrow F^\circ[u/q[v/r]]} =_2$$

where $\mathcal{D}_0$ is semishortening and belongs to $cf.EQ^1_{12}$, v occurs in q but not in p and u and v have exactly one occurrence in F° and $F^\circ[u/q]$ respectively. Then to define $\mathcal{G}_1(\mathcal{D}, r, s)$ it suffices to replace in Case 1.2 the equalities $p = q[v/r]$ and $p = q[v/s]$ by $q[v/r] = p$ and $q[v/s] = p$ respectively and use $\mathcal{G}_2(\mathcal{D}_0, p, q[v/s])$ instead of $\mathcal{G}_1(\mathcal{D}_0, p, q[v/s])$ in the induction hypothesis. Notice that in this case the definition of $\mathcal{G}_1(\mathcal{D}, r, s)$ depends on $\mathcal{G}_2(\mathcal{D}_0, p, q[v/s])$.

Case 2.3 r has the form $r^\circ[u/q]$ and $\mathcal{D}$ can be represented as:

$$\frac{\begin{array}{c}\mathcal{D}_0 \\ \Gamma' \Rightarrow F[v/r^\circ[u/p]]\end{array}}{q = p, \Gamma' \Rightarrow F[v/r^\circ[u/q]]} =_2$$

where $\mathcal{D}_0$ is semishortening and belongs to $cf.EQ^1_{12}$, $q \not\prec p$, u has exactly one occurrence in r°, but it occurs neither in F nor in s, and v has exactly one occurrence in F. Then $\mathcal{G}_1(\mathcal{D}, r, s)$ is defined as in Case 1.3 except that $p = q$

is replaced by $q = p$ and the $=_1$ and $=_1^l$-inferences are replaced by $=_2$ and $=_2^l$ inferences with operating equality $q = p$ rather than $p = q$. Since $q \not\prec p$, the derivation that is obtained is semishortening . Furthermore, it belongs to $cf.EQ_{12}^1$.

Case 1^l $\mathcal{D}$ ends with an $=_1^l$-inference. Then $\mathcal{D}$ can be represented as:

$$\frac{\begin{array}{c}\mathcal{D}_0\\ G[u/p], \Gamma' \Rightarrow F[v/r]\end{array}}{p = q, G[u/q], \Gamma' \Rightarrow F[v/r]} =_1^l$$

where $\mathcal{D}_0$ is semishortening and belongs to $cf.EQ_{12}^1$, $p \prec q$ and u and v have exactly one occurrence in G and F respectively. By induction hypothesis $\mathcal{G}_1(\mathcal{D}_0, r, s)$ is a semishortening derivation in $cf.EQ_{12}^1$ of $r = s, G[u/p], \Gamma' \Rightarrow F[v/s]$ and we can let $\mathcal{G}_1(\mathcal{D}, r, s)$ be:

$$\frac{\begin{array}{c}\mathcal{G}_1(\mathcal{D}_0, r, s)\\ r = s, G[u/p], \Gamma' \Rightarrow F[v/s]\end{array}}{p = q, r = s, G[u/q], \Gamma' \Rightarrow F[v/s]} =_1^l$$

which is semishortening, since $p \prec q$, and belongs to $cf.EQ_{12}^1$.

Case 2^l $\mathcal{D}$ ends with a $=_2^l$-inference, namely $p = q$ is replaced by $q = p$ in the endsequent of $\mathcal{D}$ as represented in Case 1^l and $q \not\prec p$. Then $\mathcal{G}_1(\mathcal{D}, r, s)$ is defined as in Case 1^l, except that $p = q$ is replaced by $q = p$ and the $=_1^l$ inferences are replaced by $=_2^l$-inferences with operating equality $q = p$ rather than $p = q$. Since $q \not\prec p$ the derivation that is obtained is semishortening. Furthermore it belongs to $cf.EQ_{12}^1$.

Clearly the argument goes through without any change under the assumption that F and G are atomic, thus establishing the last part of the claim. □

THEOREM 5.1 *Any derivation in EQ_{12} can be transformed into a cut-free semishortening derivation in EQ_{12} of its endsequent. The same holds for EQ_{12} with the equality rules restricted to be atomic.*

Proof As pointed out at the beginning of Sec 4.1, the rules $=_1^l$ and $=_2^l$ are derivable in EQ, hence every derivation in EQ_{12} can be effectively transformed into a derivation in EQ, henceforth, by Theorem 3.1, into a cut free derivation in EQ of its endsequent. The conclusion follows by the admissibility of the equality rules $=_1$ and $=_2$ in $cf.EQ_{12}$ restricted to semishortening derivations, established in Proposition 5.1. The claim for EQ_{12} with atomic equality rules follows in the same way by the claims in Theorem 3.1 and Proposition 5.1. concerning atomic equality rules. □

As for Theorem 3.2 and 4.2, from Proposition 2.2 and Theorem 5.1 we have the following final result:

THEOREM 5.2 *Any derivation in $LJ^{=}_{12}$ or $LK^{=}_{12}$ can be transformed into a cut-free semishortening derivation of its endsequent in the same calculus, whose equality inferences are atomic and precede the structural and logical ones.*

Proof By Proposition 2.2, every derivation in $LJ^{=}_{12}$ or $LK^{=}_{12}$ of a sequent $\Gamma \Rightarrow \Delta$ can be transformed into a derivation of $\Gamma \Rightarrow \Delta$, that consists of subderivations in EQ_{12} with atomic equality inferences followed only by structural and logical inferences. Then, by Theorem 5.1, such subderivations can be transformed into derivations of their endsequents, in $cf.EQ_{12}$ with atomic semishortening equality inferences. Finally the latter can be permuted upward with respect to the structural inferences, as in the proof of Theorem 3.2, without affecting their semishortness. The result is the desired derivation of $\Gamma \Rightarrow \Delta$ in $LJ^{=}_{12}$ or $LK^{=}_{12}$ in which all the equality inferences are atomic as well as semishortening and precede the structural and logical ones. □

By Proposition 4.1, $=^{l}_{1}$ and $=^{l}_{2}$ are admissible in $cf.EQ$, hence, as $=_1$ and $=_2$ they are admissibile also in $cf.EQ_{12}$ restricted to semishortening derivations. As it results from [7], in the case of nonlengthening derivations, a direct inductive proof of this admissibility result is possible, but it requires the additional assumption that $\prec$ be a strict partial order congruent with respect to substitution. It is the admissibility of $=^{l}_{1}$ and $=^{l}_{2}$ in $cf.EQ$, unnoticed in [7], that allows for the weakening of such assumption to the requirement that $\prec$ be simply antisymmetric.

6 Acknowledgment

We are very grateful to the anonimous referees for helpful comments, in particular for the suggestion of using Takeuti's calculus LK'_e in the proof of Proposition 2.2 and strengthening a somewhat weaker form of Theorem 5.2.

References

[1] M. Baaz, A. Leitsch, Methods of Cut-Elimination Trends in Logic 34 Springer (2011)

[2] H.B. Curry, Foundations of Mathematical Logic Dover (1977)

[3] A.Degtyarev, A.Voronkov, Equality Reasoning in Sequent-Based Calculi in: A.Robinson, A.Voronkov (eds) Handbook of Automated Reasoning Vol 1 pp 613-706 Elsevier (2001)

[4] G. Gentzen, Untersuchungen uber der logische Schliessen *Matematische Zeitschrift* Vol.39, pp. 176-210, 405-431 (1935)

[5] J. Gallier, Logic for Computer Science, Foundations of Automatic Theorem Proving, Harper & Row, NewYork (1986).

[6] S. Kanger, A Simplified Proof Method for Elementary Logic. In: P. Braffort, D. Hirshberg (eds) Computer Programming and Formal Systems, pp. 87-94. North-Holland, Amsterdam (1963)

[7] A.V. Lifschitz, Specialization of the form of deduction in the predicate calculus with equality and function symbols (in Russian). In Trudy MIAN, vol. 98, 5-25 (1968). English translation in: V.P. Orevkov (ed) *The Calculi of Symbolic Logic. I*, Proceedings of the Steklov Institute of Mathematics 98 (1971)

[8] V. P. Orevkov, On Nonlengthening Applications of Equality Rules (in Russian) Zapiski Nauchnyh Seminarov LOMI, 16:152-156, 1969 English translation in: A.O. Slisenko (ed) *Studies in Constructive Logic*, Seminars in Mathematics: Steklov Math. Inst. 16, Consultants Bureau, NY-London 77-79 (1971)

[9] J. von Plato, S. Negri, Cut Elimination in the Presence of Axioms. The Bulletin of Symbolic Logic 4 (4) , 418–435 (1998)

[10] J. von Plato, S. Negri, Structural Proof Theory. Cambridge University Press (2001)

[11] D. Prawitz, Natural Deduction. A Proof-Theoretical Study Almquist and Wiksell (1965)

[12] F. Parlamento, F. Previale, Cut elimination for Gentzen's Sequent Calculus with Equality and Logic of Partial Terms. Lecture Notes in Computer Science 7750, 161-172 (2013)

[13] G. Takeuti, Proof Theory. Studies in Logic and the Foundations of Mathematics vol 81, 2nd edition North Holland, Amsterdam (1987)

[14] A.S. Troelstra, H. Schwichtenberg, Basic Proof Theory. Cambridge Tracts in Theoretical Computer Science, Vol. 43 Cambridge University Press, Cambridge(1996).

[15] A.S. Troelstra, H. Schwichtenberg, Basic Proof Theory. 2nd edition Cambridge University Press, Cambridge(2000).

[16] H. Wang, Towards mechanical mathematics IBM journal of Research and Development vol 4, pp. 2-22 (1960)

7 Appendix 1

7.1 Equivalence between $\forall X(X(b) \leftrightarrow X(a))$ and $\forall X(X(b) \rightarrow X(a))$

$\forall X(X(b) \rightarrow X(a))$ can be deduced from $\forall X(X(a) \rightarrow X(b))$ by instantiating the bound predicate variable X by the lambda term $\lambda v(Z(v) \rightarrow Z(a))$, where Z is a free predicate variable, so as to obtain $(Z(a) \rightarrow Z(a)) \rightarrow (Z(b) \rightarrow Z(a))$. Then, given the deducibilty, by $\rightarrow$-introduction, of $Z(a) \rightarrow Z(a)$, an $\rightarrow$-elimination followed by a $\forall$-introduction yields $\forall X(X(b) \rightarrow X(a))$ as desired.

7.2 Derivation of $=^{(2)}\Rightarrow$ and $\Rightarrow=^{(2)}$ from the definition of $=$

Having defined $r = s$ as $\forall X(X(r) \rightarrow X(s))$, the conclusion of the rule $=^{(2)}\Rightarrow$, namely

$$\frac{\Lambda \Rightarrow F[v/r] \qquad F[v/s], \Gamma \Rightarrow \Delta}{r = s, \Lambda, \Gamma, \Rightarrow \Delta}$$

can be derived from its premisses by applying first the left introduction rule for $\rightarrow$ and then the second order left introduction rule for $\forall$, while the conclusion of $\Rightarrow=^{(2)}$, namely

$$\frac{Z(r), \Gamma \Rightarrow Z(s)}{\Gamma \Rightarrow r = s}$$

can be derived from its premiss by applying first the right introduction rule for $\rightarrow$ and then the second order right introduction rule for $\forall$.

7.3 Derivation of the definition of $=$ from $=^{(2)}\Rightarrow$ and $\Rightarrow=^{(2)}$

The sequents $r = s \Rightarrow \forall X(X(r) \rightarrow X(s))$ and $\forall X(X(r) \rightarrow X(s)) \Rightarrow r = s$ can be derived by means of $=^{(2)}\Rightarrow$ and $\Rightarrow=^{(2)}$ as follows:

$$\dfrac{\dfrac{\dfrac{Z(r) \Rightarrow Z(r) \qquad Z(s) \Rightarrow Z(s)}{r = s, Z(r) \Rightarrow Z(s)}}{r = s \Rightarrow Z(r) \rightarrow Z(s)}}{r = s \Rightarrow \forall X(X(r) \rightarrow X(s))} =^{(2)}\Rightarrow \qquad \dfrac{\dfrac{\dfrac{Z(r) \Rightarrow Z(r) \qquad Z(s) \Rightarrow Z(s)}{Z(r) \rightarrow Z(s), Z(r) \Rightarrow Z(s)}}{\forall X(X(r) \rightarrow X(s)), Z(r) \Rightarrow Z(s)}}{\forall X(X(r) \rightarrow X(s)) \Rightarrow r = s} \Rightarrow=^{(2)}$$

7.4 Equivalence between $\Rightarrow=^{(2)}$ and $\Rightarrow r = r$

The sequent $\Rightarrow r = r$ is immediately derived by $\Rightarrow=^{(2)}$ applied to the logical axiom $Z(r) \Rightarrow Z(r)$ and, conversely, $\Rightarrow=^{(2)}$, can be derived from $\Rightarrow r = r$, by using the cut rule, as follows:

$$\dfrac{\dfrac{\dfrac{Z(r),\Gamma \Rightarrow Z(s)}{\Gamma \Rightarrow Z(r) \to Z(s)}}{\Gamma \Rightarrow \forall X(X(r) \to X(s))} \qquad \dfrac{\dfrac{\Rightarrow r = r \qquad r = s \Rightarrow r = s}{r = r \to r = s \Rightarrow r = s}}{\forall X(X(r) \to X(s)) \Rightarrow r = s}}{\Gamma \Rightarrow r = s}$$

7.5 Equivalence between $LJ^{=}$ and $LJ^{(1)=}$

Derivations of $=_1$ and $=_2$:

$$\dfrac{\Gamma \Rightarrow F[v/r] \qquad F[v/s] \Rightarrow F[v/s]}{r = s, \Gamma \Rightarrow F[v/s]} \Longrightarrow$$

$$\dfrac{\Rightarrow s = s \qquad \dfrac{\Gamma \Rightarrow F[v/r] \qquad F[v/s] \Rightarrow F[v/s]}{r = s, \Gamma \Rightarrow F[v/s]} \Longrightarrow}{s = r, \Gamma \Rightarrow F[v/s]} \Longrightarrow$$

where the last inference is a correct application of $\Longrightarrow$ in which the place of F is taken by $v = s$.

Derivation of $\Longrightarrow$ from $=_1$:

$$\dfrac{\dfrac{\Lambda \Rightarrow F[v/r]}{r = s, \Lambda \Rightarrow F[v/s]} =_1 \qquad F[v/s], \Gamma \Rightarrow \Delta}{r = s, \Lambda, \Gamma \Rightarrow \Delta}$$

Derivation of $\Longrightarrow$ from $\Rightarrow=$ and $=_2$:

$$\dfrac{\dfrac{\dfrac{\Rightarrow s = s}{r = s \Rightarrow s = r} =_2 \qquad \dfrac{\Lambda \Rightarrow F[v/r]}{s = r, \Lambda \Rightarrow F[v/s]} =_2}{r = s, \Lambda \Rightarrow F[v/s]} \qquad F[v/s], \Gamma \Rightarrow \Delta}{r = s, \Lambda, \Gamma \Rightarrow \Delta}$$

Therefore it suffices to add $\Rightarrow=$ and $=_1$ or $\Rightarrow=$ and $=_2$ to LJ in order to have a system equivalent to $LJ^{(1)=}$.

7.6 Equivalence between $LJ^{(1)=}$ and $LJ^{=}$ with the axioms $\Rightarrow t = t$ and $r = s, F[v/r] \Rightarrow F[v/s]$

$$\dfrac{F[v/r] \Rightarrow F[v/r] \qquad F[v/s] \Rightarrow F[v/s]}{r = s, F[v/r] \Rightarrow F[v/s]} \Longrightarrow$$

$$\dfrac{\dfrac{\Gamma \Rightarrow F[v/r] \qquad r = s, F[v/r] \Rightarrow F[v/s]}{\Gamma, r = s \Rightarrow F[v/s]} \qquad F[v/s], \Lambda \Rightarrow \Delta}{r = s, \Gamma, \Lambda \Rightarrow \Delta}$$

7.7 Derivations of $=_1^l$ and $=_2^l$ in $LJ^{(1)=}$

$$\frac{F[v/s] \Rightarrow F[v/s] \quad F[v/r], \Gamma \Rightarrow \Delta}{s = r, F[v/s], \Gamma \Rightarrow \Delta} \Longrightarrow$$

$$\frac{\Rightarrow r = r \qquad \dfrac{F[v/s] \Rightarrow F[v/s] \quad F[v/r], \Gamma \Rightarrow \Delta}{s = r, F[v/s], \Gamma \Rightarrow \Delta} \Longrightarrow}{r = s, F[v/s], \Gamma \Rightarrow \Delta} \Longrightarrow$$

7.8 Derivation of $\Longrightarrow$ from $=_2^l$ and from $\Rightarrow=$ and $=_1^l$

$$\frac{\Gamma \Rightarrow F[v/r] \qquad \dfrac{\Lambda, F[v/s] \Rightarrow \Delta}{r = s, F[v/r], \Lambda \Rightarrow \Delta} =_2^l}{r = s, \Gamma, \Lambda \Rightarrow \Delta}$$

$$\frac{\Gamma \Rightarrow F[v/r] \qquad \dfrac{\Rightarrow s = s \qquad \dfrac{\dfrac{\dfrac{F[v/s], \Lambda \Rightarrow \Delta}{s = r, F[v/r], \Lambda \Rightarrow \Delta} =_1^l}{s = s, r = s, F[v/r], \Lambda \Rightarrow \Delta} =_1^l}{r = s, F[v/r], \Lambda \Rightarrow \Delta}}{r = s, \Gamma, \Lambda \Rightarrow \Delta}$$

By 7.7 and 7.8, it suffices to add $\Rightarrow=$ and $=_1^l$ or $\Rightarrow=$ and $=_2^l$ to LJ in order to have a system equivalent to $LJ^{(1)=}$.

7.9 Equivalence between CNG and $\Longrightarrow$

$$\frac{\Gamma \Rightarrow r = s \qquad \dfrac{\Lambda \Rightarrow F[v/r] \quad F[v/s] \Rightarrow F[v/s]}{r = s, \Lambda \Rightarrow F[v/s]} \Longrightarrow}{\Gamma, \Lambda \Rightarrow F[v/s]}$$

$$\frac{\dfrac{r = s \Rightarrow r = s \quad \Gamma \Rightarrow F[v/r]}{r = s, \Gamma \Rightarrow F[v/s]}\ \mathrm{CNG} \qquad F[v/s], \Lambda \Rightarrow \Delta}{r = s, \Gamma, \Lambda \Rightarrow \Delta}$$

Therefore $=_1$, $=_2$, CNG, $=_1^l$ and $=_2^l$ are all equivalent to $\Longrightarrow$, and therefore to each other, over $\Rightarrow=$ and the structural rules.

8 Appendix 2

In order to prove that every derivation in $LJ^=$ or $LK^=$ can be transformed into a separated derivation of its endsequent, we note first that, thanks to the cut rule, the equality rules can be derived from their special case in which the formula that they transform is atomic, and then that the cut rule is admissible over its restriction to atomic cut formulae.

We leave to the reader the proof by induction on the degree of F of the following fact:

LEMMA 8.1 *The sequents of the following form:*

a) $r = s, F[v/r] \Rightarrow F[v/s]$

b) $s = r, F[v/r] \Rightarrow F[v/s]$

have derivations whose equality inferences are atomic.

PROPOSITION 8.1 *Any non atomic equality inference in a given derivation in $LJ^=$ or $LK^=$ can be replaced by a cut between its premiss and the endsequent of a derivation that uses only atomic equality inferences. In particular any derivable sequent in $LJ^=$ or $LK^=$ has a derivation whose equality inferences are all atomic.*

Proof A non atomic $=_1$-inference of the form:

$$\frac{\Gamma \Rightarrow \Delta, F[v/r]}{r = s, \Gamma \Rightarrow \Delta, F[v/s]}$$

can be replaced by:

$$\frac{\Gamma \Rightarrow \Delta, F[v/r] \qquad \begin{array}{c}\mathcal{D}\\ r = s, F[v/r] \Rightarrow F[v/s]\end{array}}{r = s, \Gamma \Rightarrow \Delta, F[v/s]}$$

where $\mathcal{D}$ is the derivation containing only atomic equality-inferences of Lemma 8.1 *a*) for F. A non atomic $=_2$-inference is eliminated in a similar way using Lemma 8.1 *b*). □

PROPOSITION 8.2 *If $\Gamma \Rightarrow \Delta \sharp F$ and $\Lambda \sharp F \Rightarrow \Theta$ have derivations in $LJ^=$ or $LK^=$ whose equality and cut-inferences are atomic, then also $\Gamma, \Lambda \Rightarrow \Delta, \Theta$ has a derivation in the same system whose equality and cut-inferences are atomic.*

Proof Let $\mathcal{D}$ and $\mathcal{E}$ be derivations of $\Gamma \Rightarrow \Delta \sharp F$ and $\Lambda \sharp F \Rightarrow \Theta$ whose equality and cut-inferences are atomic. If $\Delta \sharp F$ coincides with Δ or $\Lambda \sharp F$ coincides with Λ, then the desired derivation of $\Gamma, \Lambda \Rightarrow \Delta, \Theta$ can be simply obtained by applying

some weakenings to the endsequent $\Gamma \Rightarrow \Delta$ of $\mathcal{D}$ or to the endsequent $\Lambda \Rightarrow \Theta$ of $\mathcal{E}$. We can therefore assume that in $\Delta \sharp F$ there are occurrences of F that are not listed in Δ and similarly for $\Lambda \sharp F$. If F occurs in Δ, then from $\Gamma \Rightarrow \Delta \sharp F$ we can derive $\Gamma \Rightarrow \Delta$ by means of contractions, and from $\Gamma \Rightarrow \Delta$ we can then derive $\Gamma, \Lambda \Rightarrow \Delta, \Theta$ as in the previous case. Similarly if F occurs in Λ. We can therefore assume that F occurs in $\Delta \sharp F$ and in $\Lambda \sharp F$ but it does not occur in Δ, Λ. Furthermore we can assume that F does not occur in Γ, Θ either, for, otherwise $\Gamma, \Lambda \Rightarrow \Delta, \Theta$ can be derived by weakening $\mathcal{D}$, if F occurs in Θ, or $\mathcal{E}$, if F occurs in Γ, and then contracting the occurrences of F in $\Delta \sharp F$ with one of the occurrences of F in Θ or the occurrences of F in $\Lambda \sharp F$ with one of the occurrences of F in Γ. Finally if F is atomic it suffices to contract the occurrence in F in $\Delta \sharp F$ and $\Lambda \sharp F$ into a single one, and then apply a cut with the atomic cut formula F, in order to obtain the desired derivation.

In the remaining cases we proceed, as in Gentzen's original proof of the cut elimination theorem, by a principal induction on the degree of F and a secondary induction on the sum of the left rank $\rho_l(F, \mathcal{D})$ of F in $\mathcal{D}$ and of the right rank $\rho_r(F, \mathcal{E})$ of F in $\mathcal{E}$, defined as the largest number of consecutive sequents in a path of $\mathcal{D}$ (of $\mathcal{E}$) starting with the endsequent, that contain F in the succedent (in the antecedent).

Besides the cases considered in Gentzen's proof, there is also the possibility that $\mathcal{D}$ or $\mathcal{E}$ end with an atomic equality inference or with an atomic cut.

Case 1 $\mathcal{D}$ ends, say, with an atomic $=_1$-inference. Since F is not atomic, F is not active in such an inference and $\mathcal{D}$ can be represented as:

$$\dfrac{\begin{array}{c}\mathcal{D}_0 \\ \Gamma' \Rightarrow \Delta' \sharp F, A[v/r]\end{array}}{r = s, \Gamma' \Rightarrow \Delta' \sharp F, A[v/s]}$$

where $r = s, \Gamma'$ coincides with Γ and $\Delta', A[v/s]$ coincides with Δ. Since $\rho_l(F, \mathcal{D}_0) < \rho_l(F, \mathcal{D})$, by induction hypothesis we have a derivation whose equality and cut-inferences are atomic of $\Gamma', \Lambda \Rightarrow \Delta', A[v/r], \Theta$, from which the desired derivation of $\Gamma, \Lambda \Rightarrow \Delta, \Theta$ can be obtained by applying the same $=_1$-inference.

Case 2 $\mathcal{D}$ ends with an atomic cut. Then $\mathcal{D}$ can be represented as:

$$\dfrac{\begin{array}{cc}\mathcal{D}_0 & \mathcal{D}_1 \\ \Gamma_1 \Rightarrow \Delta_1 \sharp F, A & A, \Gamma_2 \Rightarrow \Delta_2 \sharp F\end{array}}{\Gamma_1, \Gamma_2 \Rightarrow \Delta \sharp F}$$

where Δ coincides with Δ_1, Δ_2, so that F does not occur in Δ_1 nor in Δ_2. Since $\rho_l(F, \mathcal{D}_0) < \rho_l(F, \mathcal{D})$ and $\rho_l(F, \mathcal{D}_1) < \rho_l(F, \mathcal{D})$, by induction hypothesis applied to $\mathcal{D}_0$ and $\mathcal{E}$ and to $\mathcal{D}_1$ and $\mathcal{E}$ there are derivations whose equality and cut-inferences are atomic of $\Gamma_1, \Lambda \Rightarrow \Delta_1, A, \Theta$ and $A, \Gamma_2, \Lambda \Rightarrow \Delta_2, \Theta$, to which it suffices to apply a cut with atomic cut formula A and then some contractions to have the desired derivation of $\Gamma_1, \Gamma_2, \Lambda \Rightarrow \Delta_1, \Delta_2, \Theta$.

The cases in which it is $\mathcal{E}$ to end with an atomic equality or a cut-inference are entirely analogous. □

From Proposition 8.1 and Proposition 8.2 it follows immediately the following:

PROPOSITION 8.3 *Every derivation in $LJ^{=}$ or $LK^{=}$ can be transformed into a derivation of its endsequent, whose equality and cut-inferences are atomic.*

Remark For the proof of Proposition 8.2 it is crucial that the equality rules transform atomic formulae only. For example in case $\rho_l(F, \mathcal{D}) = 1$ and $\rho_r(F, \mathcal{E}) = 1$, if F had the form $F^\circ[v/s]$, with F° non atomic, $\mathcal{D}$ ended with an equality inference transforming $F^\circ[v/r]$ into $F^\circ[v/s]$, and $\mathcal{E}$ by a logical inference introducing $F^\circ[v/s]$ in the antecedent, then there would be no way of applying the induction hypothesis.

Note Concerning the use of $\Gamma\sharp F$, we note that when $\Delta\sharp F$ and $\Lambda\sharp F$ take the form Δ, F and Λ, F, from Proposition 8.2, we obtain directly that the derivations having only atomic equality and cut-inferences are closed under the application of the cut rule. That is a slight simplification with respect to the use of Gentzen's mix rule that eliminates all the occurrences of F, so that the use of additional weakenings may be necessary to derive the conclusion of a cut-inference.

PROPOSITION 8.4 *If $\Gamma \Rightarrow \Delta\sharp A[v/r]$ has a separated derivation in $LJ^{=}$ or $LK^{=}$, then also $r = s, \Gamma \Rightarrow \Delta, A[v/s]$ and $s = r, \Gamma \Rightarrow \Delta, A[v/s]$ have separated derivations in the same system.*

Proof Let $\mathcal{D}$ be a separated derivation of $\Gamma \Rightarrow \Delta\sharp A[v/r]$. We proceed by induction on the height $h(\mathcal{D})$ of $\mathcal{D}$. In the base case $\mathcal{D}$ reduces to an axiom and it suffices to apply an $=_1$ or an $=_2$-inference to the axiom itself. If $h(\mathcal{D}) > 0$ we have the following cases.

Case 1. $\mathcal{D}$ ends with a cut or an equality-inference. In this case $\mathcal{D}$ doesn't contain any logical inference. If $\Delta = \Delta\sharp A[v/r]$, then it suffices to weaken the endsequent of $\mathcal{D}$. Otherwise we can contract all the occurrences of $A[v/r]$ not belonging to Δ into a single one and then apply an $=_1$ or $=_2$-inference.

Case 2 $\mathcal{D}$ ends with a weak structural inference. If such an inference involves one of the occurrences of $A[v/r]$ in $\Delta\sharp A[v/r]$ not belonging to Δ, then its premiss is already of the form $\Gamma \Rightarrow \Delta\sharp A[v/r]$ so that the desired derivation is provided directly by the induction hypothesis. Otherwise the latter is obtained by applying the induction hypothesis and then the same weak structural rule.

Case 3 $\mathcal{D}$ ends with a logical inference. $A[v/r]$, being atomic, cannot be the principal formula of the inference, and the conclusion is a straightforward consequence of the induction hypothesis. □

PROPOSITION 8.5 *If* $\Gamma \Rightarrow \Delta\sharp A$ *and* $\Lambda\sharp A \Rightarrow \Theta$ *have separated derivations in* $LJ^=$ *or* $LK^=$*, then also* $\Gamma, \Lambda \Rightarrow \Delta, \Theta$ *has a separated derivation in the same system.*

Proof. Let $\mathcal{D}$ and $\mathcal{E}$ be separated derivations of $\Gamma \Rightarrow \Delta\sharp A$ and $\Lambda\sharp A \Rightarrow \Theta$ respectively. If $\Delta\sharp A = \Delta$ or $\Lambda\sharp A = \Lambda$ the desired derivation can be obtained by weakening the conclusion of $\mathcal{D}$ or of $\mathcal{E}$. If both $\mathcal{D}$ and $\mathcal{E}$ end with an equality-inference or with a cut, then $\mathcal{D}$ and $\mathcal{E}$, being separated, do not contain any logical inference. Then it suffices to contract all the occurrences of A in $\Delta\sharp A$ not occurring in Δ and, similarly, all those occurring in $\Lambda\sharp A$ but not in Λ, into a single one, and apply an atomic cut on A. If $\mathcal{D}$ or $\mathcal{E}$, say $\mathcal{D}$, ends with a weak structural inference or with a logical inference, we proceed by induction on the sum $h(\mathcal{D}) + h(\mathcal{E})$ of the heights of $\mathcal{D}$ and $\mathcal{E}$.

Case 1 $\mathcal{D}$ ends with a weak structural inference. Then the argument is the same as in Case 2 of the proof of Proposition 8.4.

Case 2 $\mathcal{D}$ ends with a logical inference. Since A is atomic, A is not the principal formula of such an inference. Then the conclusion follows by a straightforward induction on $h(\mathcal{D}) + h(\mathcal{E})$.

The cases in which it is $\mathcal{E}$ to end with a weak structural inference or with a logical inference are entirely analogous. □

PROPOSITION 8.6 *Every derivable sequent in* $LJ^=$ *or* $LK^=$ *has a separated derivation in the same system.*

Proof Assume we are given a non separated derivation $\mathcal{D}$ of $\Gamma \Rightarrow \Delta$ in $LJ^=$ or $LK^=$. By Proposition 8.3, $\mathcal{D}$ can be transformed into a derivation $\mathcal{D}'$ whose equality and cut-inferences are atomic. Then a straightforward induction on the height of $\mathcal{D}'$, based on Proposition 8.4 and 8.5, shows that $\mathcal{D}'$ can be transformed into a separated derivation of $\Gamma \Rightarrow \Delta$. □