THE NUMBER OF AXIOMS

J. P. AGUILERA, M. BAAZ, AND J. BYDŽOVSKÝ

Abstract. We derive results concerning the least possible number of distinct axioms in proofs in classical and intuitionistic sequent calculi and other related systems. In particular, we show that there is no recursive bound on the least possible number of distinct axioms in a proof of a provable sequent in terms of the sequent's length, and that there is no elementary bound on the least possible number of axioms in a cut-free proof of a provable sequent in terms of the least possible length of an arbitrary proof thereof, strengthening a classical result due to Orevkov and Statman.

1. Introduction

In this article, we explore lower bounds on the number of axioms needed to prove theorems. We deal with first-order logic, formalized as a version of the *sequent calculus* LK introduced by Gentzen [3] (see also Takeuti [8] for additional background). A *sequent* is an expression of the form

$$\Gamma \vdash \Delta \tag{1}$$

where Γ and Δ are finite *multisets* of formulae. The interpretation of (1) is "if all formulae in Γ hold, then some formula among Δ holds". In LK, one starts with axioms and infers other sequents through various *rules of inference*. We measure the length of a proof by the number of sequents that appear in it; we measure the length of a sequent by the number of symbols in it. It is well known that one cannot give a recursive bound on the least possible length of a proof of a provable sequent S in terms of the length of S itself. Below, we prove the following strengthening:

Theorem 1.1. *There is no recursive bound on the least possible number of distinct axioms in an* LK*-proof of a sequent in terms of its length.*

In the statement of the theorem, "axiom" means "logical axiom". We do not consider two occurrences of the same axiom $A(a) \vdash A(a)$ as "distinct," but we do consider as distinct different instances of the same axiom, such as $A(a) \vdash A(a)$ and $A(b) \vdash A(b)$. Theorem 1.1 says that as one considers longer sequents, their minimal proofs not only become "longer," but also "wider," and moreover so in a way that cannot be accounted for by the repetition of axioms.

Intuitionistic logic can be formalized as one of many variants of Gentzen's LJ, which is obtained from LK by adding the restriction that all sequents $\Gamma \vdash \Delta$ contain at most one formula on the right-hand side. All our arguments below apply to intuitionistic logic and to many other related systems. In particular, we have:

Theorem 1.2. *There is no recursive bound on the least possible number of distinct axioms in an* LJ*-proof of a sequent in terms of its length.*

Date: March 21, 2022 (Compiled).
2010 *Mathematics Subject Classification.* 03F03, 03F20 (Primary); 03B10, 03F07 (Secondary).

Among the usual inferences in sequent calculi figures the *cut* rule:

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

Gentzen's *Cut-Elimination Theorem* says that the cut rule is redundant, however. Cut-free proofs are useful because they have the *subformula property*: in a proof of a sequent S with no instances of the cut rule, one only finds formulae which are substitution instances of subformulae of formulae in S. This is a desirable property for automated proof search, and other applications. The main tool in the proof of Theorem 1.1 is the following lower bound on the number of axioms in cut-free proofs:

Theorem 1.3. *Let S be a provable* LK*-sequent. Denote by m the minimal length of a cut-free* LK*-proof of S and by α the minimal number of* distinct *axioms in a cut-free* LK*-proof of S. Then*

$$\sqrt[|S|^3]{\frac{1}{2^{|S|^4+1}}\log_2(m)} \leq \alpha.$$

Theorem 1.3 is proved in Section 3. Section 4 contains applications of Theorem 1.3 and, in particular, the proofs of Theorems 1.1 and 1.2. Before moving on, we mention another application of Theorem 1.3.

Recall that cut elimination has a high computational cost. A function $f : \mathbb{N} \to \mathbb{N}$ is *elementarily bounded* if it is bounded by a function of the form

$$x \mapsto 2^{2^{\cdot^{\cdot^{\cdot^{2^x}}}}}.$$

An algorithm is *elementary* if it runs in an amount of time which is elementarily bounded. A classical theorem due independently to Orevkov [5] and Statman [7] states that there can be no elementary cut-elimination algorithm for first-order logic. By inspecting Schütte's proof of Gentzen's cut-elimination theorem (see e.g., Schwichtenberg [6]), one sees that this result is optimal, in the sense that the cut-elimination theorem requires computations as simple as possible among non-elementary classes. More precisely, it is easily shown that the cut-elimination theorem is equivalent to the totality of the superexponential function (which maps a natural number n to the result of applying the exponentiation function $x \mapsto 2^x$ n times) over Elementary Arithmetic (EA) (see e.g. Beklemishev [2] for more on relevant subsystems of arithmetic); however, this leaves open the possibility of strengthening the result in other directions; namely, Orevkov and Statman's proofs show that there is a sequence of first-order sequents the nth of which has a proof of length $\mathcal{O}(n)$, but whose shortest cut-free proofs have lengths which cannot be elementarily bounded. We strengthen this result by showing that those cut-free proofs must necessarily have non-elementarily many distinct axioms.

Theorem 1.4. *There is no elementary bound on the least possible number of* distinct *axioms of a cut-free* LK*-proof of a sequent in terms of the least possible length of an* LK*-proof of the same sequent.*

With little extra work, one can adapt these theorems to proof systems with equality axioms.

2. Preliminaries

We work in first-order logic formalized as a calculus of sequents as in Takeuti [8]. For convenience and definiteness, we recall the definitions which we will require. We work in the version of Gentzen's calculus LK with atomic axioms of the form $A \vdash A$ and the following inference rules:

Propositional rules:

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta} \wedge_1\text{−}left \qquad \frac{\Gamma \vdash \Delta, A}{\Gamma \vdash \Delta, A \vee B} \vee_1\text{−}right$$

$$\frac{\Gamma, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta} \wedge_2\text{−}left \qquad \frac{\Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \vee B} \vee_2\text{−}right$$

$$\frac{\Gamma_1, A \vdash \Delta_1 \quad \Gamma_2, B \vdash \Delta_2}{\Gamma_1, \Gamma_2, A \vee B \vdash \Delta_1, \Delta_2} \vee\text{−}left \qquad \frac{\Gamma_1 \vdash \Delta_1, A \quad \Gamma_2 \vdash \Delta_2, B}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2, A \wedge B} \wedge\text{−}right$$

$$\frac{\Gamma \vdash \Delta, A}{\Gamma, \neg A \vdash \Delta} \neg\text{−}left \qquad \frac{\Gamma, A \vdash \Delta}{\Gamma \vdash \Delta, \neg A} \neg\text{−}right$$

$$\frac{\Gamma, A \vdash \Delta, B}{\Gamma \vdash \Delta, A \to B} \to_1 -right \qquad \frac{\Gamma_1 \vdash \Delta_1, A \quad \Gamma_2, B \vdash \Delta_2}{\Gamma_1, \Gamma_2, A \to B \vdash \Delta_1, \Delta_2} \to -left$$

Structural rules:

$$\frac{\Gamma \vdash \Delta}{\Gamma, A \vdash \Delta} \; weakening\text{−}left \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, A} \; weakening\text{−}right$$

$$\frac{\Gamma, A, A \vdash \Delta}{\Gamma, A \vdash \Delta} \; contraction\text{−}left \qquad \frac{\Gamma \vdash \Delta, A, A}{\Gamma \vdash \Delta, A} \; contraction\text{−}right$$

$$\frac{\Gamma_1 \vdash \Delta_1, A \quad A, \Gamma_2 \vdash \Delta_2}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2} \; Cut$$

Quantifier rules:

$$\frac{\Gamma, A(t) \vdash \Delta}{\Gamma, \forall x\, A(x) \vdash \Delta} \forall\text{−}left \qquad \frac{\Gamma \vdash \Delta, A(t)}{\Gamma \vdash \Delta, \exists x\, A(x)} \exists\text{−}right$$

$$\frac{\Gamma, A(a) \vdash \Delta}{\Gamma, \exists x\, A(x) \vdash \Delta} \exists\text{−}left \qquad \frac{\Gamma \vdash \Delta, A(a)}{\Gamma \vdash \Delta, \forall x\, A(x)} \forall\text{−}right$$

where (i) the bound variable x does not appear in $A(a)$ (resp. $A(t)$) and[1] (ii) in the case of inferences $\forall$−right and $\exists$−left, the variable a does not appear in the lower sequent of the inference. In this case, a is called the *characteristic variable* of the inference and the inference is called a *strong* quantifier inference. The other two quantifier inferences are called *weak*; in these, t is assumed to be a term, i.e., an expression built up of function symbols applied to constant symbols and *free* variables.[2]

An *occurrence* of a formula A (or term, variable, etc.) in a sequent (or proof, etc.) is the formula A together with the position it occupies in the proof (sequent,

[1]This condition precludes expressions such as $\exists x \forall x\, A(x, x)$ from occurring in sequent calculi.

[2]In particular, recall that if x is a bound variable, then expressions such as $f(x)$ are not terms, but rather only *semiterms*.

etc.). Given an occurrence of a term t in a sequent S, we call the corresponding occurrences of t in formulae below S the *successors* of this occurrence of t; the *predecessors* of this occurrence of t are defined similarly.

For the logical rules and quantifier rules, the indicated formulae are called respectively the *auxiliary* and *principal* formulae of the inferences. If a is the characteristic variable of a strong quantifier inference with principal formula A, we also say that a is the characteristic variable of A and that A is a principal formula of a. By extension, if x is the bound variable used to quantify an occurrence of a, we say that x is the principal variable of this occurrence of a, and of the inference. We also say that the corresponding occurrence of a is *quantified* by (or in) that inference. Similarly, given a weak quantifier inference with principal variable x and auxiliary term t, we say that the corresponding occurrence of t is *quantified* in that inference, but we do not say this of any occurrence of a proper subterm of t.

An LK-*proof* of a sequent S is a sequence of sequents $S_0, \dots, S_m = S$ such that for each $k \leq m$, S_k is either an axiom or is derived from some S_i, S_j with $i, j < k$ by an application of an inference rule above. We remark that, since sequents have been defined as multisets, there is no need for exchange rules; this is merely for convenience.

Every LK-proof π corresponds to a directed acyclic graph where the vertices are labelled by sequents from π and the arrows lead from the hypotheses of an inference to its conclusion. If the corresponding graph is a tree we call the proof *tree-like*.

For a formula, sequent, or term X, we write $|X|$ for the number of symbols appearing in X counting each occurrence of the symbol individually. For a proof π, we write $|\pi|$ for its length.

3. Bounds on the number of axioms

Here, we restate and prove Theorem 1.3. Although it is stated for LK, the argument is very general and applies to many other proof systems (see Section 4 below).

Theorem 3.1. *Let S be a provable* LK*-sequent. Denote by m the minimal length of a cut-free* LK*-proof of S and by α the minimal number of* distinct *axioms in a cut-free* LK*-proof of S. Then*

$$\sqrt[|S|^3]{\frac{1}{2^{|S|^4+1}}\log_2(m)} \leq \alpha.$$

To prove the theorem, we will need some more definitions. Below, we speak of *subterms* of a term – these are defined as expected. We remark that subterms of a term are themselves terms and hence contain no bound variables.

Definition 3.2. Let π be an LK-proof and t be a term.

(1) We say that an occurrence of t in π is *inessential* if one of its successors is quantified in π and t is not a subterm of any term which appears in an axiom of π.
(2) We say that a term t in π is *inessential* if all its occurrences are inessential.

Note that inessential occurrences of terms can only be introduced by weakenings.

Definition 3.3. *Let c be a constant symbol or variable. A proof π is* c-efficient *if the following hold:*

(1) *c does not appear in an axiom in π nor in the end-sequent of π;*
(2) *suppose that t is a term in π one of whose occurrences is inessential in π, then either t is equal to c, or t is a free variable and none of its occurrences are weakly quantified in π.*

Lemma 3.4. *Let π be an* LK*-proof and let c be a constant symbol or a variable not occurring in π. Then, there is an* LK*-proof σ such that*

(1) *σ has the same length, end-sequent, and axioms as π;*
(2) *σ is c-efficient.*

Proof. Let π and c be as in the statement of the lemma. We obtain σ as follows: suppose there is an inessential occurrence of a term t in π and a successor of this occurrence is weakly quantified in π; then, replace this occurrence by c.

Thus, σ differs from π only on occurrences of inessential terms and thus has the same axioms and end-sequent, as well as the same length. We observe that

(1) σ and π differ only on occurrences of terms which are weakly quantified and their predecessors; and
(2) if an occurrence of a term t is replaced by c in passing from π to σ, then all predecessors of this occurrence are also replaced.

It follows that the characteristic-variable condition holds of σ; hence, σ is a proof and is as desired. □

Remark 3.5. In the preceding argument, σ is *not* a substitution instance of π, since different occurrences of terms in π might correspond to different terms in σ.

Lemma 3.6. *Suppose that π is a cut-free c-efficient* LK*-proof and let t be a subterm of some term in π. Then t is one of the following:*

(1) *a subterm of an axiom;*
(2) *a subterm of a term in the end-sequent;*
(3) *c;*
(4) *an inessential free variable.*

Proof. If t is not in the first two categories, then t must be a subterm of some term s with an inessential occurrence in π. By c-efficiency, s must either be equal to c or to a free variable. Thus, t must either be equal to c or to an inessential free variable. □

Definition 3.7. Let a cut-free LK-proof π and an occurrence of a formula B in a sequent S in a proof π be given. Consider the the unique sequence of pairs $(B_0, S_0), \ldots, (B_n, S_n)$ in π such that

(1) $B_0 = B$, $S_0 = S$;
(2) S_n is the end-sequent of π;
(3) if $i < n$, then S_{i+1} is the sequent immediately below S_i in π;
(4) if $i < n$ and B_i is an auxiliary formula of the inference that produces S_{i+1} from S_i, then B_{i+1} is the principal formula;
(5) if $i < n$ and B_i is not an auxiliary formula of the inference that produces S_{i+1} from S_i, then B_{i+1} is the occurrence of B_i in S_{i+1}.

The *trace* of B is the set $\{B_0, \ldots, B_n\}$ (without repetitions) together with the position B_n occupies in the end-sequent of π. We refer to B_n, together with the position it occupies in π, as the *conclusion* of the trace.

Remark 3.8. We defined the trace of a formula as a set $\{B_0, \dots, B_n\}$. However, we could also have defined it as an ordered set, and the order is uniquely determined by the elements of $\{B_0, \dots, B_n\}$.

Remark 3.9. Because the trace includes the information of where the conclusion appears in π, it is possible that two different occurrences of the same formula in π have different traces which nonetheless consist of the same formulae. Similarly, due to contractions, it is possible that two different occurrences of the same formula, neither an ancestor of the other, have the same trace.

The trace of a formula essentially records its history through the proof up to the end-sequent, as well as its position therein.

Definition 3.10. Suppose $\star$ is a new symbol not appearing elsewhere.

Given a proof π and a formula D occurring in it. We define the *$\star$-abstraction* of D, $D^\star$, as the expression obtained from D by simultaneously substituting the symbol $\star$ for all inessential free variables in π.

We define the *$\star$-trace* of an occurrence B of a formula in π as the result of replacing each formula in the trace of B by its $\star$-abstraction.

The next lemma is the main ingredient for the proof of Theorem 3.1.

Lemma 3.11. *Let π be a cut-free* LK*-proof of a sequent S and let α denote the number of distinct axioms in π. Then there is a cut-free tree-like* LK*-proof of S with at most $\alpha^{|S|^3} \cdot 2^{|S|^4}$ different terms.*

Consequently, S is provable by a cut-free LK*-proof of length less than $2^{\alpha^{|S|^3} \cdot 2^{|S|^4+1}}$.*

Proof. We can assume that π is tree-like since otherwise we can transform it into a tree-like form without adding any new (different) instances of axioms. Without loss of generality, we assume that π is 0-efficient.

Now let us consider an arbitrary inessential variable a and one of its occurrences in π.

Let $\forall x B(x)$ be the principal formula of the inference in which a successor of the selected occurrence of a is quantified, and say that this inference has the form

$$\frac{\Gamma_a \vdash \Delta_a, B(a)}{\Gamma_a \vdash \Delta_a, \forall x\, B(x)}\ .$$

If so, we replace the occurrence of the variable a in this inference, as well as all its predecessors, by a new variable v_T, where T is the $\star$-trace of $B(a)$.

We repeat these transformations for all occurrences of inessential variables and denote the resulting sequence of sequents by π'. The main observation to make at this point is that π' might not be a proof because the characteristic-variable condition need not be satisfied by all strong quantifier inferences. However, this is the only potential obstacle, as all other inferences in π' are valid. In particular, it should be remarked that all instances of contraction in π remain valid in π', for when a variable a is replaced by one of the form v_T, all its predecessors are too. Before we can modify π' to obtain a correct proof some observations are in order.

Claim 3.12. *The end-sequent of π' is S.*

Proof. This is immediate from the construction, because π and π' only differ in terms that appear neither in axioms nor in the end-sequent. □

Claim 3.13. *Assume $A(a)$ and $B(b)$ are occurrences of characteristic formulae with the same principal variable and with characteristic variables a and b, respectively. Let $T_{A(a)}$ and $T_{B(b)}$ be the traces of $A(a)$ and of $B(b)$ and suppose that $T_{A(a)} \subseteq T_{B(b)}$ and that the traces have the same conclusion. Then $A(x)$ and $B(x)$ are the same formula.*

Proof. This is immediate from the definition and the fact that a formula cannot contain two separate nested quantifiers applied to the same bound variable (this follows from condition (i) in the definition of the quantifier rules in LK on page 3). □

Claim 3.14. *Suppose a is a free variable of the form v_T and that π' contains the inferences*

$$\frac{\Pi_1 \vdash \Lambda_1, B_1(a, s_1, \ldots, s_n)}{\Pi_1 \vdash \Lambda_1, \forall x_i\, B_1(x_i, s_1, \ldots, s_n)}$$

and

$$\frac{\Pi_2 \vdash \Lambda_2, B_2(a, t_1, \ldots, t_m)}{\Pi_2 \vdash \Lambda_2, \forall x_j\, B_2(x_j, t_1, \ldots, t_m)}$$

Then $B_1(a, s_1, \ldots, s_n) = B_2(a, t_1, \ldots, t_m)$.

Proof. By hypothesis, a is of the form v_T for some $\star$-trace

$$T = \{D_0, D_1, \ldots, D_k.\}$$

Because of the way the substitution of inessential variables by v_T was carried out, D_0 can be obtained by substituting $\star$ for inessential variables in B_1, or in B_2. Thus B_1 and B_2 are substitution instances of each other and can only differ in inessential variables. To show $B_1 = B_2$ assume b_1 and b_2 are inessential free variables on the same position in B_1 and B_2, respectively.

Observe that b_1 and b_2 will become the same bounded variable in D_k. By choice of π', they are strongly quantified in π'. Thus, b_1 and b_2 are characteristic variables of π' of the forms v_{T_1} and v_{T_2} for some traces T_1 and T_2.

T_1 and T_2 are both end-segments of T, so one of them must extend the other. By Claim 3.13, b_1 and b_2 have the same principal formula. The $\star$-trace T' of this principal formula has to be of the form

$$T' = D_\ell, D_{\ell+1}, \ldots, D_k$$

for some $0 < \ell \leq k$; therefore,

$$b_1 = b_2 = v_{T'},$$

as desired. This proves the claim. □

The analogue of Claim 3.14 for free variables of the form v_T which are strongly quantified on the left side is proved by a similar argument.

We now modify π' so as to obtain a proof. Consider a failure of the characteristic variable condition, say

$$(*)\ \frac{\Pi \vdash \Lambda, D(a)}{\Pi \vdash \Lambda, \forall x\, D(x)}$$

where a appears in $\Pi \vdash \Gamma, \forall x\, D(x)$. Since π was originally a proof, the only possibility is that a be of the form v_T and all its occurrences be strongly quantified in π'. By Claim 3.14, all occurrences of a in π' have the same principal formula. Thus, a branch in π' in which a appears has the following form:

$$\dfrac{\dfrac{\dfrac{\vdots}{\Pi \vdash \Lambda, D(a)}}{\Pi \vdash \Lambda, \forall x\, D(x)}\,(*)}{\dfrac{\dfrac{\vdots}{\Pi' \vdash \Lambda', D(a)}}{\Pi' \vdash \Lambda', \forall x\, D(x)}\,(**)}$$
$$\dfrac{\vdots}{\Gamma \vdash \Delta}$$

where a does not appear in $\Pi' \vdash \Lambda', \forall x\, D(x)$ or below. We modify π' by

(1) omitting all the inferences with principal formula $\forall x\, D(x)$ and auxiliary formula $D(a)$ (such as $(*)$) except for the last one, as well as any inference between (*) and (**) whose auxiliary formula is a successor of a formula that has been omitted;
(2) contracting the occurrences of $D(a)$ before $(**)$;
(3) weakening after $(**)$ to add any and all formulae missing from $\Pi' \vdash \Lambda'$.

The resulting derivation has the following form:

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\Pi \vdash \Lambda, D(a)}}{\dfrac{\vdots}{\Pi'' \vdash \Lambda'', D(a), D(a), \ldots, D(a)}}}{\dfrac{\dfrac{\Pi'' \vdash \Lambda'', D(a)}{\Pi'' \vdash \Lambda'', \forall x\, D(x)}\,(**)}{\Pi' \vdash \Lambda', \forall x\, D(x)}\,\text{weakening}}}{\dfrac{\vdots}{\Gamma \vdash \Delta}}$$

where $\Pi'' \vdash \Delta''$ is a subsequent of $\Pi' \vdash \Delta'$. Repeating this procedure for each failure of the characteristic-variable condition on each branch in the derivation tree produces a cut-free LK-proof σ of $\Gamma \vdash \Delta$ with $|\sigma| \leq |\pi|$.

We now estimate the number of terms in σ. Recall that α denotes the number of distinct axioms in π.

Claim 3.15. *The number of different terms in σ is at most $\alpha^{|S|^3} \cdot 2^{|S|^4}$.*

Proof. First observe that the terms introduced by weakening after $(**)$ are free variables of the form v_T or appear in an axiom of π or in the end-sequent, or equal to the constant 0. Now assume t is an occurrence of a maximal term in π'. Following the successors of t on the way down to the end sequent we arrive at some term $t'(x_0, x_1, \ldots, x_k)$ where $x_0, x_1, \ldots, x_k$ might be bounded variables and

so $t = t'(a_0, a_1, \dots, a_k)$ for some terms $a_0, a_1, \dots, a_k$. Moreover the number of such terms t' is bounded by $|S|$ and the terms $a_0, a_1, \dots, a_k$ can be 0, v_T for some $\star$-trace T, or a subterm of a term s which occurs in an axiom and an occurrence of s is quantified in π'. Because the axioms in π' are atomic, the number of terms of the third kind mentioned is:

$$\alpha \cdot (\text{maximal arity of an atomic formula in } S) \cdot (\text{maximal arity of a term in } S)$$

which is at most $\alpha \cdot |S|^2$. Thus the number of terms in π' is at most

$$|S| \cdot (1 + (\text{number of } \star\text{-traces}) + \alpha \cdot |S|^2)^{(\text{maximal arity of a term in } S)}$$

where the exponent is bounded by $|S| - 1$.

To finish we need to bound the number of $\star$-traces $T = \{D_0, D_1, \dots, D_n\}$ in π'. First note that similarly as in the previous paragraph, the number of terms that can appear in a $\star$-trace is at most $(1+1+\alpha \cdot |S|^2)^{(\text{maximal arity of a term in } S)}$ and the length n of the trace is at most the number of logical connectives and quantifiers in D_n thus bounded by $|S| - 1$. Moreover D_n and D_0 determines the rest of the formulas in T up to terms. Finally each term in T is a predecessor of a term in D_n and each term in D_n has at most $|S|$ predecessor. In particular, for a fixed D_n with its position in S and D_0 there is at most

$$(\text{number of terms in } D_n) \cdot (2 + \alpha \cdot |S|^2)^{|S|(|S|-1)}$$

possible combinations of terms that could appear in T. This gives at most

$$|S|^4 \cdot (2 + \alpha \cdot |S|^2)^{|S|(|S|-1)}$$

$\star$-traces originating from π'. Altogether, the number of different terms in π' is at most

$$|S| \cdot (1 + |S|^4 \cdot (2 + \alpha|S|^2)^{|S|(|S|-1)} + \alpha|S|^2)^{|S|-1} \leq |S| \cdot (|S|^4(2 + \alpha|S|^2)^{|S|^2})^{|S|-1}$$

and this is

$$\leq \alpha^{|S|^3} \cdot 2^{|S|^4}$$

since we can assume that $|S| \geq 3$. □

To prove the "consequently" part of the lemma, we note that since σ is cut-free, $3 \leq |S|$, and the number of different terms in σ is at most $\alpha^{|S|^3} \cdot 2^{|S|^4}$, it follows that there are at most

$$|S|^{\alpha^{|S|^3} \cdot 2^{|S|^4} + |S|} \leq 2^{\alpha^{|S|^3} \cdot 2^{|S|^4} + 1}$$

different sequents in σ. Since we can remove duplicate occurrences of a sequent from σ by reducing it to a proof which is not necessarily tree-like, we obtain a cut-free proof of S with at most $2^{\alpha^{|S|^3} \cdot 2^{|S|^4} + 1}$ many sequents. This completes the proof of the lemma. □

Using the lemma we can finally prove Theorem 3.1.

Proof of Theorem 3.1. Assume towards a contradiction that σ is a cut-free LK-proof of S of length m such that, letting α be the number of axioms of σ, we have

$$\alpha < \sqrt[|S|^3]{\frac{1}{2^{|S|^4+1}} \log_2(m)},$$

so

$$2^{\alpha^{|S|^3} \cdot 2^{|S|^4} + 1} < m.$$

By Lemma 3.11 applied to σ, there is an LK-proof σ' of S with

$$|\sigma'| \leq 2^{\alpha^{|S|^3} \cdot 2^{|S|^4+1}} < m.$$

contradicting the assumption on the minimality of m. □

Since an arbitrary LK-proof π of length less than m can be transformed without adding cut inferences and new sequents into a *tree-like* LK-proof π' with the same end-sequent and $|\pi'| \leq 2^{2|\pi|}$, we obtain the following corollary:

Corollary 3.16. *Let S be a provable* LK*-sequent. Denote by m the minimal length of a a tree-like cut-free* LK*-proof of S and by α the minimal number of distinct axioms in a tree-like cut-free* LK*-proof of S. Then,*

$$\sqrt[|S|^3]{\frac{(\log_2(\log_2(m)) - 1)}{2^{|S|^4+1}}} \leq \alpha.$$

4. Applications, Generalizations, and Further Remarks

4.1. **Applications.** We first derive Theorem 1.4, which we restate for convenience:

Theorem 4.1. *There is no elementary bound on the least possible number of distinct axioms of a cut-free* LK*-proof of a sequent in terms of the least possible length of an arbitrary* LK*-proof of the same sequent.*

Proof. Let $\{\pi_n : n \in \mathbb{N}\}$ be the sequences of proofs constructed by Orevkov [5] or Statman [7]. These proofs have length polynomial in n, but there is no elementary sequence of cut-free LK-proofs with the same end-sequents. The theorem follows from applying Corollary 3.16 to this sequence. □

As another application, we prove the following theorem, motivated by a question of D. J. D. Hughes (private communication). The calculus LK^{++} is as in [1].

Theorem 4.2. *There is no elementary function bounding the least number of different axioms in a cut-free* LK*-proof of a sequent in terms of the least number of axioms in a cut-free* LK^{++}*-proof of the same sequent.*

Proof. By the proof of [1, Theorem 3.3], there is a family $\{S_n : n \in \mathbb{N}\}$ of sequents with polynomial cut-free LK^{++}-proofs, but with no elementarily bounded LK-proofs (see specifically, equation (8) in [1]). By Corollary 3.16, the number of axioms in any sequence of LK-proofs of these sequents is not elementarily bounded. □

A similar result is obtained for the calculus LK^{+} from [1]. We now proceed to the proof of Theorem 1.1, which we also restate:

Theorem 4.3. *There is no recursive bound on the least possible number of distinct axioms in an* LK*-proof of a sequent in terms of its length.*

Proof. From an arbitrary proof of a sequent, Schütte's cut-elimination algorithm (see e.g., Schwichtenberg [6]) provides a cut-free proof of the same sequent whose length is primitive-recursively bounded. Thus, it suffices to show that there is no recursive bound on the least possible number of distinct axioms in a cut-free LK-proof of a sequent in terms of its length.

Suppose towards a contradiction there is a recursive function

$$f : \mathbb{N} \to \mathbb{N}$$

such that every sequent of length n has a cut-free proof with at most $f(n)$ distinct axioms. By Theorem 3.1, we could replace "distinct axioms" by "length of a proof". By a theorem of Krajíček and Pudlák [4, Theorem 2.5], the minimal number of symbols in a cut-free proof of a sequent S is bounded elementarily by the minimal number of inference steps in a cut-free proof of S. Moreover the number of inference steps and sequents in a proof are polynomially related. Therefore, if such an f existed, one would have a recursive bound on the number of symbols in a cut-free proof of a sequent in terms of its length, which would contradict the undecidability of first-order logic. □

4.2. **Equality Axioms.** One might wonder whether the results proved thus far apply to systems with equality axioms. We mention the main result in this context. Other, similar, generalizations we leave to the reader's imagination.

Definition 4.4. $\mathrm{LK}^{=}$ *is the extension of* LK *by all equality axioms of one of the following forms:*

$$\begin{aligned} &\vdash a = a \\ b = a &\vdash a = b; \\ a = b, b = c &\vdash a = c; \\ a_1 = b_1, \ldots, a_n = b_n &\vdash f(a_1, \ldots, a_n) = f(b_1, \ldots, b_n); \\ a_1 = b_1, \ldots, a_n = b_n &\vdash A(a_1, \ldots, a_n) \to A(b_1, \ldots, b_n); \end{aligned}$$

where f is a function symbol and A is an atomic formula.

Theorem 4.5. *Let* $\mathrm{LK}^{=}$ *be* LK *augmented with equality axioms. Then, there is no recursive bound on the least possible number of distinct axioms in an* $\mathrm{LK}^{=}$*-proof of a sequent in terms of its length.*

Proof. Recall that every sequent which is provable $\mathrm{LK}^{=}$ has a proof which is cut-free, except possibly for "inessential" cuts – those applied to formulas of the form $t = s$. Such a proof is obtained by applying any of the usual cut-elimination algorithms. The only case which the algorithm does not cover is that in which the cut formula is introduced directly by an equality axiom, and in fact only those formulas of the form $t = s$ which come from an equality axiom (see e.g., Takeuti [8, Theorem 7.6]).

Let us say that an $\mathrm{LK}^{=}$-proof is *almost cut-free* if all its cut formulas are atomic of the form $s = t$ and for each instance of a cut rule, at least one of the auxiliary (cut) formulas has an equality axiom as a predecessor. Almost cut-free proofs enjoy a weak subformula property whereby every formula is either a subformula of the end-sequent, or it is an instance of equality between two terms which appear in an equality axiom. To prove the theorem, it suffices to show that there is no recursive bound on the least possible number of distinct axioms in an almost cut-free $\mathrm{LK}^{=}$-proof of a sequent in terms of its length.

The idea for the proof is to prove an analogue of Lemma 3.11 for $\mathrm{LK}^{=}$:

Lemma 4.6. *There is a recursive function f with the following property: Suppose π be an almost cut-free* $\mathrm{LK}^{=}$*-proof of a sequent S and α is the number of distinct axioms in π. Then, there is an almost cut-free* $\mathrm{LK}^{=}$*proof of S with at most $f(\alpha)$ different terms.*

Proof. This is done by essentially the same argument of Lemma 3.11. Like in Lemma 3.11, we may assume that we are dealing with a c-efficient proof for some new constant symbol c; the key point here is that all terms in cut formulas appear in axioms. Hence, following the argument of Lemma 3.11, we can replace all the inessential bound variables by variables v_T indexed by $\star$-traces and from this obtain a proof with a number of terms bounded recursively in terms of α. □

From the claim, we can prove the theorem: Suppose it is false, so there is a recursive bound on the number of axioms in an almost cut-free $\mathrm{LK}^{=}$-proof in terms of the length of the end-sequent. Then by the claim there is a recursive bound on the least possible number of terms in an $\mathrm{LK}^{=}$-proof of that sequent. Arguing as in Theorem 3.1, one can recursively bound the smallest possible number of steps in an almost cut-free $\mathrm{LK}^{=}$-proof of the sequent, and thus the number of symbols, which is impossible. □

4.3. **Intuitionistic logic.** The proof of Theorem 4.1 could be made simpler by first Skolemizing the end-sequent and thus reducing the need to rename strongly quantified variables. However, the present argument can be used for calculi which do not admit Skolemization.

Theorem 4.7. *There is no elementary bound on the least possible number of distinct axioms of a cut-free* LJ*-proof of a sequent in terms of the least possible length of an arbitrary* LJ*-proof of the same sequent.*

Proof. The proof is essentially the same as that of Theorem 4.1. The only thing to mention is that the sequence $\{\pi_n : n \in \mathbb{N}\}$ can also be constructed for LJ, by making use of the sequence for LK, together with the double-negation translation. □

Theorem 4.8. *There is no recursive bound on the least possible number of distinct axioms in an* LJ*-proof of a sequent in terms of its length.*

Proof. Again, this proof is essentially the same as that of Theorem 4.3. The only points to mention are that Schütte's cut-elimination algorithm produces LJ-proofs when applied to LJ-proofs and that the main tool in the proof of Krajíček and Pudlák [4, Theorem 2.5] is a unification argument, which also applies to LJ. □

4.4. **Further remarks.** Other calculi to which the argument of Theorem 4.1 applies we leave to the reader's imagination, but we do mention that two properties of LK are essential for the proof:

(1) LK has full access to weakenings and contractions;
(2) the weak quantifier rules of LK impose no restrictions on the term structure of the quantified terms.

Consequently, the proof does not apply to proof systems that do not share these properties, such as those for linear logic. This raises two questions:

Question 4.9. *Is there an elementary bound on the least possible number of distinct axioms of a cut-free proof of a sequent in linear logic in terms of the least possible length of an arbitrary proof of the same sequent in linear logic?*

Question 4.10. *Is there a recursive bound on the least possible number of distinct axioms of a proof of a sequent in linear logic in terms of its length?*

Finally, we mention that while the sequents considered in this article are given by pairs of multisets, this was merely for convenience. A similar result holds for proof systems with exchanges, though the bounds need to be adjusted accordingly.

It is common in classical logic to gauge the complexity of a proof in terms of its Herbrand sequent. The procedure from Theorem 3.1 can be viewed as a form of "term normalization" applied to a proof, and the set of terms thus produced can be thought of as an abstract Herbrand set of terms assigned to a proof. Hence, Theorem 3.1 in a way extends this notion of complexity to non-classical systems to which it applies, such as intuitionistic logic.

Acknowledgements. The authors would like to thank D. J. D. Hughes for his motivating question and the anonymous reviewer for his or her comments and suggestions. This work was partially supported by grants P-31063, P-31955, and I4513-N from the Austrian Science Foundation. The first-listed author was additionally supported by grant 3E017319 from the Flemish Science Foundation.

References

[1] Aguilera, J. P., and Baaz, M. Unsound inferences make proofs shorter. *J. Symbolic Logic 84* (2019), 102–122.
[2] Beklemishev, L. D. Reflection principles and provability algebras in formal arithmetic. *Russian Math. Surveys 60* (2005), 197–268.
[3] Gentzen, G. K. E. Untersuchungen über das logische Schließen, I. *Math. Z. 39* (1934), 176–210.
[4] Krajíček, J., and Pudlák, P. The number of proof lines and the size of proofs in first order logic. *Arch. Math. Logic 27* (1988), 69–84.
[5] Orevkov, V. P. Lower Bounds for Increasing Complexity of Derivations after Cut Elimination (in Russian). *J. Soviet Math.* (1982), 2337–2350.
[6] Schwichtenberg, H. Proof theory: Some applications of cut-elimination. In *Handbook of Mathematical Logic*, J. Barwise, Ed. 1982, pp. 867–896.
[7] Statman, R. Lower Bounds on Herbrand's Theorem. *Proc. Amer. Math. Soc. 75* (1979), 104–107.
[8] Takeuti, G. *Proof Theory (Second ed.)*. Dover Publications, 2013.

Institute of Discrete Mathematics and Geometry, Vienna University of Technology. Wiedner Hauptstrasse 8-10, 1040 Vienna, Austria *and* Department of Mathematics, Ghent University. Krijgslaan 281 - Gebouw S8, 9000 Gent, Belgium.
E-mail address: aguilera@logic.at

Institute of Discrete Mathematics and Geometry, Vienna University of Technology. Wiedner Hauptstrasse 8-10, 1040 Vienna, Austria.
E-mail address: baaz@logic.at

Institute of Discrete Mathematics and Geometry, Vienna University of Technology. Wiedner Hauptstrasse 8-10, 1040 Vienna, Austria.
E-mail address: jan.bydz@gmail.com