Introduction to Proof Theory

Lecture 4
Cut-elimination

Marianna Girlando, Sonia Marin

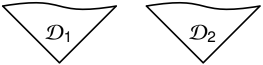
Universtiy of Amsterdam, University of Birmingham

OPLSS 2023

Eugene, Oregon, June 26 - July 8, 2023

Outline

Today’s goal



$$\text{cut}\ \frac{\Gamma \Rightarrow A \qquad A, \Gamma \Rightarrow C}{\Gamma \Rightarrow C}$$

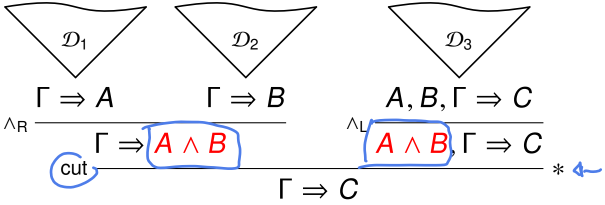
Theorem (*Hauptsatz*, Gentzen 1935)

*Every theorem of **LJ** has a proof that does not use the cut rule.*

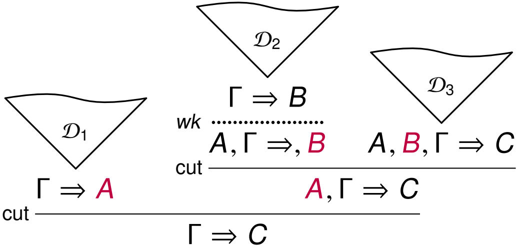
Corollary (Analyticity)

*Every theorem of **LJ** has a proof that contains only subformulas of it.*

Informal example

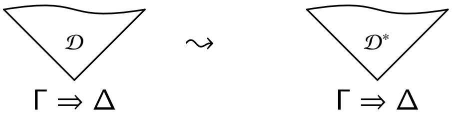


Let's eliminate the occurrence of cut marked by $*$



General strategy of the proof

LJ derivation $\rightsquigarrow$ cut-free **LJ** derivation



- Apply the cut on smaller formulas, until they disappear!
- Push the cuts upwards in the proof, until they disappear!
- We need a "measure" on formulas and on derivations, to ensure that the cut-elimination procedure terminates.

. . . The cut-elimination proof is quite complex.

We are going to sketch the proof for **LJ**.

References

Several proofs of cut-elimination exist in the literature, using slightly different procedures and for slightly different systems:

- [Buss, 1998]. *Handbook of Proof Theory*.
- [Troelstra and Schwichtenberg, 1996]. *Basic Proof Theory*.
- [Negri and von Plato, 2001]. *Structural Proof Theory*.
- ...

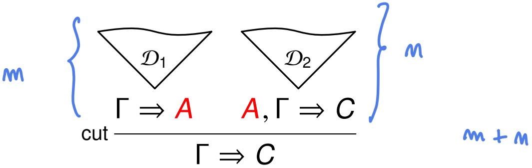
Measuring height

The height of $\mathcal{D}$, $\mathrm{ht}(\mathcal{D})$, is the length of its longest branch, minus one.

Measuring height

The height of $\mathcal{D}$, $\mathrm{ht}(\mathcal{D})$, is the length of its longest branch, minus one.

The level of a cut rule is the sum of heights of derivations of the two premisses of cut.



Measuring degree

The degree of a formula A, $\mathrm{deg}(A)$, is the number of logical connectives occurring in it.

$$\begin{aligned} \mathrm{deg}(p) &:= 0 \\ \mathrm{deg}(\bot) &:= 0 \\ \mathrm{deg}(A \star B) &:= \mathrm{deg}(A) + \mathrm{deg}(B) + 1 \\ &\text{for } \star \in \{\vee, \wedge, \rightarrow\} \end{aligned}$$

Measuring degree

The degree of a formula A, $\text{deg}(A)$, is the number of logical connectives occurring in it.

$$\begin{aligned}
\text{deg}(p) &:= 0 \\
\text{deg}(\bot) &:= 0 \\
\text{deg}(A \star B) &:= \text{deg}(A) + \text{deg}(B) + 1 \\
&\text{for } \star \in \{\vee, \wedge, \rightarrow\}
\end{aligned}$$

The rank of a cut rule is the degree of the cut formula A, plus 1.

$$\text{cut}\,\frac{\Gamma \Rightarrow A \quad A, \Gamma \Rightarrow C}{\Gamma \Rightarrow C}$$

Measuring degree

The degree of a formula A, $\text{deg}(A)$, is the number of logical connectives occurring in it.

$$
\begin{aligned}
\text{deg}(p) &:= 0 \\
\text{deg}(\bot) &:= 0 \\
\text{deg}(A \star B) &:= \text{deg}(A) + \text{deg}(B) + 1 \\
&\quad \text{for } \star \in \{\vee, \wedge, \rightarrow\}
\end{aligned}
$$

The rank of a cut rule is the degree of the cut formula A, plus 1.

$$
\text{cut}\ \frac{\Gamma \Rightarrow A \quad A, \Gamma \Rightarrow C}{\Gamma \Rightarrow C}
$$

The rank of $\mathcal{D}$, $\text{rk}(\mathcal{D})$, is the maximum of the rank of the cut formulas occurring in $\mathcal{D}$.

Measuring degree

The degree of a formula A, $\mathrm{deg}(A)$, is the number of logical connectives occurring in it.

$$
\begin{aligned}
\mathrm{deg}(p) &:= 0 \\
\mathrm{deg}(\bot) &:= 0 \\
\mathrm{deg}(A \star B) &:= \mathrm{deg}(A) + \mathrm{deg}(B) + 1 \\
&\text{for } \star \in \{\vee, \wedge, \rightarrow\}
\end{aligned}
$$

The rank of a cut rule is the degree of the cut formula A, plus 1.

$$
\text{cut}\ \frac{\Gamma \Rightarrow A \quad A, \Gamma \Rightarrow C}{\Gamma \Rightarrow C}
$$

The rank of $\mathcal{D}$, $\mathrm{rk}(\mathcal{D})$, is the maximum of the rank cut formulas occurring in $\mathcal{D}$.

$\Gamma \Rightarrow_p^m C$ means *there is a derivation of* $\Gamma \Rightarrow C$ *of height* ***at most*** *m and rank* ***at most*** *p.*

Rank of derivations (more formally)

Height and rank can be inductively defined on the structure of $\mathcal{D}$:

$$\mathcal{D} = \text{init}\,\dfrac{}{\Gamma \Rightarrow C} \qquad \mathsf{rk}(\mathcal{D}) = 0$$

$$\mathcal{D} = \text{R}\,\dfrac{\overset{\mathcal{D}_1}{\Gamma_1 \Rightarrow C_1}}{\Gamma \Rightarrow C} \qquad \mathsf{rk}(\mathcal{D}) = \mathsf{rk}(\mathcal{D}_1)$$

$$\mathcal{D} = \text{R}\,\dfrac{\overset{\mathcal{D}_1}{\Gamma_1 \Rightarrow C_1} \qquad \overset{\mathcal{D}_2}{\Gamma_2 \Rightarrow C_2}}{\Gamma \Rightarrow C} \qquad \mathsf{rk}(\mathcal{D}) = \max(\mathsf{rk}(\mathcal{D}_1), \mathsf{rk}(\mathcal{D}_2))$$

$$\mathcal{D} = \text{cut}\,\dfrac{\overset{\mathcal{D}_1}{\Gamma \Rightarrow A} \qquad \overset{\mathcal{D}_2}{A, \Gamma \Rightarrow C}}{\Gamma \Rightarrow C} \qquad \mathsf{rk}(\mathcal{D}) = \max(\mathsf{rk}(\mathcal{D}_1), \mathsf{rk}(\mathcal{D}_2), \mathsf{deg}(A) + 1)$$

Some preliminary lemmas

$\Gamma \Rightarrow C$
$\vdots$
$\Gamma', \Gamma \Rightarrow C$

$$\frac{\Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}\ wk$$

$$\frac{A, A, \Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}\ ctr$$

1. Lemma: Closure under weakening
   - If $\Gamma \Rightarrow_p^m C$, then $\Gamma', \Gamma \Rightarrow_p^m C$, for any Γ'.

   Proof. Easy induction on the height m, of the derivation.

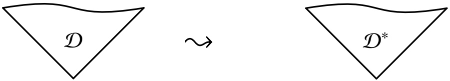
2. Lemma: Closure under contraction
   - If $A, A, \Gamma \Rightarrow_p^m A$ (handwritten correction: C), then $A, \Gamma \Rightarrow_p^m C$.

   Proof. Induction on m, using weakening (and invertibility)

NB: all the above preserve height and rank of the derivation.

The plan

LJ derivation $\rightsquigarrow$ cut-free **LJ** derivation

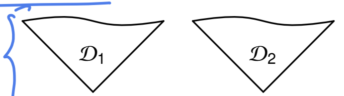


$\Gamma \Rightarrow_p C$ $\qquad$ $\Gamma \Rightarrow_0 C$

- We show how to simulate instances of cut. Lemma
- We show how to eliminate all cuts occurring in a derivation, starting with topmost cuts having maximal rank. Theorem

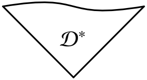
Lemma (closure under cut)

If $\Gamma \Rightarrow_0^m A$ and $A, \Gamma \Rightarrow_0^n C$



$$\text{cut}\ \frac{\Gamma \Rightarrow_0^m A \qquad A, \Gamma \Rightarrow_0^n C}{\Gamma \Rightarrow_1 C}$$

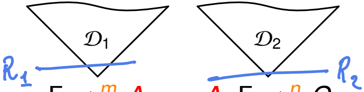
then we can construct the following derivation $\mathcal{D}^*$:



$$\Gamma \Rightarrow_0 C$$

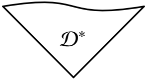
Lemma (closure under cut)

If $\Gamma \Rightarrow_0^m A$ and $A, \Gamma \Rightarrow_0^n C$



$$\dfrac{\Gamma \Rightarrow_0^m A \qquad A, \Gamma \Rightarrow_0^n C}{\Gamma \Rightarrow_1 C}\ \text{cut}$$

then we can construct the following derivation $\mathcal{D}^*$:



$$\Gamma \Rightarrow_0 C$$

Proof. Induction on $(\deg(A), m + n)$. We distinguish cases:

1. R_1 is init (R$_2$ is init)
2. A is principal in both R_1 and R_2
3. A is not principal in R_1 (A is not principal in R_2)

R_1 is init

$$\mathcal{D}_1 = \text{init}\ \frac{}{A, \Gamma' \Rightarrow_p^m A} \qquad \text{R}_2\ \frac{\begin{matrix}\mathcal{D}_2 \\ \Gamma'' \Rightarrow_p^{n-1} C''\end{matrix}}{A, \Gamma \Rightarrow_p^n C}$$

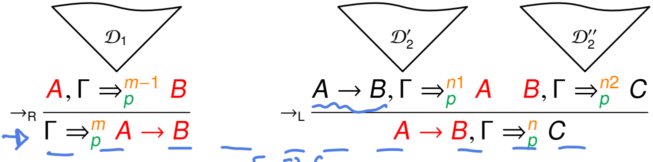
with $\Gamma = A, \Gamma'$. We construct the following derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p \Delta$:

$$\begin{array}{c} \mathcal{D}_2 \\ \text{R}_2\ \dfrac{\Gamma'' \Rightarrow C''}{A, A, \Gamma' \Rightarrow C} \\ ctr \cdots\cdots\cdots\cdots \\ A, \Gamma' \Rightarrow C \end{array}$$

A is principal in both R_1 and R_2

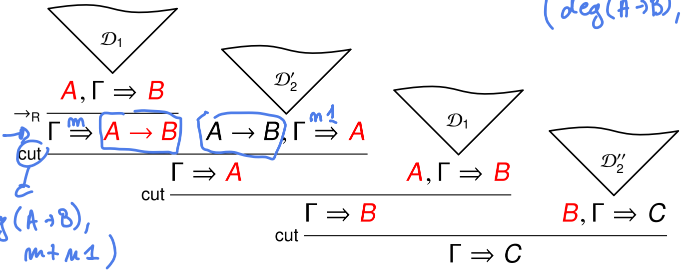
$n = c$

R_1 is $\rightarrow_R$ and R_2 is $\rightarrow_L$



with $n1, n2 < n$. We construct the following derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p \Delta$:

$(deg(A \rightarrow B), m + n)$



$(deg(A \rightarrow B), m + n1)$

$m + n1 < m + n$

Cases for the other rules . . .

A is not principal in R_1

R_1 is a one-premiss rule

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ \vdots\\ \Gamma' \Rightarrow_p^{m-1} A\end{array}}{\Gamma \Rightarrow_p^{m} A}\,R_1 \qquad\qquad \dfrac{\begin{array}{c}\mathcal{D}_2\\ \vdots\\ \Gamma'' \Rightarrow_p^{n-1} C''\end{array}}{A, \Gamma \Rightarrow_p^{n} C}\,R_2$$

We construct the following derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p \Delta$:

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_1\\ \vdots\\ \Gamma' \Rightarrow, A\end{array}}{\Gamma', \Gamma \Rightarrow A}\,wk \qquad \dfrac{\dfrac{\begin{array}{c}\mathcal{D}_2\\ \vdots\\ \Gamma'' \Rightarrow C''\end{array}}{A, \Gamma \Rightarrow C}\,R_2}{A, \Gamma', \Gamma \Rightarrow C}\,wk}{\dfrac{\Gamma', \Gamma \Rightarrow C}{\dfrac{\Gamma, \Gamma \Rightarrow C}{\Gamma \Rightarrow C}\,ctr}\,R_1}\,cut$$

R_1 is a two-premisses rule ...

End of the proof

Eliminating cut

Cut-elimination Theorem If we have a derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p C$, we can construct a derivation $\mathcal{D}^*$ of $\Gamma \Rightarrow_0 C$, that is, a derivation **where** cut **does not occur.**

Eliminating cut

Cut-elimination Theorem If we have a derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p C$, we can construct a derivation $\mathcal{D}^*$ of $\Gamma \Rightarrow_0 C$, that is, a derivation **where** cut **does not occur.**

Proof. We apply the proof transformation detailed in the Lemma to the cuts occurring in $\mathcal{D}$, starting with topmost cuts of maximal rank. The Lemma ensures us that after every proof transformation one instance of cut is eliminated.
Therefore, in finitely many steps, we obtain a derivation $\mathcal{D}^*$ of $\Gamma \Rightarrow_0 C$, where the cut rule does not occur.

Eliminating cut

Cut-elimination Theorem If we have a derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p C$, we can construct a derivation $\mathcal{D}^*$ of $\Gamma \Rightarrow_0 C$, that is, a derivation **where** cut **does not occur.**

Proof. We apply the proof transformation detailed in the Lemma to the cuts occurring in $\mathcal{D}$, starting with topmost cuts of maximal rank. The Lemma ensures us that after every proof transformation one instance of cut is eliminated.
Therefore, in finitely many steps, we obtain a derivation $\mathcal{D}^*$ of $\Gamma \Rightarrow_0 C$, where the cut rule does not occur.

☞ How many steps? $\Gamma \Rightarrow_p^m C \;\rightsquigarrow\; \Gamma \Rightarrow_0^{4_p(n)} C$

$$4_p(n) = \underbrace{4^{4^{\cdot^{\cdot^{4^n}}}}}_{p}$$

Why all this work?

- analiticity

- $\vdash_{Hcp} A \quad \rightsquigarrow \quad \vdash_{LJ} \Rightarrow A$ completenes

  - all the axioms of Hilbert system are derivable in LJ

  - MP can be simulated in LJ

$$\frac{A \qquad A \to B}{B} \qquad\qquad \frac{\begin{array}{c} \mathcal{D} \\ \Rightarrow A \end{array} \qquad \begin{array}{c} \mathcal{D} \\ A \Rightarrow B \end{array}}{\Rightarrow B}\,\text{cut}$$

Outline

Detours

In NJ a <u>detour</u> is a gadget formed by an intro. rule followed by an elim. rule

↳ its major premiss is the conclusion of the intro

$$\dfrac{\dfrac{M : A \qquad N : B}{\langle M,N\rangle : A\wedge B}\,\wedge_I}{M : A}\,\wedge_{E_1}$$

major

$$\dfrac{\dfrac{A}{\boxed{A\vee B}}\,\vee_I \qquad \begin{array}{c}A\\ \vdots\\ C\end{array} \qquad \begin{array}{c}B\\ \vdots\\ C\end{array}}{C}\,\vee_E$$

β-reduction

Outline

From NJ to LJ

From NJ to LJ

From LJ to NJ

From LJ to NJ

Back to normalisation