Basic proof theory

S. WAINER & L. WALLEN

Basic Proof Theory

S.S. WAINER,
Dept. of Pure Mathematics, University of Leeds, U.K.
L.A. WALLEN,
Computing Laboratory, University of Oxford, U.K.

§ 1 Introduction

This paper is an amalgam of two introductory lecture courses given at the Summer School. As the title suggests, the aim is to present fundamental notions of Proof Theory in their simplest settings, thus: Completeness and Cut-Elimination in Pure Predicate Logic; the Curry-Howard Correspondence and Normalization in the core part of Natural Deduction; connections to Sequent Calculus and Linear Logic; and applications to the Σ_1-Inductive fragment of arithmetic and the synthesis of primitive recursive bounding functions. The authors have tried to preserve a (readable) balance between rigour and informal lecture-note style.

§ 2 Pure Predicate Logic—Completeness

Classical first order predicate calculus (PC) is formulated here essentially in "Schütte-Ackermann-Tait" style, but with *multisets* instead of *sets* of formulas for sequents. It is kept "pure" (*i.e.*, no function symbols) merely for the sake of technical simplicity. The refinement to multiset sequents illuminates the rôle of the so-called *structural inferences* of *contraction* and *weakening* in proof-theoretic arguments.

The language of PC. The language consists of

- Individual variables: $x_0, x_1, x_2, \ldots$;
- Predicate symbols: $P_0, \bar{P}_0, P_1, \bar{P}_1, \ldots$ occurring in complementary pairs;
- Logical symbols: $\vee$ (or), $\wedge$ (and), $\exists$ (some), $\forall$ (all);
- Brackets for unique readability.

Formulas $A, B, \ldots$ are built up from atoms: $P(x_{i_1}, \ldots, x_{i_k})$, $\bar{P}(x_{i_1}, \ldots, x_{i_k})$, by applying $\vee$, $\wedge$, $\exists x$ and $\forall x$. Note that negation $\neg$ and implication $\rightarrow$ are

not included as basic logical symbols. Negation is *defined* by De Morgan's Laws: $\neg P \equiv \bar{P}$; $\neg\bar{P} \equiv P$; $\neg(A \vee B) \equiv \neg A \wedge \neg B$; $\neg(A \wedge B) \equiv \neg A \vee \neg B$; $\neg\exists x A \equiv \forall x \neg A$; $\neg\forall x A \equiv \exists x \neg A$. Implication $A \to B$ is *defined* to be $\neg A \vee B$. The reason for presenting logic in this way is that we want to exploit the duality between $\vee$ and $\wedge$, and $\exists$ and $\forall$. The price paid is that we cannot present intuitionistic logic in this way, since De Morgan's Laws are not intuitionistically valid.

Derivability in PC. Rather than deriving single formulas we shall derive finite multisets of them $\Gamma = \{A_1, A_2, ..., A_n\}$ meaning "A_1 or A_2 or ... or A_n". Γ, A means $\Gamma \cup \{A\}$ where the union sign ($\cup$) here is union of multisets.

The Proof-Rules of PC are (with any Γ and Δ):

$$\text{(Axioms)} \quad P(x_{i_1}, \ldots, x_{i_k}), \bar{P}(x_{i_1}, \ldots, x_{i_k})$$

$$(\vee) \quad \frac{\Gamma, A_0, A_1}{\Gamma, (A_0 \vee A_1)} \qquad (\wedge) \quad \frac{\Gamma, A_0 \quad \Delta, A_1}{\Gamma, \Delta, (A_0 \wedge A_1)}$$

$$(\exists) \quad \frac{\Gamma, A(x')}{\Gamma, \exists x A(x)} \qquad (\forall) \quad \frac{\Gamma, A(x')}{\Gamma, \forall x A(x)} \qquad x' \text{ not free in } \Gamma$$

$$(\mathrm{C}) \quad \frac{\Gamma, A, A}{\Gamma, A} \qquad (\mathrm{W}) \quad \frac{\Gamma}{\Gamma, A}$$

$$(\text{Cut}) \quad \frac{\Gamma, C \quad \Delta, \neg C}{\Gamma, \Delta} \qquad C \text{ is the "cut formula."}$$

We shall use $\vdash_{PC} \Gamma$ to mean that there is a PC-derivation of Γ from axioms.

EXERCISES. Show that for all Γ and all A,

(1) $\vdash_{PC} \Gamma, \neg A, A$. (Hint: prove $\vdash \neg A, A$ by induction on the "build-up" of A, then use weakening. Call this (Axiom′).

(2) If $\vdash_{PC} \Gamma, (A_0 \wedge A_1)$ then $\vdash_{PC} \Gamma, A_0$ and $\vdash_{PC} \Gamma, A_1$.

(3) If $\vdash_{PC} \Gamma, \forall x A(x)$ then $\vdash_{PC} \Gamma, A(t)$, for any term (*i.e.*, variable) t.

Alternative formulations. The ($\vee$) rule is often formulated as:

$$(\vee') \quad \frac{\Gamma, A_i}{\Gamma, (A_0 \vee A_1)} \qquad i = 0 \text{ or } 1,$$

which is easily seen to be equivalent to ($\vee$) in the presence of contraction and weakening. Moreover, the rules are usually understood as working over

sets instead of multisets. In the set approach the rules with two premises ($\wedge$) and (Cut) can be given the same contexts Γ, *i.e.*,

$$(\wedge') \quad \frac{\Gamma, A_0 \quad \Gamma, A_1}{\Gamma, (A_0 \wedge A_1)} \qquad (\mathrm{Cut}') \quad \frac{\Gamma, C \quad \Gamma, \neg C}{\Gamma}.$$

If the axioms are taken as in the above exercise, both contraction and weakening can be dropped with no change in the set of provable sequents. The system comprising (Axiom′), ($\vee'$), ($\wedge'$), ($\exists$), ($\forall$) and (Cut′), working over sets of formulas is known as the Schütte-Ackermann-Tait presentation of PC.

EXERCISE.

(4) Prove that the Schütte-Ackermann-Tait rules working over *multisets* are indeed derived rules of PC, because of Contraction and Weakening.

REMARK. In the absence of contraction and weakening, but with sequents as sets, the alternative formulations of the ($\vee$) and ($\wedge$) rules lead to distinct connectives and thence to Linear Logic. Girard (1987) calls the primed rules with implicit contraction "additive" and our original rules with distinct contexts "multiplicative."

The Semantics of PC. An *interpretation* of PC gives a fixed meaning to all the formulas and consists of a structure $\mathcal{M} = \langle M, P_0^M, P_1^M, P_2^M, \ldots \rangle$ where M is some non-empty set and P_k^M is a relation on M which has the same arity as P_k and, given any list of arguments from M, is either true or false. Thus with respect to a given interpretation, and a given assignment $x_{i_1} := m_1, \ldots, x_{i_n} := m_n$ of elements of M to the free variables, a formula $A(x_{i_1}, \ldots, x_{i_n})$ makes a statement about $\mathcal{M}$ which is either true ($\boldsymbol{t}$) or false ($\boldsymbol{f}$). If it works out $\boldsymbol{t}$ under *all* possible interpretations $\mathcal{M}$ and *all* possible assignments of elements of M to its free variables, then A is said to be (logically or universally) valid.

COMPLETENESS THEOREM (GÖDEL 1930). $\vdash_{PC} \Gamma$ *iff* Γ *is valid.*

PROOF: For *soundness:* $\vdash_{PC} \Gamma \Rightarrow \Gamma$ is valid; simply note that the axioms are valid and each of the rules preserves validity.

For *adequacy:* $\nvdash_{PC} \Gamma \Rightarrow \Gamma$ not valid; we try to construct a derivation tree for Γ by successively taking it to bits using the ($\vee$), ($\wedge$), ($\exists$), ($\forall$) rules backwards. We do not use Cut! Since we are assuming that Γ is not derivable, this procedure must fail, and out of the failure we can construct an interpretation in which Γ is false. Hence Γ is not valid. It goes thus:

First write out Γ as a sequence of formulas, starting with the atoms (if there are any). Let A denote the first non-atomic formula in the sequence and Δ the rest of Γ, thus

$$\Gamma = \text{atoms}, A, \Delta.$$

Now take A to bits using whichever one of the rules $(\vee)$, $(\wedge)$, $(\exists)$, $(\forall)$ applies. This produces one or (in the case of $\wedge$) two new sequences of formulas Γ' as follows:

$(\vee)$ If $\Gamma = \text{atoms}, (A_0 \vee A_1), \Delta$ then $\Gamma' = \text{atoms}, A_0, A_1, \Delta$;

$(\wedge)$ If $\Gamma = \text{atoms}, (A_0 \wedge A_1), \Delta$ then $\Gamma'_i = \text{atoms}, A_i, \Delta$ for each $i = 0, 1$;

$(\forall)$ If $\Gamma = \text{atoms}, \forall x A(x), \Delta$ then $\Gamma' = \text{atoms}, A(x_j), \Delta$;

$(\exists)$ If $\Gamma = \text{atoms}, \exists x A(x), \Delta$ then $\Gamma' = \text{atoms}, A(x_k), \Delta, \exists x A(x)$,

where, in $(\forall)$ x_j is any new variable not already used, and in $(\exists)$ x_k is the first variable in the list $x_0, x_1, x_2, \ldots$ which has not already been used at a previous stage to witness the same formula $\exists x A(x)$.

Repeat this process to form $\Gamma, \Gamma', \Gamma'', \ldots$ and notice that each time, Γ follows from Γ' by applying the corresponding rule. (Notice that here we are actually using the derived rule $(\wedge')$ of the Scütte-Ackermann-Tait system since we "duplicate" the context: "atoms, Δ," in each of the two sequents Γ'_i that result.) In this way we develop what looks like a derivation-tree for Γ with branching at applications of the $(\wedge)$ rule. But assuming Γ is not derivable in PC there must be at least one branch on this tree — call it $\mathcal{B}$ — which either (a) terminates in a sequence of atoms only, but is not a (Schütte-Ackermann-Tait) logical axiom, or (b) goes on forever!

From $\mathcal{B}$ we construct a "counter-interpretation",

$$\mathcal{M} = \left\langle N, P_0^M, P_1^M, P_2^M, \ldots \right\rangle$$

where $N = \{0, 1, 2, 3, \ldots\}$ and the relations P_j^M are defined as follows :

$$P_j^M(i_1, \ldots, i_n) \Leftrightarrow_{\text{Def}} \text{ the atom } P_j(x_{i_1}, \ldots, x_{i_n}) \text{ does not occur on } \mathcal{B}.$$

CLAIM: *Under the interpretation $\mathcal{M}$ and the assignment $x_i := i$ to free variables, every formula A occurring on $\mathcal{B}$ is false.*

PROOF: (of CLAIM.) By induction on the build-up of formulas A occurring on $\mathcal{B}$, noticing that as the sequence $\Gamma, \Gamma', \Gamma'', \ldots$ is developed, every non-atomic formula on $\mathcal{B}$ will eventually "come under attention" as the first non-atomic formula in some stage:

(i) $A \equiv P_j(x_{i_1}, \ldots, x_{i_n})$ gets $\boldsymbol{f}$ by definition.

(ii) $A \equiv \bar{P}_j(x_{i_1}, \ldots, x_{i_n})$ gets $\boldsymbol{f}$ because its complement $P_j(x_{i_1}, \ldots, x_{i_n})$ cannot be on $\mathcal{B}$ (otherwise $\mathcal{B}$ would terminate in an axiom) and hence $P_j(x_{i_1}, \ldots, x_{i_n})$ gets $\boldsymbol{t}$ by definition.

(iii) $A \equiv A_0 \vee A_1$. Since A comes under attention at some stage in $\mathcal{B}$, both A_0 and A_1 also occur on $\mathcal{B}$. So by the induction hypothesis, both get $\boldsymbol{f}$ and hence so does A.

(iv) $A \equiv A_0 \wedge A_1$. Again, since A must come under attention at some stage, either A_0 or A_1 is on $\mathcal{B}$. So one of them gets $\boldsymbol{f}$ and hence so does A.

(v) $A \equiv \forall x A_0(x)$. In this case $A_0(x_j)$ is also on $\mathcal{B}$ for one of the variables x_j. So $A_0(x_j)$ gets $\boldsymbol{f}$ and hence so does A.

(vi) $A \equiv \exists x A_0(x)$. Then by the construction of $\mathcal{B}$, A comes under attention infinitely often and each time a "new" $A_0(x_k)$ is introduced. Therefore *every one* of

$$A_0(x_0), A_0(x_1), A_0(x_2), A_0(x_3), \ldots$$

occurs on $\mathcal{B}$, and they all get $\boldsymbol{f}$. Hence A gets $\boldsymbol{f}$.

This completes the proof of the claim. □

Now since every formula in the set Γ we started with occurs on $\mathcal{B}$, they all get $\boldsymbol{f}$ under this interpretation. Thus Γ is not valid. □

§ 3 Pure Predicate Logic—Cut-elimination

THE CUT-ELIMINATION THEOREM (GENTZEN 1936) *If Γ is derivable in PC then it is derivable without any use of the Cut-rule.*

PROOF: *(Semantic Proof.)* If $\vdash_{PC} \Gamma$ then by the Soundness of PC, Γ is valid. But the proof of adequacy actually shows that if Γ is not derivable using only the rules $\vee, \wedge, \exists, \forall$, then Γ is not valid. Since Γ is valid, it must therefore be derivable without Cut. □

In the rest of this section we shall develop a syntactic proof of the Cut-Elimination Theorem. We shall approach the result in two steps. First we shall prove the result for a subsystem of PC called MPC (standing for Multiplicative fragment of PC). This subsystem is formed by dropping the rules for contraction (C) and weakening (W) from the system given in the previous section. The importance of this subsystem of predicate logic has been stressed by Girard (1987). MPC will help us to illuminate the rôle played by the structural rules in various results like existence and disjunction properties. The Cut-Elimination result is then extended to PC.

Cut elimination in MPC takes on a particularly simple form since the reduction and elimination of cuts from a proof *decreases* the size of a proof. This is in contrast to the situation in both Classical and Intuitionistic Logic. Cut-free proofs are therefore the smallest proofs of sequents.

REMARK. This respects the idea that cuts are "indirections" in a proof. If a proof makes recourse to indirections, one should expect its size to exceed

that of a "direct" proof. On the other hand, if having derived a sequent once it may nevertheless be used *more than once* within a derivation, we might expect the introduction of the indirection to lead to a decrease in size. Consequently, cuts may be used to shorten proofs in the presence of contraction.

Size, height and cut-rank of derivations. Each inference (Axiom), $(\vee)$, $(\wedge)$, $(\exists)$ and $(\forall)$, has the form:

$$\frac{\Gamma_i, \Phi_i}{\Gamma, \Theta} \quad (i < k) \quad \text{for some } k : 0 \leq k \leq 2,$$

where the $(\Phi_i)_{i<k}$ are the minor formulas of the inference, Θ the principal formula(s) and $\Gamma = \bigcup_{i<k} \Gamma_i$ (multiset union). In the sequel we shall use π_i, $i = 0, 1$, to denote the immediate subderivations of a derivation π and we shall supress mention of k. For example, $\sum_i f(\pi_i)$ will be used to denote the sum of the values of function f (from derivations to natural numbers, say) over the immediate subderivations of π; if π is an axiom, *i.e.*, $k = 0$, we have $\sum_i f(\pi_i) = 0$. Given this convention, we can define the *size*, $\boldsymbol{s}\pi$, of a derivation π inductively as follows:

$$\boldsymbol{s}\pi = 1 + \sum_i \boldsymbol{s}\pi_i.$$

Notice that if π is an axiom, $\boldsymbol{s}\pi = 1$, hence the size of a derivation equals the number of inferences that comprise it.

The *height*, $\mid \pi \mid$, of a derivation π is the length of its longest branch, *i.e.*,

$$\mid \pi \mid = 1 + \sup_i \mid \pi_i \mid.$$

Likewise, the height of a formula is the length of the longest branch in its formation tree (greatest nesting of connectives).

The *cut-rank*, $\boldsymbol{r}\pi$, of a derivation π is the height of the "tallest" cut-formula in π, *i.e.*,

$$\boldsymbol{r}\pi = \begin{cases} \sup(\mid C \mid, \sup_i \boldsymbol{r}\pi_i) & \pi \text{ ends in cut on } C; \\ \sup_i \boldsymbol{r}\pi_i, & \text{otherwise.} \end{cases}$$

If $\pi(x) \vdash \Gamma(x)$ denotes a proof π of sequent Γ with variable x free, and x' is a variable free for x in π, then $\pi(x')$ denotes the proof obtained from π by substitution of x' for x. Substitution has no effect on the size, height or cut-rank of a proof:

SUBSTITUTION LEMMA *If $\lambda(x) \vdash \Gamma(x)$ and x' is free for x in π, then $\lambda(x') \vdash \Gamma(x')$ with $\boldsymbol{r}\pi(x') = \boldsymbol{r}\pi(x)$, $\boldsymbol{s}\pi(x') = \boldsymbol{s}\pi(x)$ and $\mid \pi(x') \mid = \mid \pi(x) \mid$.*

The main technical tool in the Cut-Elimination argument is the following:

CUT-REDUCTION LEMMA *If* $\lambda \vdash \Gamma, A$ *and* $\rho \vdash \Delta, \neg A$, *both with cut-rank* $< r = |\, A \,|$, *then there is a derivation* $\pi \vdash \Gamma, \Delta$ *such that*

(i) $\boldsymbol{r}\pi < r$;

(ii) $\boldsymbol{s}\pi \leq \boldsymbol{s}\lambda + \boldsymbol{s}\rho$;

(iii) $|\, \pi \,| \leq |\, \lambda \,| + |\, \rho \,|$.

PROOF: By induction on $\boldsymbol{s}\lambda + \boldsymbol{s}\rho$.

Case 1. Either A is a side formula of the last inference of λ or $\neg A$ is a side formula of the last inference of ρ.

By symmetry, we may assume the former. λ is of the form:

$$\lambda \;=\; \dfrac{\begin{matrix}\lambda_i & \lambda_j \\ \Gamma_i, B_i & \Gamma_j, B_j, A\end{matrix}}{\Gamma, A} \qquad (0 \leq i \neq j < k),$$

i.e., with at least one premise, since axioms have no side formulae. Moreover, in the absence of (implicit) contraction, A is the side formula of *exactly* one premise (distinguished here as the jth).

The induction hypothesis with $\lambda_j \vdash \Gamma_j, B_j, A$ and $\rho \vdash \Delta, \neg A$ gives a proof $\pi' \vdash \Gamma_j, B_j, \Delta$ with cut-rank $< r$, size $\leq \boldsymbol{s}\lambda_j + \boldsymbol{s}\rho$ and height $\leq |\, \lambda_j \,| + |\, \rho \,|$. Consider π, given by:

$$\pi \;=\; \dfrac{\begin{matrix}\lambda_i & \pi' \\ \Gamma_i, B_i & \Gamma_j, B_j, \Delta\end{matrix}}{\Gamma, \Delta} \qquad (0 \leq i \neq j < k).$$

Thus $\boldsymbol{r}\pi = \boldsymbol{r}\pi' < r$. Also we have:

$$\boldsymbol{s}\pi \;=\; 1 + \sum_{i \neq j < k} \boldsymbol{s}\lambda_i + \boldsymbol{s}\pi' \;\leq\; 1 + \sum_{i \neq j < k} \boldsymbol{s}\lambda_i + \boldsymbol{s}\lambda_j + \boldsymbol{s}\rho \;=\; \boldsymbol{s}\lambda + \boldsymbol{s}\rho,$$

and

$$|\, \pi \,| = 1 + \sup\left(\sup_{i \neq j < k} |\, \lambda_i \,|, |\, \pi' \,|\right) \;\leq\; \sup(|\, \lambda \,|, |\, \lambda \,| + |\, \rho \,|) = |\, \lambda \,| + |\, \rho \,|.$$

Case 2. A is a principal formula of λ and $\neg A$ is a principal formula of ρ. There are six cases according to the structure of A, which are reduced by symmetry to three.

(a) $A = P(x_1, \ldots, x_n)$. (Symmetrical case: $A = \bar{P}(x_1, \ldots, x_n)$.) Then λ and ρ are of the form:

$$\lambda = A, \neg A \qquad \rho = \neg A, A.$$

Consider the proof π given by

$$\pi = \neg A, A.$$

It is clear that π has the appropriate bounds on rank, height and size.

(b) $A = B \wedge C$. (Symmetrical case: $A = B \vee C$.) Then λ and ρ are of the form:

$$\lambda = \frac{\begin{matrix}\lambda_0 & \lambda_1\\ \Gamma_0, B & \Gamma_1, C\end{matrix}}{\Gamma, B \wedge C} \qquad \rho = \frac{\rho_0}{\Delta, \neg B, \neg C}{\Delta, \neg B \vee \neg C}$$

Consider the proof

$$\pi = \frac{\begin{matrix}\lambda_0\\ \Gamma_0, B\end{matrix} \quad \frac{\begin{matrix}\lambda_1 & \rho_0\\ \Gamma_1, C & \Delta, \neg B, \neg C\end{matrix}}{\Gamma_1, \Delta, \neg B}}{\Gamma, \Delta}$$

Again the cut-rank of π is $< r$ and we have:

$$s\pi = s\lambda_0 + s\lambda_1 + s\rho_0 + 2 = s\lambda + s\rho \ ,$$

and

$$\begin{aligned} |\pi| &= 1 + \sup(|\lambda_0|, 1 + \sup(|\lambda_1|, |\rho_0|)) \\ &\leq \sup(|\lambda|, 1 + \sup(|\lambda|, |\rho|)) \\ &\leq |\lambda| + |\rho|. \end{aligned}$$

(c) $A = \forall x B$. (Symmetrical case: $A = \exists x B$.) Then λ and ρ have the form:

$$\lambda = \frac{\lambda'(y)}{\Gamma, B(y)}{\Gamma, \forall x B(x)} \qquad \rho = \frac{\rho'}{\Delta, \neg B(x')}{\Delta, \exists x \neg B(x)}$$

Consider the proof π given by:

$$\pi \quad = \quad \dfrac{\begin{array}{cc}\lambda'(x') & \rho' \\ \Gamma, B(x') & \Delta, \neg B(x')\end{array}}{\Gamma, \Delta}$$

By the Substitution Lemma, $s\lambda'(x') = s\lambda'(y)$ and $|\ \lambda'(x')\ | = |\ \lambda'(y)\ |$. Hence

$$s\pi \quad = \quad 1 + s\lambda'(x') + s\rho' \quad = \quad 1 + s\lambda' + s\rho' \quad \leq \quad s\lambda + s\rho;$$

and

$$|\ \pi\ | \ = \ 1 + \sup(|\ \lambda'(x')\ |, |\ \rho'\ |) \ \leq \ \sup(|\ \lambda\ |, |\ \rho\ |) \ \leq \ |\ \lambda\ | + |\ \rho\ |.$$

This ends the proof. □

CUT-RANK REDUCTION LEMMA. *If $\pi \vdash \Gamma$ with cut rank $r > 0$, there is a proof $\pi' \vdash \Gamma$ with strictly smaller cut-rank such that: $s\pi' < s\pi$ and $|\ \pi'\ | \leq 2^{|\pi|}$.*

PROOF: By induction on $s\pi$. Assume that the last inference of π is a cut of rank r (the result follows immediately from the induction hypothesis in the other cases; note that π cannot be an axiom.) The last inference is therefore of the form:

$$\dfrac{\begin{array}{cc}\lambda & \rho \\ \Gamma, A & \Delta, \neg A\end{array}}{\Gamma, \Delta}.$$

By the induction hypothesis on λ and ρ we get $\lambda' \vdash \Gamma, A$ and $\rho' \vdash \Delta, \neg A$ with ranks $< r$, sizes $< s\lambda$ and $< s\rho$ resp., and heights $\leq 2^{|\lambda|}$ and $\leq 2^{|\rho|}$ resp. The Cut-Reduction Lemma on λ' and ρ' yields $\pi' \vdash \Gamma, \Delta$ with rank $< r$, size $\leq s\lambda + s\rho < s\pi$ and height $\leq 2^{|\lambda|} + 2^{|\rho|} \leq 2^{1+\sup(|\lambda|,|\rho|)} = 2^{|\pi|}$. □

Define 2^k_r by:

$$2^k_r \quad = \quad \begin{cases} k & r = 0, \\ 2^{2^k_{r-1}} & r > 1. \end{cases}$$

we can now formulate the Cut-Elimination Theorem, namely,

CUT-ELIMINATION THEOREM FOR MPC. *If $\pi \vdash \Gamma$ with cut rank $r > 0$, there is a cut-free proof $\pi^* \vdash \Gamma$ such that $s\pi^* < s\pi$ and $|\ \pi^*\ | = 2^{|\pi|}_r$.*

From the Cut-Elimination Theorem it is clear that (in MPC) the elimination of cuts reduces the size of proofs as expected, but increases their height exponentially.