The Journal of Symbolic Logic
Volume 84, Number 1, March 2019

UNSOUND INFERENCES MAKE PROOFS SHORTER

JUAN P. AGUILERA AND MATTHIAS BAAZ

Abstract. We give examples of calculi that extend Gentzen's sequent calculus LK by unsound quantifier inferences in such a way that (i) derivations lead only to true sequents, and (ii) proofs therein are nonelementarily shorter than LK-proofs.

§1. Introduction.

Consider the following argument:

1. That Kurt Gödel is Austrian entails that Kurt Gödel is Austrian.
2. Hence, that Kurt Gödel is Austrian entails that everyone is Austrian.
3. That is, if Kurt Gödel is Austrian, then all people are Austrian.
4. Therefore, there exists a person such that, if that person is Austrian, then all people are Austrian.

The argument can be formalized as follows:

$$\dfrac{\dfrac{\dfrac{A(a) \vdash A(a)}{A(a) \vdash \forall y\, A(y)}}{\vdash A(a) \to \forall y\, A(y)}}{\vdash \exists x \left(A(x) \to \forall y\, A(y)\right)} \tag{1}$$

The study of formal proofs is motivated by our desire for deductive reasoning to be *correct*—we wish that the procedures involved derive only true conclusions. The traditional way of ensuring this involves restricting proofs to those satisfying the two following properties:

- Inferences are *sound*, i.e., only true conclusions result from true premises.
- Derivations are *hereditary*, i.e., initial segments of proofs are proofs themselves.

In particular, one might impose certain *characteristic-variable conditions* that restrict the circumstances under which quantifiers can be inferred. One can use unsound rules based on weak characteristic-variable conditions to define enhanced (correct) logical calculi wherein derivations such as (1) are allowed.

In our current treatment, the correctness of proofs can only be gauged by considering them in their entirety, as forfeiting soundness while maintaining correctness necessarily violates hereditariness. This is not unlike other aspects of reasoning that

Received August 26, 2016.
2010 *Mathematics Subject Classification.* 03F03, 03F05, 03F07.
Key words and phrases. unsound inferences, sequent calculus, cut elimination, eigenvariable condition.

© 2019, Association for Symbolic Logic
0022-4812/19/8401-0005
DOI:10.1017/jsl.2018.51

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

are frequently not reflected in formal derivations. For example, hereditariness is not compatible with proofs by contradiction—one would not take as correct an initial segment of a proof obtained by interrupting it after an assumption directed towards a contradiction had been made, but before the contradiction had been reached.

To illustrate this, we note that the following subproof of the opening example should not be allowed as a proof in the system:

1. That Kurt Gödel is Austrian entails that Kurt Gödel is Austrian.
2. Hence, that Kurt Gödel is Austrian entails that everyone is Austrian.

The *strong* inference from 1 to 2 is the only unsound one:

1. That Kurt Gödel is Austrian entails that Kurt Gödel is Austrian.
3′ That is, if Kurt Gödel is Austrian, then Kurt Gödel is Austrian.
4′ Therefore, there exists a person such that, if that person is Austrian, then Kurt Gödel is Austrian.

Here, the name 'Kurt Gödel' is simply manipulated syntactically, so its meaning plays no role. Moreover, the name does not even appear in sentence 4. It could well have been replaced by any other throughout the derivation with no effect—we call (the result of formalizing) this condition *substitutability*. In our example, it is what allows us to mend the proof after a 'lie' has been introduced in 2.

Formal proofs can often be taken without loss of generality to be *regular*, i.e., any variable which is the *characteristic variable*[1] of a quantifier inference can be assumed to appear only before then. Forfeiting this condition allows us to accept the following proof:

$$\dfrac{\dfrac{\dfrac{\dfrac{A(a)\vdash A(a)}{A(a)\vdash \forall x\, A(x)} \quad \dfrac{A(f(a))\vdash A(f(a))}{\forall x\, A(x)\vdash A(f(a))}}{A(a)\vdash A(f(a))}}{\vdash A(a)\to A(f(a))}}{\vdash \exists x\,\big(A(x)\to A(f(x))\big)} \tag{2}$$

In this example, a is the characteristic variable of only one quantifier inference. We call this *weak regularity*. Our main result is that a sequent calculus augmented with unsound quantifier inferences satisfying substitutability, weak regularity (or even a further weakening thereof), and a technical *side-variable condition* is correct. This means that, although inferences therein may not be sound, the three (global) conditions are enough to guarantee that any lie introduced during the proof is eventually cleared. Moreover, the calculus yields nonelementarily shorter[2] cut-free proofs than standard proof systems. This is proved in Section 3. Some consequences of the speed-up theorem are explained in Section 4.

The three characteristic-variable conditions are motivated by the rules governing possible inferences in Hilbert's ε-*calculus*, whose language contains no quantifiers (see Section 5). As a consequence, proofs therein are less restrictive. The weakened conditions thus result from attempting to pull back some extra freedom from the ε-calculus to first-order logic. They can also be thought of as a first attempt in tracking down and covering the spots where first-order logic proofs might break

[1] See Section 2 for preliminaries and definitions.
[2] Size can be measured, e.g., by counting the number of inferences in the proof.

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

down without the usual quantifier-inference restrictions. In particular, it seems plausible that the conditions here can be further weakened, and the corresponding calculi can result even faster (see Section 6).

Weakened characteristic-variable conditions essentially identical to these have been considered by Takeuti [9] in the context of infinitary proof systems.

§2. Quantifier inferences. We consider derivations in sequent calculi. Sequents are expressions of the form $\Gamma \vdash \Delta$, where Γ and Δ represent collections of formulae. The sequent $\Gamma \vdash \Delta$ is usually interpreted as 'if all of Γ hold, then at least one of the formulae in Δ holds.' Equivalently, if $\Gamma = A_0, A_1, \ldots, A_n$ and $\Delta = B_0, B_1, \ldots, B_m$, then $\Gamma \vdash \Delta$ is interpreted as

$$\bigvee_{i \leq n} \neg A_i \vee \bigvee_{i \leq m} B_i. \tag{3}$$

Inferences are expressions of the form

$$\frac{\Gamma \vdash \Delta}{\Gamma' \vdash \Delta'}$$

which are to be interpreted as 'the sequent $\Gamma' \vdash \Delta'$ follows from the sequent $\Gamma \vdash \Delta$.' Derivations are trees whose leaves are axioms and whose nonleaf nodes are sequents obtained from their predecessors by inferences; they are to be interpreted as *proofs* of the root sequent. In fact, we frequently refer to derivations as 'proofs.'

If a formula is changed by an inference, we say that the formula is an *auxiliary* formula of the inference. If so, then the resulting formula in the conclusion is called the *principal* formula. An inference is a *quantifier inference* if the (only) principal formula has a quantifier as its outermost logical symbol. Normally, quantifier inferences consist of substituting a formula[3] $Qx\, A(x)$ for an instance thereof in a sequent, e.g., as in

$$\frac{\Gamma \vdash \Delta, A(a)}{\Gamma \vdash \Delta, \forall x\, A(x)}$$

The *polarity* of a quantifier in a sequent $\Gamma \vdash \Delta$ is defined as follows: rewrite $\Gamma \vdash \Delta$ in the form (3) and then rewrite the resulting expression without implication symbols. A quantifier is *positive* if it is universal (resp. existential) and under the scope of an even (resp. odd) number of negation symbols, and *negative* otherwise. A quantifier is *strong* if it is positive and on the right-hand side of a sequent or negative and on the left-hand side of a sequent; it is *weak* otherwise.

If an inference yields a strongly quantified formula $Qx\, A(x)$ from $A(a)$, where a is a free variable, we say that a is the *characteristic variable* of the inference. We will denote free variables by letters $a, b, c, \ldots$ (by contrast, we denote bound variables by $x, y, z, \ldots$). We denote closed terms by $t, s, r, \ldots$ and variants thereof.

Let π be any derivation. We say b is a *side variable* of a in π (written $a <_\pi b$) if π contains a strong-quantifier inference of the form:

$$\frac{\Gamma \vdash \Delta, A(a, b, \vec{c})}{\Gamma \vdash \Delta, \forall x\, A(x, b, \vec{c})}$$

[3]We will frequently use Q to denote any unspecified quantifier.

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

or of the form:

$$\frac{A(a,b,\vec{c}),\Gamma \vdash \Delta}{\exists x\, A(x,b,\vec{c}),\Gamma \vdash \Delta}$$

The Skolemization of a first-order formula A is defined by replacing every strongly quantified variable y with a new function symbol $f_y^A(x_1,\ldots,x_n)$, where $x_1,\ldots,x_n$ are the weakly quantified variables such that Qy appears in the scope of their quantifiers, and removing the quantifier Qy. We write $sk(A)$ for the Skolemization of A. Recall that whether a quantifier is weak or strong depends on the side of the sequent on which it is. Thus, $sk(A)$ means different things depending on whether A appears in the antecedent or the succedent. The Skolemization of a sequent

$$A_0,\ldots,A_m \vdash A_{m+1},\ldots,A_{m+n}$$

is defined as

$$sk(A_0),\ldots,sk(A_m) \vdash sk(A_{m+1}),\ldots,sk(A_{m+n}).$$

In classical logic, a sequent is derivable if and only if its Skolemization is.

Definition 2.1 (Suitable quantifier inference). We say a quantifier inference is *suitable for a proof* π if either it is a weak-quantifier inference, or the following three conditions are satisfied:

- (substitutability) the characteristic variable does not appear in the conclusion of π.
- (side-variable condition) the relation $<_\pi$ is acyclic.
- (weak regularity) the characteristic variable is not the characteristic variable of another strong-quantifier inference in π.

In particular, if all the quantifier inferences in a proof π are suitable for it, then every free variable a in π is the characteristic variable of at most one inference, in whose principal formula a does not appear (since $a <_\pi a$ cannot hold).

We will work with various sequent calculi extending *Propositional* LK, the sequent calculus whose only axioms are $\bot \vdash$ and $A \vdash A$ (with A atomic) and whose rules are:

Structural rules:

$$\frac{\Gamma \vdash \Delta}{\Gamma, A \vdash \Delta}\ WL \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, A}\ WR$$

$$\frac{\Gamma, A, A \vdash \Delta}{\Gamma, A \vdash \Delta}\ CL \qquad \frac{\Gamma \vdash \Delta, A, A}{\Gamma \vdash \Delta, A}\ CR$$

$$\frac{\Gamma_1, A, B, \Gamma_2 \vdash \Delta}{\Gamma_1, B, A, \Gamma_2 \vdash \Delta}\ EL \qquad \frac{\Gamma \vdash \Delta_1, A, B, \Delta_2}{\Gamma \vdash \Delta_1, B, A, \Delta_2}\ ER$$

$$\frac{\Gamma \vdash \Delta, A \quad A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta}\ Cut$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

Logical rules:

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta} \wedge_1 L \qquad \frac{\Gamma \vdash \Delta, A}{\Gamma \vdash \Delta, A \vee B} \vee_1 R$$

$$\frac{\Gamma, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta} \wedge_2 L \qquad \frac{\Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \vee B} \vee_2 R$$

$$\frac{\Gamma, A \vdash \Delta \quad \Gamma, B \vdash \Delta}{\Gamma, A \vee B \vdash \Delta} \vee L \qquad \frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \wedge B} \wedge R$$

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma, B \vdash \Delta}{\Gamma, A \to B \vdash \Delta} \to L \qquad \frac{\Gamma, A \vdash \Delta, B}{\Gamma \vdash \Delta, A \to B} \to R$$

Negation $\neg A$ is defined as $A \to \bot$. The names of the structural rules stand, respectively, for 'weakening,' 'contraction,' and 'exchange.' First-order LK is the extension of Propositional LK obtained by adding quantifier inferences:

$$\frac{\Gamma, A(t) \vdash \Delta}{\Gamma, \forall x\, A(x) \vdash \Delta} \forall L \qquad \frac{\Gamma \vdash \Delta, A(t)}{\Gamma \vdash \Delta, \exists x\, A(x)} \exists R$$

$$\frac{\Gamma, A(a) \vdash \Delta}{\Gamma, \exists x\, A(x) \vdash \Delta} \exists L \qquad \frac{\Gamma \vdash \Delta, A(a)}{\Gamma \vdash \Delta, \forall x\, A(x)} \forall R$$

with the only restriction that in the inferences $\forall R$ and $\exists L$, the variable a is not allowed to appear in the conclusion. In particular, note that every quantifier inference is suitable for every regular LK-proof. Gentzen's famous *Cut-Elimination Theorem* states that the cut rule is redundant in both propositional and first-order LK.

DEFINITION 2.2 (LK^+). The calculus LK^+ is defined like LK, except that one drops the restriction preventing the characteristic variable of an inference from appearing in its conclusion, and one adds the restriction that a proof may only contain quantifier inferences that are suitable for it.

A further weakening of the characteristic-variable conditions gives rise to the notion of weak suitability:

DEFINITION 2.3 (Weakly suitable quantifier inference). A quantifier inference is *weakly suitable for a proof* π if either it is a weak-quantifier inference or it satisfies substitutability, the side-variable condition, and:

- (very weak regularity) whenever the characteristic variable is also the characteristic variable of another strong-quantifier inference in π, then it has the same principal formula.

DEFINITION 2.4 (LK^{++}). The calculus LK^{++} is the extension of LK^+ that results from allowing all weakly suitable quantifier inferences.

It is easy to find examples of sequents that are more easily provable in LK^+ than in LK:

EXAMPLE 2.5. The sequent

$$\forall x\, A(x) \to B \vdash \exists x\, (A(x) \to B)$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

is provable in LK:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A(a)\vdash A(a)}{A(a)\vdash A(a),B}}{\vdash A(a),A(a)\to B}}{\vdash A(a),\exists x\,(A(x)\to B)}}{\vdash \exists x\,(A(x)\to B),A(a)}}{\vdash \exists x\,(A(x)\to B),\forall x\,A(x)}\qquad B\vdash B}{\forall x\,A(x)\to B\vdash \exists x\,(A(x)\to B),B}}{\forall x\,A(x)\to B,A(b)\vdash \exists x\,(A(x)\to B),B}}{\forall x\,A(x)\to B\vdash \exists x\,(A(x)\to B),A(b)\to B}}{\forall x\,A(x)\to B\vdash \exists x\,(A(x)\to B),\exists x\,(A(x)\to B)}}{\forall x\,A(x)\to B\vdash \exists x\,(A(x)\to B)}$$

However, one can find a shorter LK^{+}-proof:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A(a)\vdash A(a)}{A(a)\vdash \forall x\,A(x)}\qquad B\vdash B}{A(a),\forall x\,A(x)\to B\vdash B}}{\forall x\,A(x)\to B,A(a)\vdash B}}{\forall x\,A(x)\to B\vdash A(a)\to B}}{\forall x\,A(x)\to B\vdash \exists x\,(A(x)\to B)}$$

⊣

Recall that a function on the natural numbers is *elementary* if it can be defined by a quantifier-free formula from $+$, $\cdot$, and the function $x\mapsto 2^x$. By independent results of R. Statman [8] (which were formalized in [3]) and of V. P. Orevkov [7], the sizes of the smallest cut-free LK-proofs of sequents of length n are not bounded by any elementary function on n. Our main theorem is that cut-free LK^{+}-proofs are nonelementarily shorter than cut-free LK-proofs:

THEOREM 2.6. *There is no elementary function bounding the length of the shortest cut-free* LK*-proof of a formula in terms of its shortest cut-free* LK^{+}*-proof.*

An immediate consequence is the following:

COROLLARY 2.7. *There is no elementary function bounding the length of the shortest cut-free* LK*-proof of a formula in terms of its shortest cut-free* LK^{++}*-proof.*

We prove Theorem 2.6 in Section 3. First, we consider the question of the correctness of LK^{++}. The proof of Theorem 2.8 that we present here is due to the referee:

THEOREM 2.8. *If a sequent is* LK^{++}*-derivable, then it is already* LK*-derivable.*

PROOF. Let π be an LK^{++}-proof. Replace every unsound universal quantifier inference by a $\to L$ inference:

$$\dfrac{\Gamma\vdash \Delta,A(a)\qquad \forall x\,A(x)\vdash \forall x\,A(x)}{\Gamma,A(a)\to \forall x\,A(x)\vdash \Delta,\forall x\,A(x)}$$

Similarly replace every unsound existential quantifier by an $\to L$ inference

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

$$\frac{\exists x\, A(x) \vdash \exists x\, A(x) \qquad A(a), \Gamma \vdash \Delta}{\Gamma, \exists x\, A(x), \exists x\, A(x) \to A(a) \vdash \Delta}$$

By doing this, we obtain a proof of the desired sequent, together with many formulae of the form $A(a) \to \forall x\, A(x)$ or $\exists x\, A(x) \to A(a)$ on the left-hand side. However, we can eliminate each of them by adding an existential quantifier inference and cutting with formulae of the form

$$\vdash \exists y \left(A(y) \to \forall x\, A(x)\right)$$

or of the form

$$\vdash \exists y \left(\exists x\, A(x) \to A(y)\right),$$

both of which are easily derivable. Note that the existential quantifier inferences can be carried out in a way that is permissible by LK because the initial proof satisfied substitutability, weak regularity and the side-variable condition. ⊣

COROLLARY 2.9. *If a sequent is derivable in* LK^+, *then it is already derivable in* LK.

§3. **Nonelementary speed-up.** Our strategy for proving Theorem 2.6 is to show that LK^+ simulates a strong calculus that is already nonelementarily faster than LK. It is related to *quantifier shifts*, formulae such as

$$(\forall x\, A \to B) \to \exists x\, (A \to B).$$

The main feature of the calculus is the addition of certain rules for quickly shifting quantifiers. Below, if A is a subformula of B, we say that B is a *context* of A. We generally write $\kappa[A]$ to emphasize that A occurs in a context $\kappa[\cdot]$; they should be thought as syntactical operators $A \mapsto \kappa[A]$.

DEFINITION 3.1. The calculus $\mathrm{LK}_{\mathtt{shift}}$ is obtained by extending LK with the following rules:

$$\frac{\Gamma, \kappa[Qx\, A \lhd B] \vdash \Delta}{\Gamma, \kappa[Q' x\, (A \lhd B)] \vdash \Delta} \qquad \frac{\Gamma, \kappa[A \lhd Qx\, B] \vdash \Delta}{\Gamma, \kappa[Q' x\, (A \lhd B)] \vdash \Delta}$$

$$\frac{\Gamma \vdash \Delta, \kappa[Qx\, A \lhd B]}{\Gamma \vdash \Delta, \kappa[Q' x\, (A \lhd B)]} \qquad \frac{\Gamma \vdash \Delta, \kappa[A \lhd Qx\, B]}{\Gamma \vdash \Delta, \kappa[Q' x\, (A \lhd B)]}$$

where $\kappa[\cdot]$ is a context, $\lhd \in \{\wedge, \vee, \to\}$, x does not appear in B, and $Q' = Q$, except if $\lhd$ is $\to$ and Q is taken from the antecedent, in which case Q' is opposite. We refer to these rules as *deep quantifier shifts*.

PROPOSITION 3.2. *Cut-free* LK^+ *simulates cut-free* $\mathrm{LK}_{\mathtt{shift}}$ *double-exponentially, i.e., every* $\mathrm{LK}_{\mathtt{shift}}$*-provable sequent is* LK^+*-provable and there is a double exponential function that bounds the length of the least cut-free* LK^+*-proof of a* LK^+*-provable sequent in terms of its least cut-free* $\mathrm{LK}_{\mathtt{shift}}$*-proof.*

PROOF. Let π be a cut-free $\mathrm{LK}_{\mathtt{shift}}$-proof of size n. We transform it into a cut-free LK^+-proof. Assume by induction that there is only one application of a quantifier shift and this is the last inference. We proceed by cases. For simplicity, we assume κ does not change the polarity of the quantifiers. We also assume $\lhd$ is $\to$; the other connectives are treated similarly.

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

Case I. The last inference is

$$\frac{\Gamma, \kappa[\forall x\, A(x) \to B] \vdash \Delta}{\Gamma, \kappa[\exists x\, (A(x) \to B)] \vdash \Delta}$$

so that the proof has the following structure:

$$\dfrac{\dfrac{\dfrac{\vdots\sigma_2}{\Gamma'' \vdash \Delta'', A(a)}}{\Gamma'' \vdash \Delta'', \forall x\, A(x)}\,(*) \qquad \dfrac{\vdots\sigma_3}{\Gamma', B \vdash \Delta'}}{\dfrac{\dfrac{\vdots\sigma_1}{\Gamma' \vdash \Delta', \forall x\, A(x)} \qquad \Gamma', \forall x\, A(x) \to B \vdash \Delta'}{\dfrac{\dfrac{\vdots\sigma_0}{\Gamma, \kappa[\forall x\, A(x) \to B] \vdash \Delta}}{\Gamma, \kappa[\exists x\, (A(x) \to B)] \vdash \Delta}}}$$

where the σ_i denote subproofs. The subproof σ_0 might split into several branches not indicated in the diagram, each of which could potentially include its own copy of $\forall x\, A(x)$ being inferred and its own subproof σ_3' with conclusion of the form $\Pi, B \vdash \Lambda$. Each branch can be dealt with the same way and so we will only focus on the indicated parts of π.

Similarly, the subproof σ_1 might split into several branches not indicated in the diagram, each of which could potentially include its own copy of $\forall x\, A(x)$ being inferred. This we need to bear in mind. Note that each of those copies of $\forall x\, A(x)$ is necessarily inferred from a different characteristic variable. In fact, the variable a does not appear below $(*)$, by the regularity of $\mathrm{LK}_{\mathtt{shift}}$.

We modify the proof. Our approach is as follows: we would like to merge the subproofs σ_1 and σ_2 simply by postponing the inference $(*)$ as follows:

$$\dfrac{\dfrac{\dfrac{\vdots\sigma_2 + \sigma_1}{\Gamma' \vdash \Delta', A(a)} \qquad \dfrac{\vdots\sigma_3}{\Gamma', B \vdash \Delta'}}{\Gamma', A(a) \to B \vdash \Delta'}\,(**)}{\dfrac{\vdots\sigma_0}{\Gamma, \kappa[\exists x\, (A(x) \to B)] \vdash \Delta}}$$

The problem that might arise is that some occurrence of $\forall x\, A(x)$ that would be contracted in σ_1 with the indicated occurrence is unable to be contracted. We describe how to deal with each of them.

Notice that the problematic occurrence of $\forall x\, A(x)$ must originate from a strong-quantifier inference. Omit that inference and drag the unquantified formula $A(b)$ until after $(**)$, so that at that point we have a derivation of

$$\Gamma_0', A(b), \Gamma_1', \exists x\, (A(x) \to B), \vdash \Delta',$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

where $\Gamma' = \Gamma'_0, \Gamma'_1$. Add to that some exchanges to obtain

$$\Gamma', \exists x\,(A(x) \to B), A(b) \vdash \Delta',$$

and a subproof

$$\frac{\vdots\sigma_3}{\Gamma', B \vdash \Delta'} \tag{4}$$

in order to infer

$$\Gamma', \exists x\,(A(x) \to B), A(b) \to B \vdash \Delta',$$

with an application of $\exists L$ and a contraction, we are left again with

$$\Gamma', \exists x\,(A(x) \to B) \vdash \Delta',$$

to which we can apply σ_0 as desired. The proof grows by adding a copy of (4)—of size at most n—for each problematic occurrence of $\forall x\,A(x)$ (of which there are at most n). Hence, it grows quadratically.

Case II. The last inference is

$$\frac{\Gamma, \kappa[\exists x\,A(x) \to B] \vdash \Delta}{\Gamma, \kappa[\forall x\,(A(x) \to B)] \vdash \Delta}$$

The proof has the following structure:

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots\sigma_2}{\Gamma'' \vdash \Delta'', A(t)}}{\Gamma'' \vdash \Delta'', \exists x\,A(x)}}{\dfrac{\dfrac{\vdots\sigma_1}{\Gamma' \vdash \Delta', \exists x\,A(x)} \quad \dfrac{\vdots\sigma_3}{\Gamma', B \vdash \Delta'}}{\Gamma', \exists x\,A(x) \to B \vdash \Delta'}}}{\dfrac{\dfrac{\vdots\sigma_0}{\Gamma, \kappa[\exists x\,A(x) \to B] \vdash \Delta}}{\Gamma, \kappa[\forall x\,(A(x) \to B)] \vdash \Delta}}$$

As in Case I, the subproof σ_0 could branch off into several subtrees each of which is taken care of in the same way, and so we only focus on the indicated parts of π. We would like to merge the subproofs σ_1 and σ_2 as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots\sigma_2 + \sigma_1}{\Gamma' \vdash \Delta', A(t)} \quad \dfrac{\vdots\sigma_3}{\Gamma', B \vdash \Delta'}}{\Gamma', A(t) \to B \vdash \Delta'}}{\Gamma', \forall x\,(A(x) \to B) \vdash \Delta'}}{\dfrac{\vdots\sigma_0}{\Gamma, \kappa[\forall x\,(A(x) \to B)] \vdash \Delta}}$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

As before, we face the problem of circumventing a contraction of (possibly several occurrences of) $\exists x\, A(x)$ in σ_1 and solve it in the same way. The proof grows quadratically again. We note the following: it might happen that the term t contains some free variable a. Hence, the proof of this case does not necessarily go through in LK, as a could be the characteristic variable of a strong-quantifier inference in σ_1. Substitutability holds clearly. Since no free variable in B is a characteristic variable in the proof, the side-variable condition is satisfied.

Case III. The last inference is

$$\frac{\Gamma, \kappa\big[A \to \forall x\, B(x)\big] \vdash \Delta}{\Gamma, \kappa\big[\forall x\, (A \to B(x))\big] \vdash \Delta}$$

This is analogous to Case II.

Case IV. The last inference is

$$\frac{\Gamma, \kappa\big[A \to \exists x\, B(x)\big] \vdash \Delta}{\Gamma, \kappa\big[\exists x\, (A \to B(x))\big] \vdash \Delta}$$

This is analogous to Case I.

Case V. The last inference is:

$$\frac{\Gamma \vdash \Delta, \kappa\big[\forall x\, A(x) \to B\big]}{\Gamma \vdash \Delta, \kappa\big[\exists x\, (A(x) \to B)\big]}$$

The proof has the following form, modulo qualifications as in Cases I and II:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdots\sigma_2}{\Gamma'', A(t) \vdash \Delta''}}{\Gamma'', \forall x\, A(x) \vdash \Delta''}\,(*)}{\Gamma', \forall x\, A(x) \vdash \Delta', B}\vdots\sigma_1}{\Gamma' \vdash \Delta', \forall x\, A(x) \to B}}{\Gamma \vdash \Delta, \kappa\big[\forall x\, A(x) \to B\big]}\vdots\sigma_0}{\Gamma \vdash \Delta, \kappa\big[\exists x\, (A(x) \to B)\big]}$$

We would like to postpone the inference $(*)$, thus merging σ_1 and σ_2:

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots\sigma_2 + \sigma_1}{\Gamma', A(t) \vdash \Delta', B}}{\Gamma' \vdash \Delta', A(t) \to B}\,(**)}{\Gamma' \vdash \Delta', \exists x\, (A(x) \to B)}}{\Gamma \vdash \Delta, \kappa[\exists x\, (A(x) \to B)]}\vdots\sigma_0$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

However, we might have—similarly to the previous cases—an occurrence of $\forall x\, A(x)$ that would be contracted on the left-hand side as part of σ_1. For simplicity, assume there is only one. This occurrence is inferred from a formula $A(s)$. Postpone this occurrence until after $(**)$ so that at that point we have a derivation of

$$\Gamma_0', A(s), \Gamma_1' \vdash \Delta', \exists x\, (A(x) \to B),$$

where $\Gamma' = \Gamma_0', \Gamma_1'$. Add to that some exchanges to obtain

$$\Gamma', A(s) \vdash \Delta', \exists x\, (A(x) \to B),$$

and a weakening:

$$\frac{\Gamma', A(s) \vdash \Delta', \exists x\, (A(x) \to B)}{\Gamma', A(s) \vdash \Delta', \exists x\, (A(x) \to B), B}\ WR \tag{5}$$

in order to infer

$$\Gamma' \vdash \Delta', \exists x\, (A(x) \to B), A(s) \to B,$$

with an application of $\exists R$ and a contraction, we are left again with

$$\Gamma' \vdash \Delta', \exists x\, (A(x) \to B),$$

to which we can apply the rest of σ_0 as desired. In this case the proof only grows linearly. As in Case II, the proof of this case would not necessarily work for LK, but it is correct from the point of view of LK^+, as follows from the regularity of $\mathrm{LK}_{\mathtt{shift}}$ and the fact that the conclusion of the proof is π.

Case VI. The last inference is

$$\frac{\Gamma \vdash \Delta, \kappa\left[\exists x\, A(x) \to B\right]}{\Gamma \vdash \Delta, \kappa\left[\forall x\, (A(x) \to B)\right]}$$

The proof has the following form, modulo qualifications as in Cases I and II:

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots \sigma_2}{\Gamma'', A(a) \vdash \Delta''}}{\Gamma'', \exists x\, A(x) \vdash \Delta''}\,(*)}{\dfrac{\dfrac{\vdots \sigma_1}{\Gamma', \exists x\, A(x) \vdash \Delta', B}}{\Gamma' \vdash \Delta', \exists x\, A(x) \to B}}}{\dfrac{\dfrac{\vdots \sigma_0}{\Gamma \vdash \Delta, \kappa\left[\exists x\, A(x) \to B\right]}}{\Gamma \vdash \Delta, \kappa\left[\forall x\, (A(x) \to B)\right]}}$$

As before, we would like to postpone the inference $(*)$, thus merging σ_1 and σ_2:

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdots \sigma_2 + \sigma_1}{\Gamma', A(a) \vdash \Delta', B}}{\Gamma' \vdash \Delta', A(a) \to B}}{\Gamma' \vdash \Delta', \forall x\,(A(x) \to B)}}{\Gamma \vdash \Delta, \kappa[\forall x\,(A(x) \to B)]}}{\vdots \sigma_0}$$

We deal with contractions of $\exists x\, A(x)$ on the left-hand side as in the previous case. The proof grows linearly.

Case VII. The last inference is

$$\dfrac{\Gamma \vdash \Delta, \kappa\big[A \to \forall x\, B(x)\big]}{\Gamma \vdash \Delta, \kappa\big[\forall x\,(A \to B(x))\big]}$$

This is analogous to Case VI.

Case VIII. The last inference is

$$\dfrac{\Gamma \vdash \Delta, \kappa\big[A \to \exists x\, B(x)\big]}{\Gamma \vdash \Delta, \kappa\big[\exists x\,(A \to B(x))\big]}$$

This is analogous to Case V.

Hence, we have dealt with all the cases. Note that for each quantifier shifted, the proof grows at most quadratically. Since there are at most n deep quantifier shifts, the size of the resulting proof is bounded by

$$n^{2^n} = 2^{2^n \cdot \log(n)} \approx 2^{2^n}.$$

This finishes the proof. ⊣

Theorem 2.6 is a consequence of Proposition 3.2 and the following result:

Theorem 3.3. *There is no elementary function bounding the length of the shortest cut-free* LK*-proof of a formula in terms of its shortest cut-free* $\mathrm{LK}_{\mathtt{shift}}$*-proof.*

Proof. We will make use of a very specific family of sequents $\{S_i\}_{i<\omega}$ described in [3] and due to Statman [8], and specific LK-proofs thereof. The sequents and the proofs themselves are not important for our proof. What is relevant is that they have the following properties:

1. the size of S_i is polynomial in i;
2. there is no bound on the size of their smallest cut-free LK-proofs that is elementary in i;
3. the size of these proofs (with cuts), however, is polynomially bounded in i;
4. all cut formulae are closed; we can also assume they are prenex by, e.g., [4, Theorem 3.3].

Let $\Gamma_i \vdash \Delta_i$ be one of the sequents. We modify the proof as follows: first, replace each cut

$$\dfrac{\Gamma \vdash \Delta, A \quad A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta}\ \mathit{Cut}$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

with an application of $\to L$:

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma, A \vdash \Delta}{\Gamma, A \to A \vdash \Delta} \to L$$

We are left with a cut free proof π_0 whose end sequent is of the form:

$$A_0 \to A_0, \ldots, A_m \to A_m, \Gamma_i \vdash \Delta_i. \tag{6}$$

Choose any occurrence of a quantifier in A_0 that is not in the scope of another quantifier. Since it appears once in the antecedent of $A_0 \to A_0$ and once in the consequent, it appears once with each polarity. Because A_0 is assumed to be prenex, we can apply two deep quantifier shifts, so that we are left with a proof π_0' of the sequent:

$$\forall x_0^0 \exists x_1^0 \left(A_0' \to A_0'\right), \ldots, A_m \to A_m, \Gamma_i \vdash \Delta_i. \tag{7}$$

Continue choosing any occurrence of a quantifier in A_0' that is not in the scope of another quantifier (within A_0') and applying deep quantifier shifts in pairs until we obtain a proof π_1 of the sequent:

$$\forall x_0^0 \exists x_1^0 \ldots \left(\hat{A}_0 \to \hat{A}_0\right), \ldots, A_m \to A_m, \Gamma_i \vdash \Delta_i,$$

where $\hat{A}_0$ is quantifier-free and the prefix of $\hat{A}_0 \to \hat{A}_0$ consists of alternating quantifiers. Repeat this procedure for each of the A_i to obtain a proof π_m where each formula $A_i \to A_i$ is replaced by an expression of the form:

$$\forall x_0^i \exists x_1^i \ldots \left(\hat{A}_i \to \hat{A}_i\right),$$

where $\hat{A}_i$ is quantifier free and the quantifier prefix is alternating. Let $\hat{\Gamma}^i \vdash \hat{\Delta}^i$ be the sequent:

$$\forall x_0^0 \exists x_1^0 \ldots \left(\hat{A}_0 \to \hat{A}_0\right), \ldots, \forall x_0^m \exists x_1^m \ldots \left(\hat{A}_m \to \hat{A}_m\right), \Gamma_i \vdash \Delta_i. \tag{8}$$

Note that, since the size of the initial proof was polynomial in i, there were polynomially many quantifiers in the proof; hence, we added only polynomially many deep quantifier shifts, and so the size of the resulting proof π_m of (8) is bounded polynomially in i. Moreover, it is cut-free. Consequently, it suffices to show:

CLAIM 3.4. *There is no elementary function bounding the size of the smallest cut-free* LK*-proofs of* (8).

PROOF. Let $\{\sigma_i\}_{i<\omega}$ be a sequence of such proofs. First, we transform it into a sequence of proofs of the Skolemizations of the sequents (8), so that for each sequent, each of the implications $\hat{A}_i \to \hat{A}_i$ remains only universally quantified. Then, by Herbrand's theorem, there are propositional proofs $\{\theta_i\}_{i<\omega}$ of the Herbrand sequents of (8), each of which is of the form:

$$\bigwedge_k B_k^0, \ldots, \bigwedge_k B_k^m, \Gamma' \vdash \Delta'. \tag{9}$$

(Here Γ' and Δ' are the Skolemizations of Γ_i and Δ_i.) Moreover, the lengths of the Herbrand sequents are bounded exponentially by the lengths of $\{\theta_i\}_{i<\omega}$ (see [3, Theorem 4.3]). Each conjunct B_k^p is a quantifier-free implication of the form:

$$A(t_1, \ldots, t_l) \to A(s_1, \ldots, s_l), \tag{10}$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

for some terms $t_1, \ldots, t_l, s_1, \ldots, s_l$. The key point is that, in the process of Skolemizing the sequents (8), exactly one term in each pair (t_j, s_j) was weakly quantified and the other (which was under the scope of the strong quantifier) was replaced by a Skolem function. Hence, either t_j is of the form $f(s_j, \vec{r})$ or s_j is of the form $f(t_j, \vec{r})$, for some tuple of terms $\vec{r}$.

Transform the sequent (9) as follows: pick a pair (t_j, s_j) of some disjunct B_k^p such that the element thereof that was replaced by a Skolem term—say, $f(s_j, \vec{r}) = t_j$—does not have any other of $t_1, \ldots, t_l, s_1, \ldots, s_l$ as an argument. Such a term of course exists—it is the first term whose quantifier was shifted in (7). Substitute s_j for t_j throughout the sequent.

By repeating this process sufficiently many times, each formula (10) is transformed into a propositional tautology of the form $A(\vec{t}) \to A(\vec{t})$. This means that the sequent

$$\vdash \bigwedge_k B_k^0, \ldots, \bigwedge_k B_k^m$$

is also a propositional tautology, whereby so too is $\Gamma' \vdash \Delta'$, from which follows that it is LK-provable. By [2, Theorem 2], there is an LK-proof of the unskolemized sequent of length exponential in that of $\Gamma' \vdash \Delta'$. But this is impossible, since—by assumption—there exist no short proofs of the unskolemized sequent. ⊣

This establishes Theorem 3.3 and thus Theorem 2.6. ⊣

§**4. Properties of the calculi.** We discuss some properties of the calculi augmented with unsound inferences. Some of the arguments are only sketched—we leave the details to the interested reader.

4.1. $\mathbf{LJ^+}$ **and** $\mathbf{LJ^{++}}$**.** Consider LJ^+ and LJ^{++}, the analogs of the extended calculi LK^+ and LK^{++} for intuitionistic logic. Specifically, these are the calculi obtained by restricting possible sequents to those with at most one formula on the right-hand side. The calculi are not sound for intuitionistic logic, as shown by any of the examples in the introduction. A consequence of this that they do not admit cut elimination. The argument makes use of an important consequence of cut elimination, *subformula property*—every formula appearing in a cut-free proof of a sequent S is a subformula of a formula in S.

PROPOSITION 4.1. LJ^+ *and* LJ^{++} *do not admit cut elimination*.

PROOF. Consider example (2). Since only one formula appears on the right-hand side of each sequent, this is an LJ^+-proof. Suppose towards a contradiction that there were a cut-free proof of

$$\vdash \exists x \left(A(x) \to A(f(x))\right). \tag{11}$$

By the subformula property, this derivation would consist entirely of subformulae of (11). In particular, it would contain no strong quantifier inferences. Therefore, it would already be an LJ-proof. But this is impossible, as (11) is not intuitionistically valid. ⊣

Proposition 4.1 has the consequence that LK^+ and LK^{++} do not admit cut elimination by an algorithm that resembles Gentzen's. We give a rough definition of what we mean by this, but the reader may consult the appendix of [1] for further details.

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

DEFINITION 4.2. We say that a cut-eliminating procedure is *Gentzen-style* if it is a transformation of proofs consisting of permutation of rules, substitution of free variables, reduction of cuts of a formula to cuts of its outermost subformulae, and absorption of axioms, i.e., elimination of cuts

$$\frac{\Gamma \vdash \Delta, A \quad A \vdash A}{\Gamma \vdash \Delta, A}\ \mathit{Cut}$$

by deleting the indicated occurrence of $A \vdash A$.

The point in considering Gentzen-style cut elimination algorithms is that they transform intuitionistic proofs into cut-free intuitionistic proofs. Thus, we obtain:

COROLLARY 4.3. LK^{++} *and* LK^{+} *do not admit any Gentzen-style cut elimination*.

As is well known, the following three formulae are the only quantifier shifts not derivable in LJ:

1. $\forall x\,(A \vee B(x)) \vdash A \vee \forall x\,B(x)$;
2. $(\forall x\,A(x) \to B) \vdash \exists x\,(A(x) \to B)$;
3. $(A \to \exists x\,B(x)) \vdash \exists x\,(A \to B(x))$.

They give rise to a characterization of LJ^{++}:

PROPOSITION 4.4. *A sequent is provable in* LJ^{++} *if and only if it is provable in* LJ *with all quantifier shifts added as axioms*.

PROOF. All three quantifier shifts 1–3 are provable in LJ^{++} (see, e.g., Example 2.5). We show the converse. Clearly, the sequents $\vdash \forall y\,A(y) \to \forall x\,A(x)$ and $\vdash \exists y\,A(y) \to \exists x\,A(x)$ are provable. Hence, using the quantifier shifts 2 and 3 one derives $\vdash \exists y\,(A(y) \to \forall x\,A(x))$ and $\vdash \exists y\,(\exists x\,A(x) \to A(y))$. Therefore, we can apply the proof of Theorem 2.8 to obtain the desired result. ⊣

4.2. Skolemization. The following are applications of (the proof of) Theorem 2.6:

PROPOSITION 4.5. *If* $\Gamma \vdash \Delta$ *is* LK^{+}*-derivable* (*resp.* LK^{++}*-derivable*), *then its Skolemization is* LK^{+}*-derivable* (*resp.* LK^{++}*-derivable*) *in quadratically many steps using additional cuts*.

PROOF. Let $\Gamma \vdash \Delta$ be any LK^{++}-derivable sequent. For each formula A in Δ, the sequent $A \vdash sk(A)$ is already LK-derivable, it follows that, by using one cut per formula, we can derive $\Gamma \vdash sk(\Delta)$ in LK^{++}. Similarly, $sk(A) \vdash A$ is already LK-derivable. Hence, by also using one cut for each formula in Γ, we can derive $sk(\Gamma) \vdash sk(\Delta)$. ⊣

However, as we see, the additional cuts are necessary:

PROPOSITION 4.6. *Let S be a sequent. Then, there is no elementary bound on the length of the smallest cut-free* LK^{++}*-proof of the Skolemization of S in terms of the smallest cut-free* LK^{+}*-proof of S*.

PROOF. We argue as in the proof of Claim 3.4: if the Skolemization of $\Gamma \vdash \Delta$ were cut-free LK^{++}-derivable in elementary many steps, the cut-free LK^{++}-proof would already be an LK-proof, whence by [2, Theorem 2], there would be an LK-proof of $\Gamma \vdash \Delta$ of elementary length with respect to the LK^{+}-proof. ⊣

Note, by contrast, that a cut-free proof can be Skolemized in LK at no extra cost.

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

PROPOSITION 4.7. *There is no elementary bound on the length of the Herbrand sequent of a sequent in terms of its smallest cut-free* LK^{+}*-proof* (*resp.* LK^{++}*-proof*).

PROOF. This follows from Proposition 4.6, as a cut-free LK^{++}-proof of the Skolemization of a sequent contains no strong quantifier and is hence already an LK-proof, whence we can apply Herbrand's theorem. ⊣

Although deskolemization of cut-free proofs is exponential in LK, it is linear in LK^{++}:

PROPOSITION 4.8. *If the Skolemization of* $\Gamma \vdash \Delta$ *is cut-free* LK^{++}*-derivable, then* $\Gamma \vdash \Delta$ *is cut-free* LK^{++}*-derivable in linearly many steps.*

PROOF. Let π be a cut-free LK^{++}-proof of the Skolemization of $\Gamma \vdash \Delta$. By definition, the sets of Skolem functions assigned to each formula in $\Gamma \vdash \Delta$ are disjoint. Modify π into a proof of $\Gamma \vdash \Delta$ as follows:

First, replace each Skolem term t by a new free variable a_t (recall that the term t is of the form $f_y(x_0, x_1, \ldots, x_n)$, for some variables y and $x_0, \ldots, x_n$). Then, for each such variable a_t, add to the proof a strong quantifier inference with characteristic variable a_t below each topmost sequent in which (for t as above) a_t appears and each variable among $x_0, \ldots, x_n$ is bound by a (weak) quantifier having a_t in its scope. This results in an LK^{++}-proof having $\Gamma \vdash \Delta$ as its end-sequent. Substitutability clearly holds, as does weak regularity. The side-variable condition must also hold, since all weak quantifier inferences are obtained from the new free variables indexed by Skolem terms, and the subterm relation of (Skolem) terms is acyclic. ⊣

REMARK 4.9. The proof of Proposition 4.8 goes through for LJ^{++}.

§5. Connections to the epsilon calculus. We start by recalling the sequent-calculus reformulation of Hilbert's ε*-calculus* (see [5]).

DEFINITION 5.1. The quantifier-free language $\mathcal{L}_\varepsilon$ is obtained by removing symbols $\exists$ and $\forall$ from the language of first-order logic and adding symbols ε and τ. Formulae and terms are simultaneously defined by induction in a way that the following clauses are satisfied. We leave the precise recursion to the reader.

1. constants and free variables are terms;
2. if $t_1, \ldots, t_n$ are terms and f is an n-ary function symbol, then $f(t_1, \ldots, t_n)$ is a term;
3. if $A(t)$ is a formula, t is a term, and x is a bound variable, then $\varepsilon_x A(x)$ and $\tau_x A(x)$ are terms;
4. if $t_0, \ldots, t_n$ are terms and A is an n-ary predicate symbol, then $A(t_0, \ldots, t_n)$ is a formula;
5. if A and B are formulae, then $A \wedge B$, $A \vee B$, $\neg A$, and $A \to B$ are formulae.

The ε*-sequent calculus* $\mathrm{LK}^{\varepsilon}$ is obtained by adding to Propositional LK the rules:

$$\frac{\Gamma, A(t) \vdash \Delta}{\Gamma, A(\tau_x A(x)) \vdash \Delta}\ \tau \qquad\qquad \frac{\Gamma \vdash \Delta, A(t)}{\Gamma \vdash \Delta, A(\varepsilon_x A(x))}\ \varepsilon$$

The term $\varepsilon_x A(x)$ is to be understood as a 'generic' object satisfying property A. The dual term $\tau_x A(x)$ is added for symmetry; it is in fact redundant—one can define $A(\tau_x A) = A(\varepsilon_x(\neg A))$. Terms of the form $\varepsilon_x A(x)$ are called ε*-terms* and terms of

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

the form $\tau_x A(x)$ are called τ-*terms*. The notion of an ε- or a τ-term being *weak* or *strong* is defined in analogy with quantifiers in first-order logic.

The *standard translation* of first-order logic into epsilon calculus is defined recursively by mapping atomic formulae to themselves, respecting Booleans, and dealing with quantified formulae as follows: suppose A' is the standard translation of A, then we map:

$$\exists x\, A(x) \mapsto A'(\varepsilon_x A'(x)), \qquad \forall x\, A(x) \mapsto A'(\tau_x A'(x)).$$

PROPOSITION 5.2. *A sequent is* LK*-derivable if and only if the standard translation of its Skolemization is* $\mathrm{LK}^{\varepsilon}$*-derivable.*

In the $\mathrm{LK}^{\varepsilon}$-proof of the standard translation of a sequent, ε- and τ-rules play the role of weak quantifier inferences, while the substitution of variables for ε- and τ-terms throughout the proof plays the role of strong quantifier inferences. The converse can be proved by appealing to the *Extended First Epsilon Theorem* (cf. [6, Theorem 16]) and the elimination of Skolem functions as in [2]. It can also be proved by a semantic argument.

The characteristic-variable conditions of LK^{++} ensure the following:

PROPOSITION 5.3. *If a sequent is has an* LK^{++}*-proof of length* k*, then its standard translation has an* $\mathrm{LK}^{\varepsilon}$*-proof of length* $\leq k$*.*

The proof proceeds by replacing weak-quantifier inferences with ε- or τ-rules and omitting strong quantifier inferences—instead, we substitute the appropriate ε- or τ-term for the characteristic variable throughout the proof.

Proposition 5.3 readily yields an alternative proof of Theorem 2.8: let $\Gamma \vdash \Delta$ be an LK^{++}-provable sequent. Then its Skolemization is LK^{++}-provable by Proposition 4.5. But the Skolemization is common to LK and preserves validity, so by Propositions 5.2 and 5.3, $\Gamma \vdash \Delta$ is LK-provable.

One can state the relationship between the ε-calculus and LJ^{++}.

THEOREM 5.4. *A sequent is* LJ^{++}*-provable if and only if its standard translation is* $\mathrm{LJ}^{\varepsilon}$*-provable.*

PROOF. Suppose $\Gamma \vdash D$ is LJ^{++}-provable. By Proposition 4.4 it is then LJ-provable by adding quantifier shifts as axioms. As in Proposition 5.3, the standard translation of $\Gamma \vdash D$ is $\mathrm{LJ}^{\varepsilon}$-provable by adding the standard translations of quantifier shifts as axioms. However, the standard translations of quantifier shifts are already $\mathrm{LJ}^{\varepsilon}$-provable.

Conversely, suppose the standard translation of $\Gamma \vdash D$ is $\mathrm{LJ}^{\varepsilon}$-provable, say by a proof π. We want to transform π into an LJ^{++}-proof. One of the problems is that π might have inferences arranged the following way:

$$\dfrac{\dfrac{\dfrac{\vdots}{\Pi \vdash B(s)}}{\Pi \vdash B(\varepsilon_x B(x))}}{\dfrac{\dfrac{\vdots}{\Pi' \vdash A(\varepsilon_x B(x))}}{\Pi' \vdash A(\varepsilon_x A(x))}}$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

If so, let us refer to

$$\frac{\Pi \vdash B(s)}{\Pi \vdash B(\varepsilon_x B(x))}$$

as an *inner inference*.

CLAIM 5.5. *Suppose that π is an LJ^ε-proof of the standard translation of a sequent $\Gamma \vdash D$ in which no strong quantifiers appear. Then there is a LJ^ε-proof of the same sequent in whose axioms appear no ε- or τ-terms.*

PROOF. Suppose

$$A(\varepsilon_x B, \vec{b}) \vdash A(\varepsilon_x B, \vec{b})$$

is an axiom in the proof. Replace $\varepsilon_x B$ by a new variable $x_{\varepsilon_x B}$ throughout the proof. Below each inference of the form

$$\frac{\Pi \vdash C}{\Pi \vdash B(x_{\varepsilon_x B})}$$

such that $C \neq B(x_{\varepsilon_x B})$, add an inference

$$\frac{\Pi \vdash B(x_{\varepsilon_x B})}{\Pi \vdash B(\varepsilon_x B)}$$

and proceed similarly for τ-terms. To see that the resulting proof has the same conclusion, notice that since the conclusion of π contains no strong ε- or τ-terms, any strong ε- or τ-terms must be cut or (weakly) modified by an ε- or τ-rule during the proof, and both those transformations go through after carrying out the substitution of free variables for terms. ⊣

CLAIM 5.6. *Suppose that π is an LJ^ε-proof of the standard translation of a sequent $\Gamma \vdash D$ in which no strong quantifiers appear. Then there is a LJ^ε-proof π' such that*

1. *π' is cut-free*;
2. *π' has as conclusion the standard translation of $\bar{\Gamma}, \Gamma \vdash \Delta$, where $\bar{\Gamma}$ consists of formulae of one of the following forms*:
   (a) $\forall \vec{x}\,(A(\vec{x}) \to A(\vec{x}))$,
   (b) $\forall \vec{x} \forall \vec{y}\,(A(\vec{x}, \vec{y}) \to A(\vec{x}, f(\vec{x}))$, *or*
   (c) $\forall \vec{x} \forall \vec{y}\,(A(\vec{x}, g(\vec{x})) \to A(\vec{x}, \vec{y}))$,
   where f and g are Skolem functions;
3. *π' has no inner inferences.*

PROOF. The proof is obtained by replacing occurrences of $A(\varepsilon_x A)$ by $A(f_A(x))$, for Skolem functions f_A. Suppose $A(t)$ is a cut formula in π and t is an ε- or τ-term obtained via an inner inference; e.g., $t = \varepsilon_x B(x)$ below:

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\Pi \vdash B(s)}}{\Pi \vdash B(\varepsilon_x B(x))}}{\dfrac{\vdots}{\Pi' \vdash A(\varepsilon_x B(x))}} \qquad \dfrac{\pi_0 \vdots}{A(\varepsilon_x B(x)), \Pi' \vdash E'}}{\Pi' \vdash E'}$$
$$\vdots$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

(Here, let us assume that no other terms appear in B, to simplify notation.) Modify π_0 to form a proof π_1 as follows: the indicated occurrence of $\varepsilon_x B(x)$ must have been either introduced by weakening—in which case we modify the weakening so that a Skolem constant $c_{B(s)}$ is introduced instead—or by an axiom:

$$\dfrac{\dfrac{C(\varepsilon_x B(x)) \vdash C(\varepsilon_x B(x))}{\vdots}}{A(\varepsilon_x B(x)), \Pi' \vdash E'}$$

in which case we replace it by

$$\dfrac{\dfrac{\dfrac{C(c_{B(s)}) \vdash C(c_{B(s)}) \qquad C(\varepsilon_x B(x)) \vdash C(\varepsilon_x B(x))}{C(c_{B(s)}) \to C(\varepsilon_x B(x)), C(c_{B(s)}) \vdash C(\varepsilon_x B(x))}}{\vdots}}{A(c_{B(s)}), \Pi'' \vdash E'}$$

where Π'' has the form $C(c_{B(s)}) \to C(\varepsilon_x B(x)), \Pi'$. (In the general case, $c_{B(s)}$ is of the form $f_{B(s)}(\vec{z})$.) We modify π—say, in the latter case—as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\Pi \vdash B(s)} \quad \dfrac{\vdots}{B(c_{B(s)}) \vdash B(c_{B(s)})}}{B(s) \to B(c_{B(s)}), \Pi \vdash B(c_{B(s)})}}{\vdots}}{B(s) \to B(c_{B(s)}), \Pi' \vdash A(c_{B(s)})} \quad \dfrac{\pi_1 \vdots}{A(c_{B(s)}), \Pi'' \vdash E'}}{C(c_{B(s)}) \to C(\varepsilon_x B(x)), B(s) \to B(c_{B(s)}), \Pi' \vdash E'}$$

$$\vdots$$

Repeat this for every weak ε- and τ-term inferred in each cut formula. We then make a similar modification to the proof for each inner inference. For example,

$$\dfrac{\dfrac{\dfrac{\vdots}{\Pi, \vdash B(s)}}{\Pi, \vdash B(\varepsilon_x A(x))}}{\dfrac{\dfrac{\vdots}{\Pi', \vdash A(\varepsilon_x B(x))}}{\Pi', \vdash A(\varepsilon_x A(x))}}$$

becomes

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\Pi, \vdash B(s)} \qquad B(c_{B(s)}) \vdash B(c_{B(s)})}{B(s) \to B(c_{B(s)}), \Pi, \vdash B(c_{B(s)})}}{\vdots}}{\dfrac{B(s) \to B(c_{B(s)}), \Pi', \vdash A(c_{B(s)})}{B(s) \to B(c_{B(s)}), \Pi', \vdash A(\varepsilon_x A(x))}}$$

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

After this, replace all cuts with $\to L$ rules as in the proof of Theorem 3.3 and extend the proof with one τ-rule applied to each Skolem term and to each ε- and τ-term in each of the introduced implications (in any order). Thus one is left with a proof satisfying all desired conditions. ⊣

To prove the theorem, let π be a proof in $\mathrm{LJ}^{\varepsilon}$ of the standard translation of a sequent

$$\Gamma \vdash D.$$

First, find a proof π_0 in $\mathrm{LJ}^{\varepsilon}$ of the standard translation of $sk(\Gamma) \vdash sk(D)$ as in Proposition 4.5. By the first claim, all axioms in π_0 can be assumed not to contain ε- or τ-terms. Apply the second claim to obtain a cut-free proof π_1 with no inner inferences and with conclusion $\hat{\Gamma}, \Gamma_0 \vdash D_0$, where

1. $\Gamma_0 \vdash D_0$ is the standard translation of the Skolemization of $\Gamma \vdash D$;
2. formulae in $\hat{\Gamma}$ are standard translations of formulae of one of the following forms:
   (a) $\forall \vec{x}\,(A(\vec{x}) \to A(\vec{x}))$,
   (b) $\forall \vec{x} \forall \vec{y}\,(A(\vec{x}, \vec{y}) \to A(\vec{x}, f(\vec{x}))$, or
   (c) $\forall \vec{x} \forall \vec{y}\,(A(\vec{x}, g(\vec{x})) \to A(\vec{x}, \vec{y}))$,
   where f and g are Skolem functions.

Since the proof has no strong terms in its conclusion, is cut-free, and has no inner inferences, it can be directly translated into a proof in LJ. By the argument in Proposition 4.8, there is an LJ^{++}-proof of the deskolemized sequent

$$\bar{\Gamma}, \Gamma \vdash D,$$

where $\bar{\Gamma}$ consists of formulae of one of the following forms:

1. $\forall \vec{x}\,(A(\vec{x}) \to A(\vec{x}))$,
2. $\forall \vec{x}\, \exists \vec{z} \forall \vec{y}\,(A(\vec{x}, \vec{y}) \to A(\vec{x}, \vec{z})$, or
3. $\forall \vec{x}\, \exists \vec{z} \forall \vec{y}\,(A(\vec{x}, \vec{z}) \to A(\vec{x}, \vec{y}))$.

By Proposition 4.4, these formulae are derivable in LJ^{++}, so they can be cut to obtain a proof of

$$\Gamma \vdash D,$$

as desired. ⊣

§6. Concluding remarks. Many related questions remain open. Perhaps one could investigate more general notions of proof and their formalizations. For example, one could explore the addition of different unsound rules to proof systems. Here, we gave two examples, but others are possible. It is unknown exactly how faster the calculus LK^{+} really is—the weakening of the usual characteristic-variable condition into substitutability, weak regularity, and the side-variable condition can perhaps be exploited further.

In particular, we do not know whether LK^{+} is any (or significantly) slower than LK^{++}. We also do not know precisely the relation (complexity-wise) between them and either of $\mathrm{LK}_{\mathtt{shift}}$ or $\mathrm{LK}^{\varepsilon}$.

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press

Acknowledgments. We would like to thank the anonymous referee for several comments that certainly improved the article. We also thank Stefan Hetzl and Michel Parigot for fruitful conversations. This work was partially supported by FWF grants I-2671-N35 and W1255-N23.

REFERENCES

[1] M. Baaz and S. Hetzl, *On the non-confluence of cut-elimination*, this Journal, vol. 76 (2011), pp. 313–340.

[2] M. Baaz, S. Hetzl, and D. Weller, *On the complexity of proof deskolemization*, this Journal, vol. 77 (2012), pp. 669–686.

[3] M. Baaz and A. Leitsch, *On skolemization and proof complexity*. ***Fundamenta Informaticae***, vol. 20 (1994), pp. 353–379.

[4] ———, *Cut normal forms and proof complexity*. ***Annals of Pure and Applied Logic***, vol. 97 (1999), pp. 127–177.

[5] D. Hilbert and P. Bernays, ***Grundlagen der Mathematik II***, Springer-Verlag, Berlin, 1970.

[6] G. Moser and R. Zach, *The epsilon calculus and Herbrand complexity*. ***Studia Logica***, vol. 82 (2006), no. 1, pp. 133–155.

[7] V. P. Orevkov, *Lower bounds for increasing complexity of derivations after cut elimination* (*in Russian*). ***Journal of Soviet Mathematics***, vol. 20 (1982), pp. 2337–2350.

[8] R. Statman, *Lower bounds on Herbrand's theorem*. ***Proceedings of the American Mathematical Society***, vol. 75 (1979), pp. 104–107.

[9] G. Takeuti, ***Proof Theory***, North-Holland, Amsterdam, 1980.

INSTITUTE OF DISCRETE MATHEMATICS AND GEOMETRY
VIENNA UNIVERSITY OF TECHNOLOGY
WIEDNER HAUPTSTRAßE 8–10, 1040 VIENNA, AUSTRIA
E-mail: aguilera@logic.at
E-mail: baaz@logic.at

https://doi.org/10.1017/jsl.2018.51 Published online by Cambridge University Press