

# A Proof Theorist's Guide to First Aid

## CUT-Elimination and Other Results of Classical Proof Theory

Logan Heath

*Graduate Logic Seminar*  
University of Wisconsin-Madison  
11/20/2023

**When are correct proofs also “bad” proofs?**

# Sequents

## Definition

If  $\Gamma$  and  $\Delta$  are finite sequences of logical formulas, then  $\Gamma \Rightarrow \Delta$  is a sequent. The sequences  $\Gamma$  and  $\Delta$  are, respectively, the antecedent and succedent of the sequent and either may be empty.

The intended meaning of a sequent  $\Gamma \Rightarrow \Delta$  is that if all the formulas in the antecedent are true, then at least one formula in the succedent is true.

## Definition

An axiom is a non-empty sequent of the form  $\varphi \Rightarrow \varphi$  where  $\varphi$  is a logical formula.

# Rules of Inference

## Definition (Operational Rules)

$$\frac{\varphi, \Gamma \Rightarrow \Theta}{(\varphi \wedge \psi), \Gamma \Rightarrow \Theta} \wedge L_1$$

$$\frac{\psi, \Gamma \Rightarrow \Theta}{(\varphi \wedge \psi), \Gamma \Rightarrow \Theta} \wedge L_2$$

$$\frac{\varphi, \Gamma \Rightarrow \Theta \quad \psi, \Gamma \Rightarrow \Theta}{(\varphi \vee \psi), \Gamma \Rightarrow \Theta} \vee L$$

$$\frac{\Gamma \Rightarrow \Theta, \varphi \quad \psi, \Delta \Rightarrow \Lambda}{(\varphi \rightarrow \psi), \Gamma, \Delta \Rightarrow \Theta, \Lambda} \rightarrow L$$

$$\frac{\Gamma \Rightarrow \Theta, \varphi}{\neg \varphi, \Gamma \Rightarrow \Theta} \neg L$$

$$\frac{\Gamma \Rightarrow \Theta, \varphi \quad \Gamma \Rightarrow \Theta, \psi}{\Gamma \Rightarrow \Theta, (\varphi \wedge \psi)} \wedge R$$

$$\frac{\Gamma \Rightarrow \Theta, \varphi}{\Gamma \Rightarrow \Theta, (\varphi \vee \psi)} \vee R_1$$

$$\frac{\Gamma \Rightarrow \Theta, \psi}{\Gamma \Rightarrow \Theta, (\varphi \vee \psi)} \vee R_2$$

$$\frac{\varphi, \Gamma \Rightarrow \Theta, \psi}{\Gamma \Rightarrow \Theta, (\varphi \rightarrow \psi)} \rightarrow R$$

$$\frac{\varphi, \Gamma \Rightarrow \Theta}{\Gamma \Rightarrow \Theta, \neg \varphi} \neg R$$

# Rules of Inference Continued

## Definition (Operational Rules Continued)

$$\frac{\varphi(t), \Gamma \Rightarrow \Theta}{\forall x \varphi(x), \Gamma \Rightarrow \Theta} \forall L$$

$$! \frac{\Gamma \Rightarrow \Theta, \varphi(a)}{\Gamma \Rightarrow \Theta, \forall x \varphi(x)} \forall R$$

$$! \frac{\varphi(a), \Gamma \Rightarrow \Theta}{\exists x \varphi(x), \Gamma \Rightarrow \Theta} \exists L$$

$$\frac{\Gamma \Rightarrow \Theta, \varphi(t)}{\Gamma \Rightarrow \Theta, \exists x \varphi(x)} \exists R$$

In the inference rules above, the formulas occurring in  $\Gamma$ ,  $\Theta$ ,  $\Delta$ , and  $\Lambda$  are called side formulas. The remaining formulas in the premise(s) are called auxiliary formulas and the remaining formulas in the conclusion are called principal formulas.

# Rules of Inference Continued

## Definition (Structural Rules)

$$\frac{\Gamma \Rightarrow \Theta}{\varphi, \Gamma \Rightarrow \Theta} \text{ WL}$$

$$\frac{\Gamma \Rightarrow \Theta}{\Gamma \Rightarrow \Theta, \varphi} \text{ WR}$$

$$\frac{\varphi, \varphi, \Gamma \Rightarrow \Theta}{\varphi, \Gamma \Rightarrow \Theta} \text{ CL}$$

$$\frac{\Gamma \Rightarrow \Theta, \varphi, \varphi}{\Gamma \Rightarrow \Theta, \varphi} \text{ CR}$$

$$\frac{\Delta, \varphi, \psi, \Gamma \Rightarrow \Theta}{\Delta, \psi, \varphi, \Gamma \Rightarrow \Theta} \text{ IL}$$

$$\frac{\Gamma \Rightarrow \Theta, \psi, \varphi, \Lambda}{\Gamma \Rightarrow \Theta, \varphi, \psi, \Lambda} \text{ IR}$$

$$\frac{\Gamma \Rightarrow \Theta, \varphi \quad \varphi, \Delta \Rightarrow \Lambda}{\Gamma, \Delta \Rightarrow \Theta, \Lambda} \text{ CUT}$$

Formulas appearing in  $\Gamma$ ,  $\Theta$ ,  $\Delta$ , and  $\Lambda$  will again be called side formulas. The remaining formulas are called principal formulas. Structural rules do not have auxiliary formulas. The principal formula of a CUT inference will be called the CUT-formula of the inference.

## Definition (LK Proofs)

A proof in LK is a finite tree of sequents where the leaf nodes are axioms and all other nodes in the tree follow from the node(s) immediately above by a rule of inference. A sequent  $\Gamma \Rightarrow \Delta$  is provable in LK if it is the root node of some LK proof. A formula  $\varphi$  is provable in LK if the sequent  $\Rightarrow \varphi$  is provable in LK.

# The Subformula Property

We will think of proofs utilizing CUT as “unhealthy” in some sense, while CUT-free proofs will have a number of nice properties. The most useful for us will be the following:

## Lemma (The Subformula Property)

*Let  $\pi$  be a CUT-free proof with end-sequent  $\Gamma \Rightarrow \Delta$  and let  $\Theta \Rightarrow \Lambda$  be any sequent occurring in  $\pi$ . Then every formula in  $\Theta, \Lambda$  occurs as a sub-formula of some formula in  $\Gamma, \Delta$ .*

We will use the Subformula Property and the equivalence of LK with LK-CUT to establish the consistency of LK.



# The Consistency of LK

## Corollary (LK is Consistent)

*If LK is equivalent to LK-CUT, then there is no formula  $\varphi$  such that both  $\varphi$  and  $\neg\varphi$  are provable in LK.*

## Proof.

If there were such a formula  $\varphi$  with proofs  $\pi_1$  and  $\pi_2$  of  $\varphi$  and  $\neg\varphi$ , respectively, then there would also be a proof of the empty sequent:

$$\frac{\begin{array}{c} \vdots \\ \vdots \pi_2 \\ \Rightarrow \neg\varphi \end{array} \quad \frac{\begin{array}{c} \vdots \\ \vdots \pi_1 \\ \Rightarrow \varphi \end{array}}{\neg\varphi \Rightarrow \varphi} \neg L}{\Rightarrow} \text{CUT}$$

From this we could obtain a CUT-free proof of the empty sequent, but such a proof can consist only of empty sequents by The Subformula Property. Contradiction. □

# The CUT-Elimination Theorem

## Theorem

*If  $\Gamma \Rightarrow \Delta$  is provable in LK, then  $\Gamma \Rightarrow \Delta$  has a proof in LK-CUT.*

# The CUT-Elimination Theorem

## Theorem

*If  $\Gamma \Rightarrow \Delta$  is provable in LK, then  $\Gamma \Rightarrow \Delta$  has a proof in LK-CUT.*

## Definition (MIX)

We define a new structural rule

$$\frac{\Gamma \Rightarrow \Theta \quad \Delta \Rightarrow \Lambda}{\Gamma, \Delta^* \Rightarrow \Theta^*, \Lambda} \text{ MIX}$$

where  $\Delta$  and  $\Theta$  are both assumed to contain at least one occurrence of a given formula  $\varphi$ , the MIX-formula of the inference, and  $\Delta^*$  and  $\Theta^*$  are obtained from  $\Delta$  and  $\Theta$ , respectively, by removing all occurrences of  $\varphi$ .

# The CUT-Elimination Theorem

## Theorem

*If  $\Gamma \Rightarrow \Delta$  is provable in LK, then  $\Gamma \Rightarrow \Delta$  has a proof in LK-CUT.*

## Definition (MIX)

We define a new structural rule

$$\frac{\Gamma \Rightarrow \Theta \quad \Delta \Rightarrow \Lambda}{\Gamma, \Delta^* \Rightarrow \Theta^*, \Lambda} \text{ MIX}$$

where  $\Delta$  and  $\Theta$  are both assumed to contain at least one occurrence of a given formula  $\varphi$ , the MIX-formula of the inference, and  $\Delta^*$  and  $\Theta^*$  are obtained from  $\Delta$  and  $\Theta$ , respectively, by removing all occurrences of  $\varphi$ .

## Lemma

*A sequent  $\Gamma \Rightarrow \Delta$  is provable in LK iff it is provable in LK-CUT+MIX*

# The CUT-Elimination Theorem

## Theorem

*If  $\Gamma \Rightarrow \Delta$  is provable in  $LK-CUT+MIX$ , then  $\Gamma \Rightarrow \Delta$  has a proof in  $LK-CUT$ .*

# The CUT-Elimination Theorem

## Theorem

*If  $\Gamma \Rightarrow \Delta$  is provable in LK-CUT+MIX, then  $\Gamma \Rightarrow \Delta$  has a proof in LK-CUT.*

## Proof.

We will proceed by showing that it is always possible to remove, say, the topmost, leftmost MIX. The result then follows by induction on the number of MIX inferences occurring in the proof. Since we can focus our attention on sub-proofs containing a single MIX inference as their last inference, we can actually just focus on proofs of this form.  $\square$

# The Proof

## Lemma

*If  $\Gamma \Rightarrow \Delta$  has a proof in LK-CUT+MIX containing exactly one MIX inference as its last inference, then  $\Gamma \Rightarrow \Delta$  has a proof in LK-CUT.*

# The Proof

## Lemma

*If  $\Gamma \Rightarrow \Delta$  has a proof in LK-CUT+MIX containing exactly one MIX inference as its last inference, then  $\Gamma \Rightarrow \Delta$  has a proof in LK-CUT.*

## Proof (Sketch).

We will associate to our proof a pair of natural numbers and then “push” the MIX inference upwards through the tree. As we do this we will sometimes obtain new MIX inferences, but the sub-proofs ending in these new MIX inferences will have associated to them pairs of natural numbers which are less in the lexicographic ordering than the pair we started with. Induction then takes over and we are done. Essentially we “push” our MIX inferences upwards until we have “pushed” them all the way out of our tree. □



# The Proof: What do We Induct On?

## Proof (Sketch).

We need to assign a measure of complexity to our proofs so that we can carry out the desired induction. Let  $\pi$  be a proof containing exactly one MIX inference as its last inference.

# The Proof: What do We Induct On?

## Proof (Sketch).

We need to assign a measure of complexity to our proofs so that we can carry out the desired induction. Let  $\pi$  be a proof containing exactly one MIX inference as its last inference.

- Define the degree  $d(\pi)$  of  $\pi$  to be the degree of the MIX-formula of  $\pi$ 's MIX inference.

# The Proof: What do We Induct On?

## Proof (Sketch).

We need to assign a measure of complexity to our proofs so that we can carry out the desired induction. Let  $\pi$  be a proof containing exactly one MIX inference as its last inference.

- Define the degree  $d(\pi)$  of  $\pi$  to be the degree of the MIX-formula of  $\pi$ 's MIX inference.
- Define the left rank  $rk_l(\pi)$  of  $\pi$  to be the maximum number of consecutive sequents which each contain the MIX formula in their succedent and constitute a path terminating in the left premise of the MIX. Define the right rank  $rk_r(\pi)$  of  $\pi$  similarly.

# The Proof: What do We Induct On?

## Proof (Sketch).

We need to assign a measure of complexity to our proofs so that we can carry out the desired induction. Let  $\pi$  be a proof containing exactly one MIX inference as its last inference.

- Define the degree  $d(\pi)$  of  $\pi$  to be the degree of the MIX-formula of  $\pi$ 's MIX inference.
- Define the left rank  $rk_l(\pi)$  of  $\pi$  to be the maximum number of consecutive sequents which each contain the MIX formula in their succedent and constitute a path terminating in the left premise of the MIX. Define the right rank  $rk_r(\pi)$  of  $\pi$  similarly.
- Define the rank  $rk(\pi)$  of  $\pi$  to be  $rk_l(\pi) + rk_r(\pi)$ .

# The Proof: What do We Induct On?

## Proof (Sketch).

We need to assign a measure of complexity to our proofs so that we can carry out the desired induction. Let  $\pi$  be a proof containing exactly one MIX inference as its last inference.

- Define the degree  $d(\pi)$  of  $\pi$  to be the degree of the MIX-formula of  $\pi$ 's MIX inference.
- Define the left rank  $rk_l(\pi)$  of  $\pi$  to be the maximum number of consecutive sequents which each contain the MIX formula in their succedent and constitute a path terminating in the left premise of the MIX. Define the right rank  $rk_r(\pi)$  of  $\pi$  similarly.
- Define the rank  $rk(\pi)$  of  $\pi$  to be  $rk_l(\pi) + rk_r(\pi)$ .

Our proof then will be by induction on pairs  $(d(\pi), r(\pi))$  ordered lexicographically. □

# The Proof: How Does it Work?

Consider the following proof  $\pi$  where  $F(x)$  is atomic:

$$\frac{\frac{\frac{\frac{F(t) \Rightarrow F(t)}{F(t) \Rightarrow \boxed{\exists x F(x)}_1} \exists R}{\neg \exists x F(x), \boxed{\exists x F(x)}_4 \Rightarrow} \neg L}{\neg \exists x F(x), \boxed{\exists x F(x)}_3 \Rightarrow G(s)} \text{WR}}{\frac{\frac{\frac{\frac{G(s) \Rightarrow G(s)}{\boxed{\exists x F(x)}_4 \Rightarrow G(s)} \text{WL}}{G(s), \boxed{\exists x F(x)}_3, G(s) \Rightarrow G(s)} \text{IL}}{G(s), \boxed{\exists x F(x)}_3, G(s) \Rightarrow G(s)} \text{VL}}{\neg \exists x F(x) \vee G(s), \boxed{\exists x F(x)}_2 \Rightarrow G(s)} \text{IL}}{F(t), \neg \exists x F(x) \vee G(s) \Rightarrow G(s)} \text{MIX}$$

We have  $d(\pi) = 1$  since the MIX-formula,  $\exists x F(x)$ , has only one logical symbol. We also have  $rk(\pi) = rk_l(\pi) + rk_r(\pi) = 1 + 5 = 6$ .

Our first step will be to push the MIX upwards past the IL inference to reduce the right rank while keeping the degree and left rank fixed.

# The Proof: How Does it Work?

This results in

$$\frac{
 \frac{
 \frac{
 F(t) \Rightarrow F(t)
 }{
 F(t) \Rightarrow \boxed{\exists x F(x)}_1
 }{\exists R}
 \quad
 \frac{
 \frac{
 \frac{
 \boxed{\exists x F(x)}_4 \Rightarrow \exists x F(x)
 }{\neg \exists x F(x), \boxed{\exists x F(x)}_3 \Rightarrow
 }{\neg L}
 \quad
 \frac{
 \frac{
 G(s) \Rightarrow G(s)
 }{
 \boxed{\exists x F(x)}_3, G(s) \Rightarrow G(s)
 }{WL}
 \quad
 \frac{
 G(s), \boxed{\exists x F(x)}_2 \Rightarrow G(s)
 }{IL}
 }{
 G(s), \boxed{\exists x F(x)}_2 \Rightarrow G(s)
 }{\forall L}
 }{
 \neg \exists x F(x), \boxed{\exists x F(x)}_2 \Rightarrow G(s)
 }{WR}
 }{
 \neg \exists x F(x) \vee G(s), \boxed{\exists x F(x)}_1 \Rightarrow G(s)
 }{MIX}
 }{
 F(t), \neg \exists x F(x) \vee G(s) \Rightarrow G(s)
 }$$

$$d(\pi) = 1$$

$$rk(\pi) = rk_l(\pi) + rk_r(\pi) = 1 + 4 = 5$$

Our next step will be to push our MIX inference up past the  $\forall L$  inference.

# The Proof: How Does it Work?

Since  $\forall L$  has two premises we obtain a proof with two MIX inferences. However, in each of the corresponding sub-proofs the right ranks will be less than before and the degrees and left ranks will stay the same:

$$\left\{ \begin{array}{l}
 \frac{F(t) \Rightarrow F(t)}{F(t) \Rightarrow \boxed{\exists x F(x)}_1} \exists R \quad \frac{\frac{\frac{\boxed{\exists x F(x)}_3 \Rightarrow \exists x F(x)}{\neg \exists x F(x), \boxed{\exists x F(x)}_2 \Rightarrow} \neg L}{\neg \exists x F(x), \boxed{\exists x F(x)}_1 \Rightarrow G(s)} \text{WR}}{\frac{F(t), \neg \exists x F(x) \Rightarrow G(s)}{\neg \exists x F(x), F(t) \Rightarrow G(s)} \text{IL}} \text{MIX} \\
 \frac{F(t), \neg \exists x F(x) \Rightarrow G(s)}{\neg \exists x F(x), F(t) \Rightarrow G(s)} \text{IL}
 \end{array} \right\} \pi_1$$

$$\left\{ \begin{array}{l}
 \frac{F(t) \Rightarrow F(t)}{F(t) \Rightarrow \boxed{\exists x F(x)}_1} \exists R \quad \frac{\frac{\frac{G(s) \Rightarrow G(s)}{\boxed{\exists x F(x)}_2, G(s) \Rightarrow G(s)} \text{WL}}{G(s), \boxed{\exists x F(x)}_1 \Rightarrow G(s)} \text{IL}}{\frac{F(t), G(s) \Rightarrow G(s)}{G(s), F(t) \Rightarrow G(s)} \text{IL}} \text{MIX} \\
 \frac{F(t), G(s) \Rightarrow G(s)}{G(s), F(t) \Rightarrow G(s)} \text{IL}
 \end{array} \right\} \pi_2$$

$$\frac{\frac{\neg \exists x F(x) \vee G(s), F(t) \Rightarrow G(s)}{F(t), \neg \exists x F(x) \vee G(s) \Rightarrow G(s)} \text{IL}}{\frac{F(t), \neg \exists x F(x) \vee G(s) \Rightarrow G(s)}{F(t), \neg \exists x F(x) \vee G(s) \Rightarrow G(s)} \forall L}$$

$$d(\pi_1) = d(\pi_2) = 1$$

$$rk(\pi_1) = rk_l(\pi_1) + rk_r(\pi_1) = 1 + 3 = 4$$

$$rk(\pi_2) = rk_l(\pi_2) + rk_r(\pi_2) = 1 + 2 = 3$$



# The Proof: How Does it Work?

We push our MIX inferences upwards again to reduce the ranks of our sub-proofs:

$$\pi_1 \left\{ \frac{\frac{\frac{F(t) \Rightarrow F(t)}{F(t) \Rightarrow \boxed{\exists x F(x)}_1} \exists R}{\frac{F(t), \neg \exists x F(x) \Rightarrow}{F(t), \neg \exists x F(x) \Rightarrow G(s)} \text{WR}}{\frac{F(t), \neg \exists x F(x) \Rightarrow G(s)}{\neg \exists x F(x), F(t) \Rightarrow G(s)} \text{IL}} \text{MIX}}{\frac{\frac{\boxed{\exists x F(x)}_2 \Rightarrow \exists x F(x)}{\neg \exists x F(x), \boxed{\exists x F(x)}_1 \Rightarrow} \neg L}{\frac{F(t), \neg \exists x F(x) \Rightarrow G(s)}{\neg \exists x F(x), F(t) \Rightarrow G(s)} \text{IL}} \text{MIX}} \text{WL}}{\frac{F(t) \Rightarrow F(t)}{F(t) \Rightarrow \boxed{\exists x F(x)}_1} \exists R}{\frac{F(t), G(s) \Rightarrow G(s)}{G(s), F(t) \Rightarrow G(s)} \text{IL}} \text{MIX}} \text{WL}} \right\} \pi_2$$

$$\frac{\frac{\frac{\frac{F(t), \neg \exists x F(x) \Rightarrow}{F(t), \neg \exists x F(x) \Rightarrow G(s)} \text{WR}}{\frac{F(t), \neg \exists x F(x) \Rightarrow G(s)}{\neg \exists x F(x), F(t) \Rightarrow G(s)} \text{IL}} \text{MIX}}{\frac{\frac{\boxed{\exists x F(x)}_2 \Rightarrow \exists x F(x)}{\neg \exists x F(x), \boxed{\exists x F(x)}_1 \Rightarrow} \neg L}{\frac{F(t), \neg \exists x F(x) \Rightarrow G(s)}{\neg \exists x F(x), F(t) \Rightarrow G(s)} \text{IL}} \text{MIX}} \text{WL}}{\frac{F(t) \Rightarrow F(t)}{F(t) \Rightarrow \boxed{\exists x F(x)}_1} \exists R}{\frac{F(t), G(s) \Rightarrow G(s)}{G(s), F(t) \Rightarrow G(s)} \text{IL}} \text{MIX}} \text{WL}} \text{IL}}$$

$$d(\pi_1) = d(\pi_2) = 1$$

$$rk(\pi_1) = rk_l(\pi_1) + rk_r(\pi_1) = 1 + 2 = 3$$

$$rk(\pi_2) = rk_l(\pi_2) + rk_r(\pi_2) = 1 + 1 = 2$$

# The Proof: How Does it Work?

Now we can eliminate one of the MIX inferences entirely and reduce the rank of the other:

$$\pi_1 \left\{ \frac{\frac{F(t) \Rightarrow F(t)}{F(t) \Rightarrow \boxed{\exists x F(x)}_1} \exists R \quad \boxed{\exists x F(x)}_1 \Rightarrow \exists x F(x)}{\frac{F(t) \Rightarrow \exists x F(x)}{\neg \exists x F(x), F(t) \Rightarrow} \neg L} \text{MIX} \quad \frac{G(s) \Rightarrow G(s)}{F(t), G(s) \Rightarrow G(s)} \text{WL}}{\frac{\frac{\frac{\neg \exists x F(x), F(t) \Rightarrow}{\neg \exists x F(x), F(t) \Rightarrow G(s)} \text{WR} \quad \frac{F(t), G(s) \Rightarrow G(s)}{G(s), F(t) \Rightarrow G(s)} \text{IL}}{G(s), F(t) \Rightarrow G(s)} \text{VL}}{\frac{\neg \exists x F(x) \vee G(s), F(t) \Rightarrow G(s)}{F(t), \neg \exists x F(x) \vee G(s) \Rightarrow G(s)} \text{IL}}$$

$$d(\pi_1) = 1$$

$$rk(\pi_1) = rk_l(\pi_1) + rk_r(\pi_1) = 1 + 1 = 2$$

# The Proof: How Does it Work?

Finally, we remove the remaining MIX inference.

$$\frac{\frac{\frac{F(t) \Rightarrow F(t)}{F(t) \Rightarrow \exists x F(x)} \exists R}{\neg \exists x F(x), F(t) \Rightarrow} \neg L}{\frac{\neg \exists x F(x), F(t) \Rightarrow G(s)}{\neg \exists x F(x) \vee G(s), F(t) \Rightarrow G(s)} \text{WR}} \text{IL} \quad \frac{\frac{G(s) \Rightarrow G(s)}{F(t), G(s) \Rightarrow G(s)} \text{WL}}{G(s), F(t) \Rightarrow G(s)} \text{IL}}{\frac{\neg \exists x F(x) \vee G(s), F(t) \Rightarrow G(s)}{F(t), \neg \exists x F(x) \vee G(s) \Rightarrow G(s)} \text{IL}} \vee L$$

# Peano Arithmetic

We work in the signature  $\{0, s, +, \cdot\}$ . We obtain the system **PA** by adding to LK a rule of inference for induction

$$! \frac{\varphi(a), \Gamma \Rightarrow \Theta, \varphi(s(a))}{\varphi(0), \Gamma \Rightarrow \Theta, \varphi(t)} \text{ CJ}$$

and admitting as axioms the following sequents:

$$\text{PA1 } \Rightarrow t = t$$

$$\text{PA2 } t_1 = t_2 \Rightarrow t_2 = t_1$$

$$\text{PA3 } t_1 = t_2, t_2 = t_3 \Rightarrow t_1 = t_3$$

$$\text{PA4 } s(t) = 0 \Rightarrow$$

$$\text{PA5 } t_1 = t_2 \Rightarrow s(t_1) = s(t_2)$$

$$\text{PA6 } s(t_1) = s(t_2) \Rightarrow t_1 = t_2$$

$$\text{PA7 } \Rightarrow t + 0 = t$$

$$\text{PA8 } \Rightarrow t_1 + s(t_2) = s(t_1 + t_2)$$

$$\text{PA9 } \Rightarrow t \cdot 0 = 0$$

$$\text{PA10 } \Rightarrow t_1 \cdot s(t_2) = s(t_1 \cdot t_2) + t_1$$

# The Big Idea

Our strategy to proving that **PA** is consistent will proceed from two ideas:

# The Big Idea

Our strategy to proving that **PA** is consistent will proceed from two ideas:

- 1 Sequents like  $\Rightarrow 5 + 2 = 2 + 5$  should always be provable without using quantifiers or induction. In fact, we should be able to produce proofs of these sorts of sequents without using any operational rules.

# The Big Idea

Our strategy to proving that **PA** is consistent will proceed from two ideas:

- 1 Sequents like  $\Rightarrow 5 + 2 = 2 + 5$  should always be provable without using quantifiers or induction. In fact, we should be able to produce proofs of these sorts of sequents without using any operational rules.
- 2 Proofs with sequents containing only atomic sentences should never have  $\Rightarrow 0 = 1$  as an end-sequent.

# Simple Proofs

## Definition

A proof in **PA** is simple if the only formulas occurring in it are atomic sentences and the only rules of inference occurring in it are structural.



# Simple Proofs

## Definition

A proof in **PA** is simple if the only formulas occurring in it are atomic sentences and the only rules of inference occurring in it are structural.

In the consistency proof for **PA**, simple proofs will be the analogues of CUT-free proofs:

## Definition

A proof in **PA** is simple if the only formulas occurring in it are atomic sentences and the only rules of inference occurring in it are structural.

In the consistency proof for **PA**, simple proofs will be the analogues of CUT-free proofs:

- In LK we could not find CUT-free proofs of both  $\varphi$  and  $\neg\varphi$  and in **PA** we will not be able to find a simple proof of  $0 = 1$ .

## Definition

A proof in **PA** is simple if the only formulas occurring in it are atomic sentences and the only rules of inference occurring in it are structural.

In the consistency proof for **PA**, simple proofs will be the analogues of CUT-free proofs:

- In LK we could not find CUT-free proofs of both  $\varphi$  and  $\neg\varphi$  and in **PA** we will not be able to find a simple proof of  $0 = 1$ .
- In LK we could transform any proof into a CUT-free proof of the same end-sequent and in **PA** we will be able to transform any proof with end-sequent containing only atomic sentences into a simple proof of the same end-sequent.

# Consistency of Simple Proofs

## Definition

Define the value  $val(t)$  of a closed term  $t$  recursively as follows:

- $val(0) = 0$
- $val(s(t)) = val(t) + 1$
- $val(r + t) = val(r) + val(t)$
- $val(r \cdot t) = val(r) \cdot val(t)$

An atomic sentence  $r = t$  is true if  $val(r) = val(t)$ , and false otherwise. A sequent containing only atomic sentences is true if there is a false sentence in its antecedent or a true sentence in its succedent.

## Lemma

*Every sequent in a simple proof is true. In particular, there is no simple proof with end-sequent  $\Rightarrow 0 = 1$ .*

# Existence of Certain Simple Proofs

The following will be useful when removing CJ inferences from our proofs:

## Lemma

*Let  $t$  be a closed term and  $val(t) = n$ . Then there is a simple proof with end-sequent  $\Rightarrow t = n$ .*

## Proof (Sketch).

Induction. Lots and lots of induction. The proof is made considerably easier if we add as axioms to **PA** the sequents

$$r = t, u = v \Rightarrow r + u = t + v \text{ and } r = t, u = v \Rightarrow r \cdot u = t \cdot v.$$

If the terms involved in these sequents are taken to be closed, then it is easy to verify that the sequents are true and the lemma above regarding the consistency of simple proofs remains true.  $\square$

# The Almost Simple Part

Given a proof in **PA** with end-sequent containing only atomic sentences there will be part of the proof that “looks like a simple proof.” This part of the proof will contain

- The end-sequent.
- Some other stuff immediately above the end-sequent which “looks simple.”

Our goal is to transform the proof so that this part expands to include the entire proof. We make precise the notion of “looks like a simple proof.”

# Successors and Predecessors

## Definition

Let  $I$  be an inference in a proof  $\pi$ ,  $\varphi$  be a formula appearing in the conclusion of  $I$ , and  $\psi$  be a formula appearing in a premise of  $I$ . Then  $\varphi$  is the successor of  $\psi$  (and  $\psi$  is a predecessor of  $\varphi$ ) if one of the following holds:

- Both  $\varphi$  and  $\psi$  are corresponding occurrences of the same formula appearing in the side formulas of  $I$ .
- $I$  is an operational inference with principal formula  $\varphi$  and auxiliary formula  $\psi$ .
- $I$  is either a contraction or interchange and  $\varphi$  and  $\psi$  are the same formula.
- $I$  is a CJ inference and  $\varphi$  and  $\psi$  are the principal and auxiliary formulas in the succedents (antecedents), respectively.

# Bundles and Boundaries

## Definition

Let  $\pi$  be a proof and  $\varphi_0, \dots, \varphi_n$  be a sequence of formulas appearing in  $\pi$  such that  $\varphi_0$  has no predecessor,  $\varphi_{i+1}$  is the successor of  $\varphi_i$  for  $i < n$ , and  $\varphi_n$  has no successor. Such a sequence is called a bundle. A bundle  $\varphi_0, \dots, \varphi_n$  is implicit if  $\varphi_n$  is a CUT-formula and explicit otherwise. An inference in  $\pi$  is implicit if its principal formula belongs to an implicit bundle and is explicit otherwise. Lowermost implicit operational inferences are called boundary inferences.

Note that if  $\varphi_0, \dots, \varphi_n$  is an explicit bundle, then  $\varphi_n$  is a formula appearing in the end-sequent of  $\pi$ .



# The End-part of a Proof

## Definition

The end-part of a proof is the smallest part of the proof such that:

# The End-part of a Proof

## Definition

The end-part of a proof is the smallest part of the proof such that:

- 1 The end-sequent belongs to the end-part.

# The End-part of a Proof

## Definition

The end-part of a proof is the smallest part of the proof such that:

- 1 The end-sequent belongs to the end-part.
- 2 If the conclusion of an inference belongs to the end-part, so do the premises, unless the inference is a boundary inference.

# The End-part of a Proof

## Definition

The end-part of a proof is the smallest part of the proof such that:

- 1 The end-sequent belongs to the end-part.
- 2 If the conclusion of an inference belongs to the end-part, so do the premises, unless the inference is a boundary inference.

In the case of proofs with end-sequents containing only atomic sentences, all operational inferences must be implicit and so boundary inferences turn out to be lowermost operational inferences. Thus the end-part of such a proof “looks simple”: it contains only structural and CJ inferences.

# Expanding the End-part

We will successively apply three different reduction steps to expand the end-parts our proofs. These are:

- 1 Removal of CJ inferences
- 2 Removal of weakenings
- 3 Reduction of CUTs

The second sort of reduction is needed for technical reasons which we will comment on later. For now we focus our attention on steps 1 and 2.

# Removing CJ Inferences

Suppose  $\pi$  is a proof with end-sequent containing only atomic sentences and that there is a CJ inference in the end-part of  $\pi$ . Then there is a lowermost such CJ inference and we may assume that its principal formula  $\varphi(t)$  is a sentence. So  $t$  is a closed term and  $val(t)$  is defined. Suppose  $val(t) = n$  and consider the sub-proof of  $\pi$  ending in this lowermost CJ inference of the end-part:

$$\begin{array}{c} \vdots \\ \vdots \pi'(a) \\ \vdots \\ \varphi(a) \Rightarrow \varphi(s(a)) \\ \hline \varphi(0) \Rightarrow \varphi(t) \end{array} \text{ CJ}$$

We will replace the above sub-proof with what are essentially several copies of the proof above the CJ inference.

# Removing CJ Inferences

Consider the following proof  $\pi''(n)$ :

$$\begin{array}{c}
 \begin{array}{c} \vdots \\ \pi'(0) \\ \vdots \\ \varphi(0) \Rightarrow \varphi(1) \end{array} \quad \begin{array}{c} \vdots \\ \pi'(1) \\ \vdots \\ \varphi(1) \Rightarrow \varphi(2) \end{array} \\
 \hline
 \varphi(0) \Rightarrow \varphi(2) \quad \text{CUT} \\
 \hline
 \begin{array}{c} \vdots \\ \pi'(3) \\ \vdots \\ \varphi(2) \Rightarrow \varphi(3) \end{array} \\
 \hline
 \varphi(0) \Rightarrow \varphi(3) \quad \text{CUT} \\
 \hline
 \begin{array}{c} \vdots \\ \vdots \\ \vdots \\ \varphi(0) \Rightarrow \varphi(n-1) \end{array} \quad \begin{array}{c} \vdots \\ \pi'(n) \\ \vdots \\ \varphi(n-1) \Rightarrow \varphi(n) \end{array} \\
 \hline
 \varphi(0) \Rightarrow \varphi(n) \quad \text{CUT}
 \end{array}$$

Since  $val(t) = n$ , there is a simple proof of  $\Rightarrow t = n$  and we can use this to build a proof of  $\varphi(n) \Rightarrow \varphi(t)$  without CJ inferences or complex CUTs. Taking a CUT of this with  $\pi''(n)$  gives us a way to remove the CJ inference.

# Removing CJ Inferences

$$\frac{\begin{array}{c} \vdots \\ \vdots \pi''(n) \\ \vdots \end{array} \quad \varphi(0) \Rightarrow \varphi(n) \quad \varphi(n) \Rightarrow \varphi(t) \quad \begin{array}{c} \vdots \\ \vdots \end{array}}{\varphi(0) \Rightarrow \varphi(t)} \text{ CUT}$$

So if we replace

$$! \frac{\begin{array}{c} \vdots \\ \vdots \pi'(a) \\ \vdots \end{array} \quad \varphi(a) \Rightarrow \varphi(s(a))}{\varphi(0) \Rightarrow \varphi(t)} \text{ CJ}$$

with the proof just constructed, we remove a CJ inference. But what else happens?



# Removing CJ Inferences

Looking at  $\pi''(n)$  again we see that we have essentially introduced  $n + 1$  new copies of  $\pi'(a)$  to our proof.

$$\begin{array}{c}
 \begin{array}{c} \vdots \\ \pi'(0) \\ \vdots \end{array} \quad \begin{array}{c} \vdots \\ \pi'(1) \\ \vdots \end{array} \\
 \hline
 \varphi(0) \Rightarrow \varphi(1) \quad \varphi(1) \Rightarrow \varphi(2) \quad \text{CUT} \\
 \hline
 \varphi(0) \Rightarrow \varphi(2)
 \end{array}
 \quad
 \begin{array}{c}
 \begin{array}{c} \vdots \\ \pi'(3) \\ \vdots \end{array} \\
 \hline
 \varphi(2) \Rightarrow \varphi(3) \quad \text{CUT} \\
 \hline
 \varphi(0) \Rightarrow \varphi(3)
 \end{array}
 \\
 \vdots \\
 \begin{array}{c} \varphi(0) \Rightarrow \varphi(n-1) \end{array}
 \quad
 \begin{array}{c}
 \begin{array}{c} \vdots \\ \pi'(n) \\ \vdots \end{array} \\
 \hline
 \varphi(n-1) \Rightarrow \varphi(n) \quad \text{CUT} \\
 \hline
 \varphi(0) \Rightarrow \varphi(n)
 \end{array}
 \end{array}$$

So any part of the end-part of  $\pi$  contained in  $\pi'(a)$  is reproduced  $n + 1$  times and we may end up introducing more CJ inferences to the end-part of our proof than we remove.

# Removing CJ Inferences

## Definition

Let  $\pi$  be a proof with end-sequent containing only atomic sentences. An induction chain in  $\pi$  is a sequence  $I_0, \dots, I_k$  where each  $I_j$  is a CJ inference in the end-part of  $\pi$ ,  $I_{j+1}$  occurs below  $I_j$ , no CJ inference occurs between  $I_{j+1}$  and  $I_j$ , and no CJ inference occurs below  $I_k$ . We take  $m(\pi)$  to be the maximum length among induction chains in  $\pi$  and  $o(\pi)$  to be the number of induction chains in  $\pi$  of length  $m(\pi)$ .

In the procedure described above for removing CJ inferences we can always choose to remove a CJ inference which is the last term of an induction chain of maximal length. When we do this, any new induction chains introduced are of length strictly shorter so either the number of induction chains of maximal length decreases by 1, or the maximal length of induction chains in  $\pi$  decreases.

# Reduction of CUTs

After applying the first two reduction steps it is not difficult to show that if our proof is not simple, then the end-part contains a complex CUT where both CUT-formulas descend from boundary inferences. For instance, if the CUT-formula is  $\neg\varphi$ , then sub-proof ending in the complex CUT would look like the following where dashed lines indicate boundary inferences:

$$\begin{array}{ccc} \begin{array}{c} \vdots \\ \pi_1 \\ \vdots \\ \varphi \Rightarrow \\ \hline \neg\varphi \end{array} & \neg R & \begin{array}{c} \vdots \\ \pi_2 \\ \vdots \\ \neg\varphi \Rightarrow \\ \hline \varphi \end{array} \\ \begin{array}{c} \vdots \\ \pi'_1 \\ \vdots \\ \Rightarrow \neg\varphi \end{array} & \Rightarrow & \begin{array}{c} \vdots \\ \pi'_2 \\ \vdots \\ \neg\varphi \Rightarrow \\ \hline \text{CUT} \end{array} \end{array}$$

# Reduction of CUTs

We replace the sub-proof above with the following:

$$\begin{array}{c}
 \begin{array}{c} \vdots \\ \pi_1 \\ \Rightarrow \\ \neg\varphi \\ \Rightarrow \\ \neg\neg\varphi \end{array} \neg R \quad \begin{array}{c} \vdots \\ \pi_2 \\ \Rightarrow \\ \varphi \\ \Rightarrow \\ \neg\varphi \end{array} \text{WL} \\
 \hline
 \begin{array}{c} \vdots \\ \pi'_1 \\ \Rightarrow \\ \neg\varphi \end{array} \quad \begin{array}{c} \vdots \\ \pi''_2 \\ \Rightarrow \\ \varphi \end{array} \text{CUT} \\
 \hline
 \Rightarrow \varphi
 \end{array}
 \quad
 \begin{array}{c}
 \begin{array}{c} \vdots \\ \pi_1 \\ \Rightarrow \\ \varphi \\ \Rightarrow \\ \neg\neg\varphi \end{array} \text{WR} \quad \begin{array}{c} \vdots \\ \pi_2 \\ \Rightarrow \\ \neg\varphi \\ \Rightarrow \\ \neg\neg\varphi \end{array} \neg L \\
 \hline
 \begin{array}{c} \vdots \\ \pi''_1 \\ \Rightarrow \\ \varphi \\ \Rightarrow \\ \neg\varphi \end{array} \quad \begin{array}{c} \vdots \\ \pi'_2 \\ \Rightarrow \\ \neg\varphi \end{array} \text{CUT} \\
 \hline
 \varphi \Rightarrow \quad \text{CUT} \\
 \hline
 \Rightarrow
 \end{array}$$

# Reduction of CUTs

We replace the sub-proof above with the following:

$$\begin{array}{c}
 \begin{array}{c} \vdots \\ \pi_1 \\ \Rightarrow \\ \neg\varphi \end{array} \xRightarrow{\neg R} \begin{array}{c} \vdots \\ \neg\neg\varphi \end{array} \quad \begin{array}{c} \vdots \\ \pi_2 \\ \Rightarrow \\ \varphi \end{array} \xRightarrow{WL} \begin{array}{c} \vdots \\ \neg\varphi \Rightarrow \varphi \end{array} \quad \begin{array}{c} \vdots \\ \pi_1 \\ \Rightarrow \\ \varphi \end{array} \xRightarrow{WR} \begin{array}{c} \vdots \\ \varphi \Rightarrow \neg\varphi \end{array} \quad \begin{array}{c} \vdots \\ \pi_2 \\ \Rightarrow \\ \neg\varphi \end{array} \xRightarrow{\neg L} \begin{array}{c} \vdots \\ \neg\neg\neg\varphi \end{array} \\
 \begin{array}{c} \vdots \\ \pi'_1 \\ \Rightarrow \\ \neg\varphi \end{array} \quad \begin{array}{c} \vdots \\ \pi''_2 \\ \Rightarrow \\ \neg\varphi \end{array} \quad \begin{array}{c} \vdots \\ \pi''_1 \\ \Rightarrow \\ \varphi \end{array} \quad \begin{array}{c} \vdots \\ \pi'_2 \\ \Rightarrow \\ \neg\varphi \end{array} \\
 \hline
 \begin{array}{c} \Rightarrow \\ \varphi \end{array} \quad \text{CUT} \quad \begin{array}{c} \Rightarrow \\ \varphi \end{array} \quad \text{CUT} \\
 \hline
 \Rightarrow
 \end{array}$$

But this doesn't look particularly helpful: we now have three CUT inferences and only one is of lower degree than what we started with.

# Reduction of CUTs

We replace the sub-proof above with the following:

$$\begin{array}{c}
 \begin{array}{c} \vdots \\ \pi_1 \\ \Rightarrow \neg\varphi \end{array} \neg R \quad \begin{array}{c} \vdots \\ \pi_2 \\ \Rightarrow \varphi \end{array} \text{WL} \quad \begin{array}{c} \vdots \\ \pi_1 \\ \Rightarrow \varphi \\ \varphi \Rightarrow \neg\varphi \end{array} \text{WR} \quad \begin{array}{c} \vdots \\ \pi_2 \\ \Rightarrow \neg\varphi \end{array} \neg L \\
 \hline
 \begin{array}{c} \vdots \\ \pi'_1 \\ \Rightarrow \neg\varphi \end{array} \text{CUT} \quad \begin{array}{c} \vdots \\ \pi''_2 \\ \neg\varphi \Rightarrow \varphi \end{array} \text{CUT} \quad \begin{array}{c} \vdots \\ \pi''_1 \\ \varphi \Rightarrow \neg\varphi \end{array} \text{CUT} \quad \begin{array}{c} \vdots \\ \pi'_2 \\ \neg\varphi \Rightarrow \neg\varphi \end{array} \text{CUT} \\
 \hline
 \begin{array}{c} \Rightarrow \varphi \end{array} \quad \begin{array}{c} \varphi \Rightarrow \end{array} \\
 \hline
 \Rightarrow
 \end{array}$$

But this doesn't look particularly helpful: we now have three CUT inferences and only one is of lower degree than what we started with. However, now we have brought parts of both  $\pi_1$  and  $\pi_2$  into the end-part of our proof, thus raising the boundary and allowing us to apply our other reduction procedures to more of the proof.

# Reduction of CUTs

We replace the sub-proof above with the following:

$$\begin{array}{c}
 \begin{array}{c} \vdots \\ \pi_1 \\ \vdots \\ \neg\varphi \Rightarrow \\ \hline \neg\varphi \\ \hline \Rightarrow \neg\varphi \end{array} \neg\text{R} \quad \begin{array}{c} \vdots \\ \pi_2 \\ \vdots \\ \Rightarrow \varphi \\ \hline \cancel{\neg\varphi} \Rightarrow \varphi \end{array} \cancel{\text{WL}} \\
 \hline
 \begin{array}{c} \vdots \\ \pi'_1 \\ \vdots \\ \Rightarrow \neg\varphi \end{array} \quad \begin{array}{c} \vdots \\ \pi''_2 \\ \vdots \\ \cancel{\neg\varphi} \Rightarrow \varphi \end{array} \text{CUT} \\
 \hline
 \Rightarrow \varphi
 \end{array}
 \quad \Rightarrow \quad
 \begin{array}{c}
 \begin{array}{c} \vdots \\ \pi_1 \\ \vdots \\ \varphi \Rightarrow \\ \hline \varphi \Rightarrow \cancel{\neg\varphi} \end{array} \cancel{\text{WR}} \quad \begin{array}{c} \vdots \\ \pi_2 \\ \vdots \\ \Rightarrow \varphi \\ \hline \neg\varphi \Rightarrow \varphi \end{array} \neg\text{L} \\
 \hline
 \begin{array}{c} \vdots \\ \pi''_1 \\ \vdots \\ \varphi \Rightarrow \cancel{\neg\varphi} \end{array} \quad \begin{array}{c} \vdots \\ \pi'_2 \\ \vdots \\ \neg\varphi \Rightarrow \varphi \end{array} \text{CUT} \\
 \hline
 \varphi \Rightarrow \text{CUT}
 \end{array}$$

But this doesn't look particularly helpful: we now have three CUT inferences and only one is of lower degree than what we started with. However, now we have brought parts of both  $\pi_1$  and  $\pi_2$  into the end-part of our proof, thus raising the boundary and allowing us to apply our other reduction procedures to more of the proof.

# Reduction of CUTs

We replace the sub-proof above with the following:

$$\frac{\begin{array}{c} \vdots \\ \pi_2 \\ \vdots \\ \Rightarrow \varphi \\ \hline \cancel{\neg\varphi} \Rightarrow \varphi \end{array} \text{WL}}{\begin{array}{c} \vdots \\ \pi_2'' \\ \vdots \\ \cancel{\neg\varphi} \Rightarrow \varphi \\ \hline \Rightarrow \varphi \end{array} \text{CUT}} \Rightarrow \frac{\begin{array}{c} \vdots \\ \pi_1 \\ \vdots \\ \varphi \Rightarrow \\ \hline \varphi \Rightarrow \cancel{\neg\varphi} \end{array} \text{WR}}{\begin{array}{c} \vdots \\ \pi_1'' \\ \vdots \\ \varphi \Rightarrow \cancel{\neg\varphi} \\ \hline \varphi \Rightarrow \end{array} \text{CUT}} \text{CUT}$$

But this doesn't look particularly helpful: we now have three CUT inferences and only one is of lower degree than what we started with. However, now we have brought parts of both  $\pi_1$  and  $\pi_2$  into the end-part of our proof, thus raising the boundary and allowing us to apply our other reduction procedures to more of the proof.



# Please, Make it Stop

Once again we need to show that if we continue applying our reduction procedures, then we will eventually reach a proof to which our reduction procedures cannot be applied. Such a proof will be a simple proof and our argument will be complete. We induct on the ordinal  $\varepsilon_0$ :

$$\varepsilon_0 = \sup\{\omega_n(0) : n < \omega\}$$

where  $\omega_n(\alpha)$  is given by

$$\omega_0(\alpha) = \alpha$$

$$\omega_{n+1}(\alpha) = \omega^{\omega_n(\alpha)}$$

for  $\alpha$  an ordinal. So  $\varepsilon_0$  is the limit of the sequence  $0, \omega^0, \omega^{\omega^0}, \omega^{\omega^{\omega^0}}, \dots$

# Please, Make it Stop

## Definition

Let  $\pi$  be a proof in **PA**.

# Please, Make it Stop

## Definition

Let  $\pi$  be a proof in **PA**.

- 1 If  $S$  is an initial sequent of  $\pi$ , then  $o(S, \pi) = 1$ .

# Please, Make it Stop

## Definition

Let  $\pi$  be a proof in **PA**.

- 1 If  $S$  is an initial sequent of  $\pi$ , then  $o(S, \pi) = 1$ .
- 2 If  $I$  is structural inference other than CUT in  $\pi$  with premise  $S$ , then  $o(I, \pi) = o(S, \pi)$ .

## Definition

Let  $\pi$  be a proof in **PA**.

- 1 If  $S$  is an initial sequent of  $\pi$ , then  $o(S, \pi) = 1$ .
- 2 If  $I$  is structural inference other than CUT in  $\pi$  with premise  $S$ , then  $o(I, \pi) = o(S, \pi)$ .
- 3 If  $I$  is an operational inference with one premise  $S$ , then  $o(I, \pi) = o(S, \pi) + 1$ .

# Please, Make it Stop

## Definition

Let  $\pi$  be a proof in **PA**.

- 1 If  $S$  is an initial sequent of  $\pi$ , then  $o(S, \pi) = 1$ .
- 2 If  $I$  is structural inference other than CUT in  $\pi$  with premise  $S$ , then  $o(I, \pi) = o(S, \pi)$ .
- 3 If  $I$  is an operational inference with one premise  $S$ , then  $o(I, \pi) = o(S, \pi) + 1$ .
- 4 If  $I$  is an inference with two premises  $S, S'$ , then  $o(I, \pi) = \max\{o(S, \pi), o(S', \pi)\} + 1$ .

# Please, Make it Stop

## Definition

Let  $\pi$  be a proof in **PA**.

- 1 If  $S$  is an initial sequent of  $\pi$ , then  $o(S, \pi) = 1$ .
- 2 If  $I$  is structural inference other than CUT in  $\pi$  with premise  $S$ , then  $o(I, \pi) = o(S, \pi)$ .
- 3 If  $I$  is an operational inference with one premise  $S$ , then  $o(I, \pi) = o(S, \pi) + 1$ .
- 4 If  $I$  is an inference with two premises  $S, S'$ , then  $o(I, \pi) = \max\{o(S, \pi), o(S', \pi)\} + 1$ .
- 5 If  $I$  is a CJ inference with premise  $S$  and  $o(S, \pi)$  has Cantor Normal Form  $\omega^{\alpha_0} + \dots + \omega^{\alpha_n}$ , then  $o(I, \pi) = \omega^{\alpha_0+1}$ .

# Please, Make it Stop

## Definition

Let  $\pi$  be a proof in **PA**.

- 1 If  $S$  is an initial sequent of  $\pi$ , then  $o(S, \pi) = 1$ .
- 2 If  $I$  is structural inference other than CUT in  $\pi$  with premise  $S$ , then  $o(I, \pi) = o(S, \pi)$ .
- 3 If  $I$  is an operational inference with one premise  $S$ , then  $o(I, \pi) = o(S, \pi) + 1$ .
- 4 If  $I$  is an inference with two premises  $S, S'$ , then  $o(I, \pi) = \max\{o(S, \pi), o(S', \pi)\} + 1$ .
- 5 If  $I$  is a CJ inference with premise  $S$  and  $o(S, \pi)$  has Cantor Normal Form  $\omega^{\alpha_0} + \dots + \omega^{\alpha_n}$ , then  $o(I, \pi) = \omega^{\alpha_0+1}$ .
- 6 If  $S$  is the conclusion of inference  $I$ , and  $l$  and  $k$  are the maximum degrees of all CUTs and CJ inferences below  $S$  and the premises of  $I$ , respectively, then  $o(S, \pi) = \omega_{k-l}(o(I, \pi))$ .



# Please, Make it Stop

For  $\pi$  a proof with end-sequent  $S$  we take  $o(\pi) = o(S, \pi)$ . All that is left to do now is show that whenever we apply one of our reduction procedures, the ordinal corresponding to our proof decreases. But this is just bookkeeping.

# References

Paolo Mancosu, Sergio Galvan, and Richard Zach. *An Introduction to Proof Theory: Normalization, Cut-Elimination, and Consistency Proofs*. Oxford University Press, Oxford, 2021.