Cut-Elimination in the implicative fragment $\rightarrow$G3mi of an intuitionistic G3-Gentzen-System and its Computational Meaning

Master's Thesis

by

Clemens Grabmayer

at the

Institute for Logic, Language and Computation
Universiteit van Amsterdam

October 1999

Contents

1 Vestergaard's "Computational Anomaly" 2

1.1 A typed $\to$**G3m**-system $\mathcal{G}^+$ 6

1.2 A map Φ from $\mathcal{G}_0^+$+Cut-derivations to derivations a typed$\to$**N[mi]**-system 11

1.3 Cut-elimination for $\mathcal{G}^+$ 15

1.4 Vestergaard's "Anomaly" 25

1.5 Closer analysis of the "problematic" cut-elimination step in $\mathcal{G}^+$ 34

1.6 An alternative system $\to$**G2′mi**^{e*} 39

2 Strong Cut-Elimination 44

2.1 A Strong Cut-Elimination Theorem for $\to$**G3mi** and $\to\perp$**G3i** 47

2.2 A more general version of a Strong Cut-Elimination Theorem for $\to$**G3mi** and $\to\perp$**G3i** 69

Acknowledgement 75

Appendix A: Handout at the Defense 77

Bibliography 81

Chapter 1

A "Computational Anomaly" discovered by Vestergaard in a typed $\to$G3[mi]-System[1]

In [TS96] the Gentzen-systems **G3[mic]** (for minimal, intuitionistic and classical logic) are presented as a formulation of sequent-calculi (proof-systems that were developed by G. Gentzen) with the andecedents and succedents of sequents consisting of multisets of formulas, where the structural rules weakening and contraction do not appear as explicit rules of the systems. In relation to the **G3**-systems these rules occur only as derived rules, that is, as lemmas about derivability. This contrasts with the basic Gentzen-systems **G1[mic]** and partly also with the systems **G2[mic]** defined in [TS96]: Whereas in the **G1**-systems (that remain closest to Gentzen's original sequent-calculi LK and LJ) explicit weakening and contraction rules are part of the systems, weakening does not longer appear as a derivation rule in the **G2**-systems (it has instead been absorbed into the other rules and become a derived rule), but contraction is still present there as a formal rule.

The designation $G3$ for Gentzen-systems without explicit weakening and contraction rules originated with S.C. Kleene, who in [Kl52] presented a sequent-calculus under this name. The formulation of the **G3**-systems in [TS96] owes much—at least in the intuitionistic case—to a Gentzen-system $GHPC$ for intuitionistic logic given by A.G. Dragalin in [Drag79], in which also the succedents of sequents are permitted to consist of mulitsets of

[1] I do want to thank Prof. A.S. Troelstra for his suggestion to investigate the two different, but related topics in proof-theory treated in this thesis, for his careful reading of my drafts and for many ideas about how to improve both the mathematical precision and then also the style of expression and exposition; I do think that I have learned much through this from him. Prof. Troelstra has also found many errors and mistypes for me, for which I feel very thankful. (If other misprints or more severe errors have nevertheless slipped through or there appear shortcomings in precision and the exposition-style, this only shows that I have been negligent in watching my responsibilities. I do want to learn to do better in the future.)

formulas. In *GHPC* more than one formula may occur in the succedent of a sequent (but subject to restrictions implicit in the specific formulation of the rules). This has certain advantages for the exposition of a proof for cut-elimination in *GHPC*, but on the other hand it is an uncommon formulation of a sequent-calculus for intuitionistic logic. For this reason the system was reformulated in the **G3[mi]**-systems with exactly one formula in the succedent of every sequent (as in Gentzen's *LJ* and Kleene's *G*3) by A.S. Troelstra in [TS96].

The fact that the **G3**-systems do not contain structural derivation rules has noteworthy effects on the structure of possible cut-elimination procedures in these systems. In proofs of cut-elimination for the **G1**-systems and in Gentzen's original proof for cut-elimination in *LK* and *LJ* the local transformation steps applied to a derivation, that are needed for removing a cut or for reducing the depth of at least one subderivation of a currently treated cut, depend heavily on the use of the structural rules weakening and contraction[2]. This is no longer in the same way possible in proofs for cut-elimination for the **G3**-systems. Here two ways of carrying out such a proof are practicable: Either (1) in situations, where a weakening or a contraction is necessary to link a derivation $\mathcal{D}$ to other derivations by rules of the system or by cut, certain lemmas have to be relied on, that state, given $\mathcal{D}$ is a cutfree derivation, another derivation $\mathcal{D}'$ can effectively be found, in which the corresponding weakening or contraction has (in relation to $\mathcal{D}$) already taken place. Or (2), explicit weakening and contraction rules are again permitted to occur *temporarily* during the course of performing cut-elimination in a derivation for the purpose of making some of the involved local transformation steps possible, but have to be removed later separately (and also regularly as part of the entire procedure at many occasions).

To refer to these matters more precisely, the definition of the **G3[mi]**-systems in the special case of their implicative fragments will be repeated here as well as the most important properties of these calculi (that they still admit weakening and contraction to be derived rules). In the case of the implicative fragments the most important particularity of the **G3**-systems is the asymmetric formulation of the L→-rule.

Definition 1.1 (G3[mi]'s implicative fragments →G3m and →G3i). The formal system **→G3i** , the implicative fragment of the system **G3i** in [TS96], is defined by the following axioms and rules:

$$\text{Ax} \quad P, \Gamma \Rightarrow P \quad (P \text{ atomic})$$

$$\text{R}\to \; \frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow A \to B}$$

$$\text{L}\to \; \frac{A \to B, \Gamma \Rightarrow A \qquad B, \Gamma \Rightarrow C}{A \to B, \Gamma \Rightarrow C}$$

[2] In the case of *LK* and *LJ* also "exchange" is necessary, since Gentzen considered the antecedents (as well as the succedents in the classical case) to be lists of formulas instead of sets or mulitsets.

The implicative fragment **→G3m** of **G3m** is the same system as **→G3i**; **→G3m** and **→G3i** will here together be referred to as **→G3mi**.

The cut-rule Cut relative to these systems for minimal and intuitionistic logic takes on the shape

$$\text{Cut}\ \frac{\Gamma \Rightarrow D \qquad D, \Pi \Rightarrow C}{\Gamma\Pi \Rightarrow C} \quad .$$

Whenever **→G3mi** will be considered to be enriched by the additional presence of Cut, the extended system will be denoted as either of **→G3mi**+Cut, respectively. ⊠

The system **→G3mi** does not contain weakening and contraction rules

$$\text{W}\ \frac{\Gamma \Rightarrow\ C}{A, \Gamma \Rightarrow\ C} \qquad \text{and} \qquad \text{C}\ \frac{A, A, \Gamma \Rightarrow\ C}{A, \Gamma \Rightarrow\ C} \quad ,$$

but has instead been formulated in such a manner, that these rules are derived or admissible rules of the systems. This is the content of the following lemma.

Lemma 1.1. *Suppose that the notation $\vdash_n$ symbolically designates the notion of derivability in* **→G3mi** *by a deduction of depth $\leq n$. Then for all A, C, Γ, Δ it holds:*

(i) If $\vdash_n \Gamma \Rightarrow C$, then $\vdash_n \Gamma\Delta \Rightarrow C$.

(ii) If $\vdash_n A, A, \Gamma \Rightarrow C$, then $\vdash_n A, \Gamma \Rightarrow C$.

The *proof* of (i) can be done by an immediate induction on n; (ii) can also be shown in this way, but there is one non-obvious case: This occurs, if the formula to be contracted happens to be also the principal formula of an application of L→ in the last step, i.e. in the following situation:

$$\frac{\begin{array}{c}\mathcal{D}_0\\ \vdash_n B \to D, B \to D, \Gamma \Rightarrow B\end{array} \qquad \begin{array}{c}\mathcal{D}_1\\ \vdash_n D, B \to D, \Gamma \Rightarrow C\end{array}}{\vdash_{n+1} B \to D, B \to D, \Gamma \Rightarrow C}\ \text{L}\!\to \quad . \tag{1.1}$$

The induction hypothesis is applicable to $\vdash_n B \to D, B \to D, \Gamma \Rightarrow B$ (and gives $\vdash_n B \to D, \Gamma \Rightarrow B$), but not directly to $\vdash_n D, B \to D, \Gamma \Rightarrow C$. For the purpose of treating this premise of L→ in (1.1) accordingly an additional lemma (right-sided inversion with respect to L→) is usually applied first.

Lemma 1.2 (Inversion lemma with respect to the rule L→). *For* **→G3mi** *the following holds for all A, B, C, Γ in the notation of the preceding lemma:*

If $\vdash_n A \to B, \Gamma \Rightarrow C$, then also $\vdash_n B, \Gamma \Rightarrow C$.

Proof. By a straightforward induction on n. □

The argument in the case (1.1) for the induction-step in the proof of (ii) in Lemma 1.1 can now be carried through: By an application of Lemma 1.2 to the right premise of L$\rightarrow$ in (1.1) $\vdash_n D, B \rightarrow D, \Gamma \Rightarrow C$ also $\vdash_n D, D, \Gamma \Rightarrow C$ follows. The induction hypothesis can then be applied to this latter statement and implies $\vdash_n D, \Gamma \Rightarrow C$. Together with the already established statement $\vdash_n B \rightarrow D, \Gamma \Rightarrow C$ the desired result

$$\frac{\vdash_n B \rightarrow D, \Gamma \Rightarrow B \qquad \vdash_n D, \Gamma \Rightarrow C}{\vdash_{n+1} B \rightarrow D, \Gamma \Rightarrow C}\,\mathrm{L}\!\rightarrow$$

for the completion of the induction step follows.

Lemma 1.2 can also be interpreted as stating that the rule

$$\frac{A \rightarrow B, \Gamma \Rightarrow C}{B, \Gamma \Rightarrow C}\,\mathrm{Inv} \quad ,$$

called *inversion* of L$\rightarrow$ (with respect to its right premise), is an admissible rule in the system **G3[mi]**.

The proof of cut-elimination for the **G3**-systems in [TS96] (there Theorem 4.1.2, p.77) relies in some cases of local transformation-steps on the possibility to perform contractions in a given derivation effectively; that is, a form of Lemma 1.1 is applied in some situations[3]. Thereby for the **G3[mi]**-systems also an inversion-lemma with respect to the rule L$\rightarrow$, a version for the system considered comparable to its special case Lemma 1.2 for $\rightarrow$**G3mi**, comes in implicitly, since the proof of the fact that contraction is a derived rule in $\rightarrow$**G3mi**, depends on such an inversion-lemma (as described above). The proof of a cut-elimination theorem for $\rightarrow$**G3mi** is actually a special case of the proof of cut-elimination for **G3[mi]** in [TS96], in which Lemma 1.1 (which relies on Lemma 1.2 implicitly) can be used to perform weakenings and contractions to given derivations.

There is a very natural (many-to-one, but surjective) map from derivations in an intuitionistic (or minimal) sequent calculus to natural-deduction derivations, which was first described and utilized in the context of his discovery of "normalization" for natural-deduction derivations by D. Prawitz in [Pra65]. The relation between sequent- and natural-deduction calculi under such a map and the exact connection between the concepts of cut-elimination and normalization in these systems was first deeply investigated in a paper [Zu74] of J. Zucker.

He found out, that for a suitable LJ-near sequent-calculus $\mathcal{S}$ and relative to a surjective map Φ from $\mathcal{S}$-derivations to natural-deduction derivations (essentially a map like the one used by Prawitz) cut-elimination steps in a $\mathcal{S}$-derivation $\mathcal{D}$ and normalization-steps on $\Phi(\mathcal{D})$ can simulate each other (with respect to the connection between these formal systems as given by Φ), that is, (1) if for a $\mathcal{S}$-derivation $\mathcal{D}$ a derivation $\mathcal{D}'$ is the result of

[3](A form of this lemma that is true for the considered **G3**-system (and not just for the implicative fragments $\rightarrow$**G3i**, $\rightarrow$**G3m** of the systems **G3[mi]** as Lemma 1.1).)

a cut-elimination step applied to $\mathcal{D}$, then $\Phi(\mathcal{D}')$ can be the result of a finite sequence of normalization-steps performed in $\Phi(\mathcal{D})$, and (2) if for a $\mathcal{S}$-derivation $\mathcal{D}$ a derivation $\tilde{\mathcal{D}}'$ is the result of a normalization-step performed in $\Phi(\mathcal{D})$, then there exists a $\mathcal{S}$-derivation $\mathcal{D}'$ such that $\mathcal{D}'$ is the result of a finite number of cut-elimination steps carried out starting from $\mathcal{D}$ and $\Phi(\mathcal{D}') = \tilde{\mathcal{D}}'$. (Zucker showed this result relative to a completely specified list of cut-elimination steps (the normalization-steps instead had already been given in the form of a—almost entirely—fixed list by Prawitz), but for the negative fragment $\mathcal{S}^-$ of his sequent-calculus $\mathcal{S}$ only. By work of G. Pottinger in [Pott77] this result could be generalized to cover a Gentzen-system for full intuitionistic logic (and even Zucker's system $\mathcal{S}$) by an appropriate choice of the possible cut-elimination steps.)

R. Vestergaard in [Vest99] is interested in whether the execution of cut-elimination steps to a derivation $\mathcal{D}$ in the system $\rightarrow$**G3mi** (according to a usually applied procedure in these systems) can interfere with the "computational meaning" of $\mathcal{D}$ in an irregular manner. In a typed $\rightarrow$**G3mi**-like calculus he gives an example of a sequence $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ of pairwisely different derivations, such that every $\mathcal{D}_n$ (for $n \in \mathbb{N}$) is taken to the same derivation $\mathcal{D}'$ by a cut-elimination procedure very near to the usual one for the untyped system $\rightarrow$**G3mi**, but where all derivations $\mathcal{D}_n$ have different "meanings". If these "computational meanings" were interpreted as the natural-deduction images $\Phi(\mathcal{D}_n)$ of $\mathcal{D}_n$ under a map Φ similar to the one implicitly given by Prawitz, this result would suggest that the very smooth relationship between normalization on natural-deduction derivations and cut-elimination on Gentzen-system derivations—as it exists in the above sketched form with respect to Zucker's LJ-near system $\mathcal{S}$—could be seriously disturbed for the **G3[mi]**-systems.

1.1 A typed →G3m-system $\mathcal{G}^+$

R. Vestergaard in [Vest99] considers a typed system of the implicative fragment $\rightarrow$**G3m** of **G3m**, a system, that will be referred to here as $\mathcal{G}^+_{\boldsymbol{v}}$ (in notational similarity to the system $\mathcal{G}^+$, that will here be presented and used instead), where (1) the type-expressions t assigned to a formula C in the succedent of a conclusion-sequent $x_1 : A_1, \ldots , x_n : A_n \Rightarrow t : C$ ($n \in \mathbb{N}$) of a derivation $\mathcal{D}$ describe a corresponding natural-deduction derivation $\mathcal{D}^*$ with the conclusion C from the marked assumption-classes $[A_1]^{x_1}, \ldots , [A_n]^{x_n}$ very directly, and (2) cut-elimination in this typed system can still be done as suggested by the proof of the cut-elimination theorem for the untyped **G3[mi]**-systems in [TS96]. Vestergaard is interested in the "computational meaning" of derivations in the **G3[mi]**-systems (one could understand the related natural-deduction derivation here as this "meaning") and in how the usual cut-elimination procedure for these systems interferes with this meaning. His system is therefore tailor-made for the purpose of describing cut-elimination on a given derivation as a stepwise process of locally applied transformations (a process that

Vestergaard later describes as one that can be executed according to rules of an appropriate rewrite-rule system). Since—as indicated earlier—the necessity of performing weakenings, contractions and applications of an inversion-lemma often arises during the course of a cut-elimination process for a derivation in a $\rightarrow$**G3mi**-system, explicit rules for such operations had to be devised and taken into the system. These are explicit additional rules that on the one hand will allow to represent cut-elimination as a sequence of local transition-steps, but that on the other hand must be treated separately and ultimately have to be removed completely to arrive at a cut-free derivation.

Vestergaard's system $\boldsymbol{\mathcal{G}}_{\boldsymbol{v}}^{+}$ is very close to the following system $\boldsymbol{\mathcal{G}}^{+}$, that will be used here instead.

Definition 1.2 (The derivation-term annotated systems $\boldsymbol{\mathcal{G}}^{+}$,$\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$). The formal system $\boldsymbol{\mathcal{G}}^{+}$, a typed version of **G3m** 's or **G3i** 's implicative fragment $\rightarrow$**G3mi** is defined as follows: The antecedent of a sequent in this system is a multiset of variables of formula-type (written as variable-annoted formulas), the succedent consists of a (rigidly) typed derivation-term, whose free type-variables occur in the antecedent. $\boldsymbol{\mathcal{G}}^{+}$ has the axioms and rules as listed below:

$$\text{Ax} \quad x : P, \Gamma \Rightarrow \mathbf{ax}_{x^P;\Gamma} : P \quad (P \text{ atomic})$$

$$\text{R}\!\rightarrow \frac{[x : A], \Gamma \Rightarrow t : B}{\Gamma \Rightarrow \boldsymbol{\lambda} x^A.t^B : A \rightarrow B}$$

$$\text{L}\!\rightarrow \frac{x : A \rightarrow B, \Gamma \Rightarrow t_0 : A \qquad [y : B], \Gamma \Rightarrow t_1 : C}{x : A \rightarrow B, \Gamma \Rightarrow \mathbf{let}_{y^B}(t_1^C, x^{A\rightarrow B} t_0^A) : C}$$

$$\text{mW} \frac{\Gamma \Rightarrow t : C}{x_1 : A, \ldots, x_n : A_n, \Gamma \Rightarrow \mathbf{W}^{\bigoplus_{i=1}^{n}\{x_i^{A_i}\}}(t^C) : C}$$

$$\text{mC} \frac{(x_1 : A_1)^2, \ldots, (x_n : A_n)^2, \Gamma \Rightarrow t : C}{x_1 : A_1, \ldots, x_n : A_n, \Gamma \Rightarrow \mathbf{C}^{\bigoplus_{i=1}^{n}\{x_i^{A_i}\}}(t^C) : C}$$

$$\text{Inv} \frac{x : A \rightarrow B, \Gamma \Rightarrow t : C}{y : B, \Gamma \Rightarrow \mathbf{I}_{x^{A\rightarrow B}, y^B}(t^C) : C}$$

Here the following abbreviations and conventions were used:

- The operator $\bigoplus_{i=(.)}^{(.)}$ denotes the union of multisets.
- Typed variables and terms are used in both the notations x^A, t^B and $x : A$, $t : B$, which are considered to be syntactically the same, but the longer versions $x : A$ and $t : B$ informally refer to an assumption A labelled by x (in a corresponding natural-deduction derivation, cf. section 1.2) or a proof-term t of a $\boldsymbol{\mathcal{G}}^{+}$-derivation with B in the antecedent of its conclusion-sequent.

- The notation with a typed variable in brackets [...] is here always to be understood as in the following example: $[x:A],\Gamma$ refers to one of the multisets $\Gamma \oplus \bigoplus_{i=1}^{n}\{x:A\}$ (with $n \in \mathbb{N}$), where $x:A$ is assumed not to be an element of the multiset Γ ($n=0$ is excluded here, i.e. $x:A$ occurs at least once in $[x:A],\Gamma$).

The system $\boldsymbol{\mathcal{G}}_0^+$ has the same axioms as $\boldsymbol{\mathcal{G}}^+$, but contains only the *logical rules* R$\rightarrow$ and L$\rightarrow$ of $\boldsymbol{\mathcal{G}}^+$, not also inversion Inv and the *structural rules* mW and mC of this system.

The succedents of the sequents appearing in a $\boldsymbol{\mathcal{G}}^+$-derivation will be called *derivation-terms* of $\boldsymbol{\mathcal{G}}^+$.

Any of the two systems S defined above can be enriched by the additional presence of the cut-rule

$$\text{Cut}\,\frac{\Gamma \Rightarrow t_0 : D \qquad [x:D],\Pi \Rightarrow t_1 : C}{\Gamma\Pi \Rightarrow t_1^C[\![x^D := t_0^D]\!] : C}$$

to the system S+Cut. $\boxtimes$

The multiple-weakening rule mW could have been written more concisely in the form

$$\text{mW}\,\frac{\Gamma \Rightarrow t : C}{\Delta\Gamma \Rightarrow \mathbf{W}^{\Delta}(t^C) : C}\quad ,$$

but was formulated more explicitly in the above definition so as to allow comparison with the more restrictive (with respect to the form of its active formulas) multiple-contraction rule mC.

There are some noteworthy aspects, in which the systems $\boldsymbol{\mathcal{G}}^+$ defined above differ formally and conceptionally from the system (here abbreviated with:) $\boldsymbol{\mathcal{G}}_v^+$ considered in [Vest99]:

(i) In Vestergaard's system $\boldsymbol{\mathcal{G}}_v^+$ also axioms $x:A,\Gamma \Rightarrow x:A$ are permitted, where the principal formula A does not need to be atomic. For $\boldsymbol{\mathcal{G}}^+$ the stronger condition on axioms Ax as in the **G3**-systems from [TS96] (to refer to atomic principal formulas only) was taken over. (This has no immediate consequence with respect to Vestergaard's result in the case of a typed system with antecedents consisting of *multisets*.)

(ii) Although Vestergaard's system also contains a weakening rule, the derivation-terms do not account for its presence in a derivation. This is because it looks as if the phenomenon treated in [Vest99] has nothing to do with the fact that weakening is not a formal rule in the **G3**-systems. Instead derivations are treated there as "equivalence classes" up to applications of weakening, which—although this could perhaps be made more precise—seems a bit unclear [to me, C.G.]. For this reason (and because explicit treatment of weakening does not cause too much additional work and notation) the term notation in $\boldsymbol{\mathcal{G}}^+$ has been designed to reflect also the effects of weakening by the introduction of a multiple-weakening rule (in a form that will be useful for the description of the cut-elimination procedure implicit in [TS96] (there Theorem 4.1.2 on p. 72) in the special case for the systems $\boldsymbol{\mathcal{G}}^+$).

(iii) To make it possible that a derivation-term describes a $\boldsymbol{\mathcal{G}}^+$-derivation completely, also the variable-annotated formulas in the contexts of axioms Ax had to be formally taken into the term-notation.

(iv) With the exception of the multiple-weakening construct $\mathbf{W}^{(\cdot)}$ basically the same term-expressions are used in [Vest99] to designate applications of rules (only slightly different expressions **let** $x := yt_0$ in t_1 are used there instead of $\mathbf{let}_x(t_1, yt_0)$, the notation used in [TS96] to describe the application of a L$\rightarrow$-rule).

(v) Vestergaard does not consider the succedent of a sequent in his system to be a rigidly typed derivation-term. In his system a succedent consists of a single formula that is annotated by a term t, that is not a type-expression, although it can also describe a derivation in his system precisely (up to occurrences and effects of weakenings, which are neglected). Derivation-(describing-)terms in Vestergaard's system $\boldsymbol{\mathcal{G}}_{\boldsymbol{v}}^+$ are only looked upon as expressions that describe a derivation in his system constructed from assumption variables x, y, z (and also of f, g, h, which he uses exclusively for non-atomic formulas) with the use of term-constructors referring to applications of rules. (This difference has some consequences, that will be explained below. However, these consequences have no bearing on the phenomenon presented in [Vest99].)

A few more things have to be said about the last item: The sequents in Vestergaard's system all have the form

$$x_1 : A_1, x_2 : A_2, \ldots, x_n : A_n \Rightarrow t : C \quad , \tag{1.2}$$

where $n \in \mathbb{N}$, $x_1, \ldots, x_n$ are untyped variables, $A_1, \ldots, A_n, C$ are formulas, and t is a derivation-term formed from untyped variables and derivation-terms by the use of term-constructors inductively as expressions

$$\boldsymbol{\lambda} x.t_0 \ , \ \ \mathbf{let}\, y := xt_0 \text{ in } t_1 \ , \ \ \mathbf{C}^{\Delta}(t_0) \ , \ \ \mathbf{I}_{x,y}(t_0) \text{ or } \ t_1[\![x := t_0]\!] \tag{1.3}$$

(where x, y are variables, t_0, t_1 are terms (by the induction-hypothesis of the definition)). If the formula in the succedent of (1.2) is regarded as the type of the derivation-term C, it could be said, that in this sequent all displayed variables and terms in (1.2) carry types, but subterms of the typed term t^C do not.

Sequents (1.2) of his system are furthermore assumed to be of the special kind that no variable annotates two *different* formulas in the antecedent (a restriction that is taken into the definition of the axioms and rules of his system in the form of a tacit side-condition); but an annotated formula $x : A$ may occur there several times (the antecedents therefore really may be proper multisets).

Vestergaard's motivation for this side-condition in his system is that he intends to abstract away from the derivations as far as possible and that he wants to consider them only in the form of term-representations instead (and for this aim the side-condition really makes sense); in Appendix A of [Vest99] he gives a proof for a statement that every term (inductively definable from variables by the constructors $\boldsymbol{\lambda}$, **let**, **C**, **I** and $\cdot[\![\cdot := \cdot]\!]$ in the above sketched way) really represents—under some mild restrictions on the use of bound variables—a derivation in his system (here called:) $\boldsymbol{\mathcal{G}_v^+}$. Although Vestergaard apparently uses a tacit convention on the use of variables with the letters f, g exclusively for the annotation of non-atomic formulas (contrary to his use of variables like x, y, z) some serious doubts about this statement Lemma 17 on p.8 of [Vest99] seem justified, where [to me, C.G.] it looks as if the proof referring to Lemma 14 runs into troubles in the case of the two-premise rules L$\rightarrow$ and Cut. – This claim for an inverse map from terms (in Vestergaard's notion as described shortly in (1.3)) to derivations in $\boldsymbol{\mathcal{G}_v^+}$ is not essential for his main argument, since by the definition of the axioms and rules in his system $\boldsymbol{\mathcal{G}_v^+}$ every derivation is nevertheless represented by *some* derivation-term (which is obvious from the definition of the rules in this system).

These doubts about whether terms of Vestergaard's system $\boldsymbol{\mathcal{G}_v^+}$ really always represent derivations led to the formulation of the systems $\boldsymbol{\mathcal{G}^+}$ in Definition 1.2. In these systems every derivation $\mathcal{D}$ really is uniquely determined by the derivation-term in the conclusion-sequent. It can be checked on the basis of an inspection of the term-notation given together with the rules of $\boldsymbol{\mathcal{G}^+}$ in Definition 1.2, that every derivation $\begin{array}{c}\mathcal{D}\\ \Gamma \Rightarrow t : C\end{array}$ in $\boldsymbol{\mathcal{G}^+}$ can be reconstructed from the succedent term t^C (which in the rules is written as $t : C$) in the conclusion of $\mathcal{D}$ inductively. Thereby the term t^C also allows to rebuild the antecedent Γ of the conclusion-sequent $\Gamma \Rightarrow t : C$ of $\mathcal{D}$ inductively.

Definition 1.3. The operation ant on $\boldsymbol{\mathcal{G}^+}$+Cut-derivation-terms is defined inductively as follows (where derivation-terms, typed variables and multisets of typed variables occuring below are assumed to be arbitrary such objects appearing within a $\boldsymbol{\mathcal{G}^+}$-derivation according to Definition 1.2):

$$\begin{aligned}
\mathrm{ant}(\mathbf{ax}_{x^A;\Gamma}) &:= \{x : A\} \oplus \Gamma \;\; ; \\
\mathrm{ant}(\boldsymbol{\lambda} x^A.t^B) &:= \mathrm{ant}(t^B) \ominus [x : A] \;\; ; \\
\mathrm{ant}(\mathbf{let}_{y^B}(t_1^C, x^{A\to B} t_0^A)) &:= \mathrm{ant}(t_0^B) \;\; ; \\
\mathrm{ant}(\mathbf{W}^{\Delta}(t^C)) &:= \mathrm{ant}(t^C) \oplus \Delta \;\; ; \\
\mathrm{ant}(\mathbf{C}^{\Delta}(t^C)) &:= \mathrm{ant}(t^C) \ominus \Delta \;\; ; \\
\mathrm{ant}(\mathbf{I}_{x^{A\to B}, y^B}(t^C)) &:= (\mathrm{ant}(t^C) \ominus \{x : A \to B\}) \oplus \{y : B\} \;\; ;
\end{aligned}$$

$$\mathrm{ant}(t_1^C[\![x^D := t_0^D]\!])) := \mathrm{ant}(t_0^D) \oplus (\mathrm{ant}(t_1^C) \ominus [x : D]) \ .$$

(Here $\oplus, \ominus$ denote multiset-union and multiset-subtraction[4] respectively. An expression like $\Gamma \ominus [x : A]$ means the deletion of all occurrences of $x : A$ from the multiset Γ. The outermost types of the terms (which can be reconstructed in an obvious way) on the right side of the definition have been dropped for legibility.) $\boxtimes$

Lemma 1.3. *For every ($\mathcal{G}^+$+Cut)-derivation-term t^C there is exactly one derivation $\mathcal{D}$ in ($\mathcal{G}^+$+Cut) such that $\mathcal{D}$ is of the form*

$$\begin{array}{c} \mathcal{D} \\ \Gamma \Rightarrow t : C \end{array}$$

(Γ a multiset of formulas); for this derivation $\mathcal{D}$ moreover $\Gamma = ant(t^C)$ holds

Proof. By induction on the syntactical depth of t^C, thereby examining all rules of $\mathcal{G}^+$+Cut for the induction-step. $\square$

1.2 A map Φ from $\mathcal{G}_0^+$+Cut-derivations to derivations a typed$\to$N[mi]-system

Derivations in intuitionistic and minimal sequent-calculi can be associated with a corresponding natural-deduction derivation in a very immediate and straightforward way as was first described precisely by D. Prawitz in [Pra65]:

> "A proof in a calculus of sequents can be looked upon as an instruction on how to construct a corresponding natural deduction. This is particularly evident in the case of intuitionistic or minimal logic. A top-sequent then corresponds to a natural deduction consisting of just the formula that occur both in the antecedent and the succedent. As we go downwards in the proof in the calculus of sequents, we successively enlarge in two directions the corresponding natural deductions at the bottom, applying the corresponding I-rules; when we come to applications of antecedent rules, we usually enlarge the corresponding natural deductions at the top, applying the corresponding E-rules." [...]
>
> "The proof in the calculus of sequents can in this way be said to prescribe (to some extent) a certain order in which a corresponding natural deduction can be constructed. This order is often irrelevant and is only partially mirrored in the corresponding natural deduction that results from the construction. Different proofs in the calculus of sequents may therefore correspond (in the way indicated) to the same natural deduction."

[4] $A \ominus B$ means the result of a deletion process, where from the multiset A all elements of the multiset B are removed as often as they occur in B (an element of B can but naturally only be removed from A if it occurs in (is element of) A at all).

In the case of the system $\mathcal{G}^+$ a derivation-term t^C in an end-sequent of a derivation $\mathcal{D}$ does not only allow us to describe $\mathcal{D}$ completely, but gives also—in most cases—clear instructions about how a natural-deduction derivation $\mathcal{D}'$ corresponding to $\mathcal{D}$ under Prawitz' map can be built. This is possible for all $\mathcal{G}^+$-derivations, that do not contain applications of inversion Inv; in the more special case of $\mathcal{G}_0^+$-derivations this is the content of the definition below.

Definition 1.4 (The maps Φ, ϕ, Φ_0 between $\mathcal{G}_0^+$+Cut and $\to$N[mi]*). The map Φ is an operation that takes derivations of the systems $\mathcal{G}_0^+$+Cut (that is, derivations in $\mathcal{G}^+$+Cut containing only applications of logical rules and Cut) to natural-deduction derivations $\Phi(\mathcal{D})$ in a term-calculus $\to$**N[mi]*** for $\to$**N[mi]** (cf. [TS96], Def. 2.2.2, p. 37 for a term calculus for the full systems **N[mi]**, whose special case for **N[mi]** is here referred to as $\to$**N[mi]***). The derivation $\Phi_0(\mathcal{D})$ in $\to$**N[mi]** will denote the deduction $\Phi(\mathcal{D})$ in $\to$**N[mi]*** without the occurrences of term-labels for formulas.

Φ will be defined in parallel with a map ϕ that maps an arbitrary $\mathcal{G}_0^+$+Cut-derivations $\mathcal{D}$ to the term-representation $\phi(\mathcal{D})$ of the $\to$**N[mi]**-derivation $\Phi_0(\mathcal{D})$ underlying $\Phi(\mathcal{D})$. Both Φ and ϕ will be given by an inductive definition on the depth of $\mathcal{D}$ by transitions of the following structure:

$$\begin{array}{c} \mathcal{D} \\ \Gamma \Rightarrow t : C \end{array} \qquad \text{goes to} \qquad \begin{array}{c} [A_1]^{x_1} \dots [A_n]^{x_n} \\ \Phi(\mathcal{D}) \\ \phi(\mathcal{D}) : C \end{array} \quad ,$$

where $\{x_1 : A_1, \dots, x_n : A_n\} \subseteq \mathrm{set}(\Gamma)$, i.e. $\mathrm{set}(\Gamma)$ is the set resulting from the multiset Γ by dropping multiple occurrences of elements in Γ.

If $\mathcal{D}$ consists of an axiom only, then Φ and ϕ are defined by the transition

$$x : A, \Gamma \Rightarrow \mathbf{ax}_{x^A;\Gamma} : A \qquad \text{goes to} \qquad x : A \quad (\text{as marked assumption: } A^x) \ .$$

If $\mathcal{D}$ ends with an application of R$\to$, then $\Phi(\mathcal{D})$ and $\phi(\mathcal{D})$ are defined by the transition

$$\dfrac{\begin{array}{c} \mathcal{D}_0 \\ [x : A], \Gamma \Rightarrow t_0 : B \end{array}}{\Gamma \Rightarrow \lambda x^A.t_0^B : A \to B}\,\mathrm{R}\!\to \qquad \text{goes to} \qquad \dfrac{\begin{array}{c} [A]^x \\ \Phi(\mathcal{D}_0) \\ \phi(\mathcal{D}_0) : B \end{array}}{\boldsymbol{\lambda} x^A.\phi(\mathcal{D}_0)^B : A \to B}\,\to\!\mathrm{I},x \quad .$$

In the case of $\mathcal{D}$ ending with an application of L$\to$, $\Phi(\mathcal{D})$ and $\phi(\mathcal{D})$ are defined by the transition

$$\dfrac{\begin{array}{cc}\mathcal{D}_0 & \mathcal{D}_1\\ x:A\to B,\Gamma\Rightarrow t_0:A & [y:B],\Gamma\Rightarrow t_1:C\end{array}}{x:A\to B,\Gamma\Rightarrow \mathbf{let}_{y^B}(t_1^C, x^{A\to B}t_0^A):C}\,\mathrm{L}{\to} \quad \text{goes to} \quad \begin{array}{c}\dfrac{(A\to B)^x \qquad \begin{array}{c}\langle A\to B\rangle^x\\ \Phi(\mathcal{D}_0)\\ \phi(\mathcal{D}_0):A\end{array}}{[x^{A\to B}\phi(\mathcal{D}_0)^A:B]}\,{\to}\mathrm{E}\\ (\Phi(\mathcal{D}_1))^*[y^B/(x^{A\to B}\phi(\mathcal{D}_0)^A)^B]\\ (\phi(\mathcal{D}_1))^*[y^B/(x^{A\to B}\phi(\mathcal{D}_0)^A)^B]:C\end{array} \quad ,$$

where the stars * indicate that a renaming in the bound variables occurring in the terms within $\Phi(\mathcal{D}_1)$ and in the term $\phi(\mathcal{D}_1)$ has to be performed to make the substitution of $x^{A\to B}\phi(\mathcal{D}_0)^B$ for y^B in $\phi(\mathcal{D}_1)$ possible. Furthermore, the angle-notation $\langle A\to B\rangle^x$ in $\Phi(\mathcal{D})$ above is intended to refer to only a part of the assumption-class $[A\to B]^x$, namely to that part, which consists just of all occurrences of marked assumptions $(A\to B)^x$ in $\Phi(\mathcal{D})$ originating from the assumption-class $[A\to B]^x$ in $\Phi(\mathcal{D}_0)$ (whereas the full assumption-class $[A\to B]^x$ in $\Phi(\mathcal{D})$ contains all occurrences of $(A\to B)^x$ in $\Phi(\mathcal{D})$ originating from $\Phi(\mathcal{D}_0)$ and from $\Phi(\mathcal{D}_1)$ as well as the single additional assumption $(A\to B)^x$ in the major premise of the explicitly shown application of $\to$E). (This notation will be used in similar meaning also for comparable situations.)

If $\mathcal{D}$ ends with a cut, $\Phi(\mathcal{D})$ and $\phi(\mathcal{D})$ are defined according to:

$$\dfrac{\begin{array}{cc}\mathcal{D}_0 & \mathcal{D}_1\\ \Gamma\Rightarrow t_0:D & [x:D],\Pi\Rightarrow t_1:C\end{array}}{\Gamma\Pi\Rightarrow t_1^C[\![x^D:=t_0^D]\!]:C}\,\mathrm{Cut} \quad \text{goes to} \quad \begin{array}{c}\Phi(\mathcal{D}_0)\\ [\phi(\mathcal{D}_0):D]\\ ((\Phi(\mathcal{D}_1))^*[x^D/\phi(\mathcal{D}_0)^D]\\ ((\phi(\mathcal{D}_1))^*[x^D/\phi(\mathcal{D}_0)^D]:C\end{array} \quad .$$

Again, the stars indicate a necessary renaming process in the bound variables, carried out simultaneously in $\Phi(\mathcal{D}_1)$ and $\phi(\mathcal{D}_1)$ to make the substitution of $\phi(\mathcal{D}_0)^D$ for x^D possible.

The derivation $\Phi_0(\mathcal{D})$ in $\boldsymbol{\to}\mathbf{N[mi]}$ is defined from $\Phi(\mathcal{D})$ by dropping the term-expressions in all formulas, that do not occur in a leaf at the top of the derivation $\Phi(\mathcal{D})$ (there the terms are retained as assumption-markers). ⊠

It would also have been possible to extend the maps Φ and ϕ to cover derivations $\mathcal{D}$ in the systems $\boldsymbol{\mathcal{G}}^+$, if $\mathcal{D}$ contains applications of weakening mW and contraction mC, too, but no inversion Inv. This could be done by taking the transitions

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma\Rightarrow t_0:C\end{array}}{\Gamma\Rightarrow \mathbf{W}^{\Delta}(t_0):C}\,\mathrm{mW} \quad \text{and} \quad \dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma\Rightarrow t_0:C\end{array}}{\Gamma\Rightarrow \mathbf{C}^{\Delta}(t_0):C}\,\mathrm{mC} \quad \text{go to} \quad \begin{array}{c}\Phi(\mathcal{D}_0)\\ \phi(\mathcal{D}_0):C\end{array}$$

into Definition 1.4.

Vestergaard does not refer to such a map Φ from derivations in his system $\boldsymbol{\mathcal{G}_v^+}$ to natural-deduction derivations explicitly, he but bases his argument merely on formal observations about terms, which—with the concepts and the terminology used here—can be looked upon as derivation-terms $\phi(\mathcal{D})$ describing natural-deduction derivations $\Phi_0(\mathcal{D})$ associated with derivations $\mathcal{D}$ in the system $\boldsymbol{\mathcal{G}^+}$ (for Vestergaard these derivations are derivations in the system $\boldsymbol{\mathcal{G}_v^+}$, which here was only taken as the basis for the formulation of $\boldsymbol{\mathcal{G}^+}$ in Definition 1.2 above). He takes these derivation-terms $\phi(\mathcal{D})$ of natural-deduction derivations $\Phi_0(\mathcal{D})$ to be the "computational meaning" of derivations $\mathcal{D}$ in the typed sequent-calculus. – Definition 1.4 was set up with the intention of following Vestergaard's paper as closely as possible, but also with the aim of looking at the phenomenon he describes from a slightly different (namely a proof-theoretic) angle.

There is one noticeable feature of the map Φ as defined above (also by way of following the argument of Vestergaard) above, which distinguishes it from an analogous map used by J. Zucker in [Zu74]: In the case of a derivation $\mathcal{D}$ ending with the cut-rule, having $\mathcal{D}_0$ and $\mathcal{D}_1$ as immediate subderivations, the transition in Definition 1.4 necessitates the amalgamation in $\Phi(\mathcal{D})$ of open assumption-classes $[A]^x$ with $A^x \not\equiv D^x$, that occur both in the derivations $\Phi(\mathcal{D}_0)$ and $\Phi(\mathcal{D}_1)$ (given by the induction-hypothesis of the definition of Φ). In $\Phi(\mathcal{D})$, which then can be written in the form

$$\begin{array}{c}
\langle A\rangle^x \\
\Phi(\mathcal{D}_0) \\
[\phi(\mathcal{D}_0) : D] \qquad \langle A\rangle^x \\
((\Phi(\mathcal{D}_1))^*[x^D/\phi(\mathcal{D}_0)^D] \\
((\phi(\mathcal{D}_1))^*[x^D/\phi(\mathcal{D}_0)^D] : C
\end{array} \quad ,$$

a new assumption-class $[A]^x$ is formed, that now consists of all occurrences of A^x at places $\langle A\rangle^x$ (which here in $\Phi(\mathcal{D})$ stand for the occurrences of marked assumptions A^x in the subderivation $\Phi(\mathcal{D}_0)$ and in the part originating from $\Phi(\mathcal{D}_1)$.)

In the case of Zucker's sequent-calculus $\mathcal{S}$ such an amalgamation of assumption-classes by his (rather similar) map Φ does not happen, which is due to a very special—indeed careful—way of the formulation of the logical rules and the cut-rule in $\mathcal{S}$ based on a special indexing system for antecedent-formulas. An identification of different assumption-classes is in this system only possible in the image under Φ of a $\mathcal{S}$-derivation ending with an application of (an unrestricted version of) the contraction-rule.

A situation, similar to a $\boldsymbol{\mathcal{G}^+}$-derivation ending with Cut, arises for a derivation $\mathcal{D}$ ending with L$\rightarrow$, that has immediate subderivations $\mathcal{D}_0$ and $\mathcal{D}_1$. There, too, identifications of assumption-classes from $\Phi(\mathcal{D}_0)$ and $\Phi(\mathcal{D}_1)$ take place implicitly in the respective transition of Definition 1.4, that is, all occurrences of marked assumptions C^z (if $C^z \not\equiv (A \to B)^x$) in $\Phi(\mathcal{D})$, that originate from open assumptions C^z in $\Phi(\mathcal{D}_0)$ or $\Phi(\mathcal{D}_1)$, are taken to form the new open assumption-class $[C]^z$ of $\Phi(\mathcal{D})$. The open assumption class $[A \to B]^x$ of $\Phi(\mathcal{D})$

contains all open assumptions of the form $[A \to B]^x$, that originate from open assumptions in $\Phi(\mathcal{D}_0)$ or $\Phi(\mathcal{D}_1)$ and one additional occurrence of $(A \to B)^x$ (as apparent from the definition of $\Phi(\mathcal{D})$).

While in the case of Zucker's sequent-calculus $\mathcal{S}$ each two-premise-rule R (including Cut) was formulated in such a way that in the image $\Phi(\mathcal{D})$ of a $\mathcal{S}$-derivation $\mathcal{D}$ ending with R identifications of assumption-classes originating from $\Phi(\mathcal{D}_0)$ and $\Phi(\mathcal{D}_1)$ ($\mathcal{D}_0$ and $\mathcal{D}_1$, the immediate subderivations of $\mathcal{D}$) never take place (during the formation of $\Phi(\mathcal{D})$ from $\Phi(\mathcal{D}_0)$ and $\Phi(\mathcal{D}_1)$), this is not in the same way possible for a typed calculus of a $\to$**G3mi**-system: The formulation of the L$\to$-rule with a multiset Γ appearing simultaneously in the antecedents of both premises and in the conclusion of the rule leads to the necessity of amalgamating assumption-classes from $\mathcal{D}_0$ and $\mathcal{D}_1$ in the image $\Phi(\mathcal{D})$ of a derivation ending with L$\to$ ($\mathcal{D}_0$ and $\mathcal{D}_1$ are here again the immediate subderivations of $\mathcal{D}$).

1.3 Cut-elimination for $\mathcal{G}^+$

A procedure for cut-elimination in the system $\mathcal{G}^+$ relies for some steps on the possibility of renaming variables in the antecedents of a sequent and throughout the immediate subderivations appropriately. This could be done as a local process by the introduction of a new renaming-construct in addition to the rules of $\mathcal{G}^+$ just like mW, mC and Inv, and by a separate treatment of applications of this new construct for the purpose of cut-elimination in a derivation.

Since Vestergaard describes cut-elimination in his system as a process of successive applications of rewrite-rules on derivation-terms, and because substitution is a familiar notion for terms, he refers for this matter on substitution-lemmas like the below ones instead:

Lemma 1.4. *Let $\mathcal{D}$ be a derivation of the form*

$$\begin{array}{c}\mathcal{D}\\ [(x:A)^n],\Gamma \Rightarrow t:C\end{array}$$

(where $n \in \mathbb{N}$) in $\mathcal{G}^+$, that contains rule-applications of logical rules of $\mathcal{G}^+$ only. Then for all type-variables y^A, that are distinct from all bound variables in t^C

$$\begin{array}{c}\mathcal{D}[x^A/y^A]\\ (y:A)^n,\Gamma \Rightarrow t[x^A/y^A]:C\end{array}$$

holds.

Proof. By an induction on the depth $|\mathcal{D}|$ of $\mathcal{D}$.

One key case for the validity of this lemma (a case involving the principal formula of an application of L$\to$, that is the root of trouble in many similar situations) shall be shown here. If $x : A$ is the principal annotated formula $x : B \to D$ (i.e. if $A \equiv B \to D$) of a bottom-most application of L$\to$ in $\mathcal{D}$ and if $n = l + 1$, then $\mathcal{D}$ is of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_0 \\ [x : B \to D, (x : B \to D)^l], \Gamma \Rightarrow t_0 : B\end{array} \qquad \begin{array}{c}\mathcal{D}_1 \\ [z : D], [(x : B \to D)^l], \Gamma \Rightarrow t_1 : C\end{array}}{[(x : B \to D)^n], \Gamma \Rightarrow \mathbf{let}_{z^D}(t_1^C, x^{B\to D} t_0^B) : C}\,\mathrm{L}\!\to \quad ,$$

where $l \in \mathbb{N}_0$. If $l > 0$, then the induction-hypothesis can be applied to both $\mathcal{D}_0$ and $\mathcal{D}_1$, and due to $z^D \not\equiv y^{B\to D}$ the results $\mathcal{D}_0[x^{B\to D}/y^{B\to D}]$ and $\mathcal{D}_1[x^{B\to D}/y^{B\to D}]$ can be linked together again by L$\to$ with the derivation

$$\dfrac{\begin{array}{c}\mathcal{D}_0[x^{B\to D}/y^{B\to D}] \\ y : B \to D, (y : B \to D)^l, \Gamma \Rightarrow t_0[x/y] : B\end{array} \qquad \begin{array}{c}\mathcal{D}_1[x^{B\to D}/y^{B\to D}] \\ [z : D], (y : B \to D)^l, \Gamma \Rightarrow t_1[x/y] : C\end{array}}{(y : B \to D)^n, \Gamma \Rightarrow \big(\mathbf{let}_{z^D}(t_1^C, x^{B\to D} t_0^B)\big)[x^{B\to D}/y^{B\to D}] : C)}\,\mathrm{L}\!\to$$

as outcome. If $l = 0$ the induction-hypothesis has to be applied only to $\mathcal{D}_0$ and the resulting derivation $\mathcal{D}_0[x^{B\to D}/y^{B\to D}]$ can then be linked together again with $\mathcal{D}_1$ by L$\to$ to reach a derivation of the desired form. $\square$

Lemma 1.5. *For all $\boldsymbol{\mathcal{G}}^+$-derivations* $\begin{array}{c}\mathcal{D} \\ \Gamma \Rightarrow t : C\end{array}$ *and all typed variables x^A a $\boldsymbol{\mathcal{G}}^+$-derivation $\mathcal{D}_{(x^A)}$ and thus also a derivation-term $t_{(x^A)}$ can effectively be found (by renaming annotated variables in the antecedents and in derivation-terms only) such that*

$$\begin{array}{c}\mathcal{D}_{(x^A)} \\ \Gamma \Rightarrow t_{(x^A)} : C\end{array}$$

and x^A does not occur among the bound variables of $t_{(x^A)}$ and in $\mathcal{D}_{(x^A)}$.

Proof. By induction on the depth of $\mathcal{D}$, carrying out a renaming of x^A to another variable $(x')^A$ not previously occuring in the derivation (which can be done by an appropriate use of Lemma 1.4) in the induction-step, whenever x^A appears as a bound variable. $\square$

Lemma 1.6. *Let $\mathcal{D}$ be a derivation of the form* $\begin{array}{c}\mathcal{D} \\ [(x : A)^n], \Gamma \Rightarrow t : C\end{array}$ *(where $n \in \mathbb{N}$) in $\boldsymbol{\mathcal{G}}^+$, that only contains rule-applications of logical rules. Then for all variables y*

$$\begin{array}{c}\mathcal{D}_{(y^A)}[x^A/y^A] \\ (y : A)^n, \Gamma \Rightarrow t_{(y^A)}[x^A/y^A] : C\end{array}$$

holds, where $\mathcal{D}_{(y^A)}$ and $t_{(y^A)}$ are in relation to $\mathcal{D}, t$ and y^A defined (and implicitly constructed in the proof of) Lemma 1.5.

Proof. This is immediate from Lemma 1.4 and Lemma 1.5. $\square$

Theorem 1.1. *Cut-elimination holds for $\mathcal{G}^+$.*

More precisely, every derivation $\mathcal{D}$ in ($\mathcal{G}^+$+Cut) can be transformed by a finite sequence of successively applied local reduction-steps with a cut-free derivation in $\mathcal{G}_0^+$ as result, i.e. a derivation $\mathcal{D}'$, that contains neither applications of the cut-rule nor of the rules multiple-weakening mW, multiple contraction mC or inversion Inv.

Furthermore the process of cut-elimination for a derivation $\mathcal{D}$ in $\mathcal{G}^+$ can be completely simulated on derivation-terms by applications of rules from an appropriate rewrite-rule system starting at the derivation-term t of $\mathcal{D}$; these rule-applications have to respect a certain order, in which single rewrite-rule steps are successively executed.

Proof. The proof of this theorem relies on two lemmas below, that together deal with the case of a $\mathcal{G}^+$-derivation $\mathcal{D}$, i.e. a derivation not containing Cut but arbitrary many applications of weakening mW, contraction mC and inversion Inv (Lemma 1.7 and Lemma 1.8 below give—with the help of an immediate induction on the number of applications of mW, mC and Inv in $\mathcal{D}$—that every such derivation $\mathcal{D}$ can effectively be transformed in the desired manner to a derivation $\mathcal{D}'$ in $\mathcal{G}_0^+$, thus to a derivation $\mathcal{D}'$, that possesses only applications of logical rules). It suffices therefore here to show that every derivation $\mathcal{D}$ terminating with an application of Cut, such that the immediate subderivations $\mathcal{D}_0$ and $\mathcal{D}_1$ of $\mathcal{D}$ contain only logical rules, can be transformed by stepwise and local transformations to a cut-free derivation $\mathcal{D}'$ in $\mathcal{G}_0^+$, i.e. a derivation having only applications of logical rules. (The theorem then follows by induction on the number of applications mW, mC, Inv or Cut in $\mathcal{D}'$, an induction, in which always topmost occurrences of these rules are treated and removed.)

This can be shown by an induction on one plus the logical depth of the formula A within the annotated cut-formula $x : A$ in the cut at the bottom of $\mathcal{D}$, which is called the *rank* of the cut, together with a subinduction on the *level* $|\mathcal{D}_0| + |\mathcal{D}_1|$ of this cut, where $\mathcal{D}_0$ and $\mathcal{D}_1$ are its immediate subderivations.

The proof is very similar to that of the cut-elimination theorem for the **G3**-systems (cf. Theorem 4.1.2 on p. 77 in [TS96]), more precisely, it is analogous to the proof of a cut-elimination theorem for the implicative fragments **→G3mi** of **G3[mi]**.

For a derivation $\mathcal{D}$ in $\mathcal{G}^+$+Cut of the form

$$\text{Cut}\ \frac{\begin{matrix}\mathcal{D}_0\\ \Gamma \Rightarrow t_0 : D\end{matrix} \qquad \begin{matrix}\mathcal{D}_1\\ [x : D], \Pi \Rightarrow t_1 : C\end{matrix}}{\Gamma\Pi \Rightarrow t_1^C[\![x^D := t_0^D]\!] : C}\ \text{S}$$

with $\mathcal{D}_0, \mathcal{D}_1$ $\mathcal{G}_0^+$-derivations (thus containing only applications of logical rules) three cases are distinguished and treated separately: (1) If one of the premises of the cut S consists of an axiom, then a reduction which removes the cut in one step can be performed. (2) If both premises are not axioms, and the cut-formula is not principal in at least one of the rule-applications S_0, S_1 immediately preceding S, then the cut can be permuted upwards over the logical rule S_i ($i = 0, 1$)) in that respective premise, thereby reducing the level of the resulting cut(s) by at least one; in the case of a two-premise rule two new cuts of lower level may appear. The induction-hypothesis can then be applied to show that the new cut(s) is (are) transformable to a cut-free form. The resulting derivation(s) can then be linked together again (either with each other, in the case of two new cuts, or otherwise with a weakened subderivation of $\mathcal{D}$) by a logical rule of the same type as S_i to build the result of cut-elimination for $\mathcal{D}$. (3) If both of the premises of the cut at the bottom of $\mathcal{D}$ are not axioms, and the cut-formula is principal in both premises, then a "fork-reduction" (a term from [Drag79] for a similar reduction) can take place. In the perhaps most frequent case (the cut-formula does occur just once in the antecedent of the left premise of the cut at the bottom of $\mathcal{D}$) a reduction can be performed by splitting the cut S in one cut S'_1 of the same rank, but lower level, and two succeeding cuts S'_2 and S'_3 of lower rank than that of S, cuts, which are then followed by a number of contractions. Th subinduction-hypothesis can be applied to the derivation $\tilde{\mathcal{D}}'_1$ terminating with S'_1 for showing that $\tilde{\mathcal{D}}'_1$ can be transformed to a cut-free derivation in $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$ in the desired way. Then the induction-hypothesis can be used twice and successively to show that S'_2 and S'_3 can also be removed in this way. The succeeding contractions can then be done away with by an appeal to Lemma 1.8 to arrive ultimately at a derivation in $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$. – In the second case of a fork-reduction for a ($\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$+Cut)-derivation $\mathcal{D}$, that occurs if the cut-formula of the cut S at the bottom of $\mathcal{D}$ appears more than once in the antecedent of its left premise, a similar reduction is performed: Now S is split into two cuts S''_1 and S''_2 of the same rank as S, but of lower level, that are linked together by a cut S''_3 and followed by a cut S''_4, both of lower level than S, the latter of which is then succeeded by a number of contractions in the resulting derivation $\mathcal{D}''$. The subinduction-hypothesis and the induction-hypothesis can then be applied similar as before to show that $\mathcal{D}''$ is transformable to a cut-free derivation in $\boldsymbol{\mathcal{G}}^{+}$ in the desired way.

All these reduction-steps are rather straightforward to perform and—except for one case, that will be shown here below—analogous[5] to the ones for an untyped $\rightarrow$**G3mi**^{e}-system (defined precisely in Definition 2.1). For this system the reduction-steps are displayed in the lists A–C in chapter 2, following Definition 2.1, where they were gathered for the formulation of a strong cut-elimination theorem.

Since derivation-terms in $\boldsymbol{\mathcal{G}}^{+}$ uniquely represent derivations in these systems, the reduction-steps referred to in this proof can be given in the form of rewrite-rules on derivation-terms. These respective rules will be given at the end of this proof.

In case (1) the situation of a derivation $\mathcal{D}$ of the form

[5](with the sole exception of the second case of a "fork-reduction"-step just described, which is an analogue to a cut-elimination-step necessary for dealing with a similar situation in the Kleene-System $\rightarrow$**GK3mi**)

$$\dfrac{x : P, \Gamma \Rightarrow \mathbf{ax}_{x^P;\Gamma} : P \qquad \begin{array}{c}\mathcal{D}_1\\ {[(y:P)^n], \Pi \Rightarrow t_1 : C}\end{array}}{x : P, \Gamma\Pi \Rightarrow t_1^C[\![y^P := \mathbf{ax}_{x^P;\Gamma}]\!] : C}\ \text{Cut}$$

can occur. $\mathcal{D}$ can be then be transformed to a cut-free derivation $\mathcal{D}'$

$$\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}(\mathcal{D}_1)_{(x^P)}[y^P/x^P]\\ (x:P)^n, \Pi \Rightarrow (t_1)_{(x^P)}[y^P/x^P] : C\end{array}}{(x:P)^{n-1}, \Pi \Rightarrow \mathbf{C}^{\{x^P\}}\big((t_1)_{(x^P)}[y^P/x^P]\big) : C}\ \text{mC}}{\vdots}\ \text{mC}}{x : P, \Pi \Rightarrow \underbrace{\mathbf{C}^{\{x^P\}}\big(\ldots \mathbf{C}^{\{x^P\}}}_{n-1}\big((t_1)_{(x^P)}[y^P/x^P]\big)\ldots\big) : C}\ \text{mC}}{x : P, \Gamma\Pi \Rightarrow \mathbf{W}^{\Gamma}\big(\mathbf{C}^{\{x^P\}}\big(\ldots \mathbf{C}^{\{x^P\}}\big((t_1)_{(x^P)}[y^P/x^P]\big)\ldots\big)\big) : C}\ \text{mW}$$

with the help of an application of Lemma 1.6. $\mathcal{D}'$ can then be seen to be transformable to a derivation, which only contains applications of logical rules (or consists of a single axiom) by $n-1$ successive appeals to Lemma 1.8 (for the contractions) and one to Lemma 1.7 (for the weakening).

The rewrite-rules originating from cut-elimination steps in this proof are gathered in the lists A–C below[6]:

A. Axiomatic Cut-Reduction Rewrite-Rules

a. $t_1^C[\![y^P := \mathbf{ax}_{x^P;\Gamma}]\!] \ \longrightarrow_{(\text{Cut})}\ \mathbf{W}^{\Gamma}\big(\underbrace{\mathbf{C}^{\{x^P\}}\big(\ldots \mathbf{C}^{\{x^P\}}}_{n-1}\big((t_1^C)_{(x^P)}[y^P/x^P]\big)\ldots\big)\big)$,
where[7] $n := \text{mult}(y^P, \text{ant}(t_1^C))$,
and $(t_1^C)_{(x^P)}$ is defined in relation to t_1^C and x^P by an application of Lemma 1.5.

b. $\mathbf{ax}_{y^P;\Pi}[\![y^P := t_0^P]\!] \ \longrightarrow_{(\text{Cut})}\ \mathbf{W}^{\Pi}(t_0^P)$

c. $\mathbf{ax}_{x^P;y^D,\Pi_0}[\![y^D := t_0^D]\!] \ \longrightarrow_{(\text{Cut})}\ \mathbf{ax}_{x^P;\text{ant}(t_0^D)\Pi_0}$.

B. Rewrite-Rules for Upwards-Permutation of Cut

a. $t_1^C[\![y^D := \mathbf{let}_{z^B}(t_{01}^D, x^{A\to B} t_{00}^A)]\!] \ \longrightarrow_{(\text{Cut})}$

$$\mathbf{let}_{(z')^B}\big(t_1^C[\![y^D := t_{01}^D[z^B/(z')^B]]\!], x^{A\to B}\mathbf{W}^{\text{ant}(t_1^C)\ominus[y:D]}(t_{00}^A)\big)$$

for z' such that $\big[(z')^B \notin \text{ant}(t_1^C) \ominus [y:D] \ \wedge$
$\wedge\ \big(z \equiv z' \ \vee\ (z')^B$ does not occur in $t_{01}^D\big)\big]$.

[6] The symbols used here are essentially meta-language symbols (as "by default" throughout the thesis), which has the consequence that variables x and y, formulas A and B or typed variables x^A, y^B need not stand for different variables or formulas in general expressions, except this is explicitly stated using formulations like for instance $x \not\equiv y$, $A \not\equiv B$ or $x^A \not\equiv y^B$.

[7] Where for every multiset Γ and every object a the expression $\text{mult}(a, \Gamma)$ means the *multiplicity* of a in Γ, i.e. the number of occurrences of a in Γ.

b. $(\boldsymbol{\lambda} x^A.t_{10}^B)[\![y^D := t_0^D]\!] \quad \longrightarrow_{(\mathrm{Cut})} \quad \boldsymbol{\lambda}(x')^A.((t_{10}^B[x^A/(x')^A])[\![y^D := t_0^D]\!])$

for x' such that $\big[(x')^A \notin \mathrm{ant}(t_0^D) \ \wedge$

$\wedge \ \big(x \equiv x' \ \vee \ (x')^A$ does not occur in $t_{10}^B\big)\big]$.

c. Whenever $y^D \not\equiv x^{A\to B}$:

$$\big(\mathbf{let}_{z^B}(t_{11}^C, x^{A\to B}t_{10}^A)[\![y^D := t_0^D]\!]\big) \quad \longrightarrow_{(\mathrm{Cut})}$$

$$\mathbf{let}_{(z')^B}\big((t_{11}^C[z^B/(z')^B])[\![y^D := t_0^D]\!], x^{A\to B}t_{10}^A[\![y^D := t_0^D]\!]\big)$$

for z' such that $\big[(z')^B \notin \mathrm{ant}(t_0^D) \wedge$

$\wedge \ \big(z \equiv z' \ \vee \ (z')^B$ does not occur in $t_{11}^C\big)\big]$.

C. Fork Reduction Rewrite-Rule

$$\mathbf{let}_{(z^B)}(t_{11}^C, y^{A\to B}t_{10}^A)[\![y^{A\to B} := \boldsymbol{\lambda} x^A.t_{00}^B]\!] \quad \longrightarrow_{(\mathrm{Cut})}$$

$$\begin{cases} \mathbf{C}^{\Gamma\Pi}\big(t_{11}^C[\![z^B := t_{00}^B[\![x^A := t_{10}^A[\![y^{A\to B} := \boldsymbol{\lambda} x^A.t_{00}^B]\!]]\!]]\!]\big) \\ \qquad\qquad \ldots \ \text{if } \Pi = \Pi_0 \ \text{(or equivalently if: } \mathrm{mult}(y^{A\to B}, t_{10}^A) = 1) \\ \mathbf{C}^{\Gamma}\big(\mathbf{C}^{\Gamma\Pi_0}\big((t_{11}^C[\![y^{A\to B} := \boldsymbol{\lambda} x^A.t_{00}^B]\!])[\![z^B := t_{00}^B[\![x^A := t_{10}^A[\![y^{A\to B} := \boldsymbol{\lambda} x^A.t_{00}^B]\!]]\!]]\!]\big)\big) \\ \qquad\qquad \ldots \ \text{if } \Pi \supsetneqq \Pi_0 \ \text{(or equivalently if: } \mathrm{mult}(y^{A\to B}, t_{10}^A) > 1) \end{cases},$$

where $\Gamma := \mathrm{ant}(t_{00}^B) \ominus [x : A]$,
$\Pi := \mathrm{ant}(t_{10}^A) \ominus \{y : A \to B\}$ and
$\Pi_0 := \mathrm{ant}(t_{10}^A) \ominus [y : A \to B]$.

(It is immaterial to understand the reason for the particular form of the subscript-notation used in derivation-terms in the above rewrite-rules for their application. Yet, this notation comes from one for subderivations of a given derivation $\mathcal{D}$, where for $n \in \mathbb{N}$ the derivation $\mathcal{D}_{i_1,\ldots,i_n}$ ($i_1, \ldots, i_n \in \mathbb{N}_0$) is inductively defined as the subderivation of $\mathcal{D}$ leading to the $(i_n + 1)$-th premise from the left of the bottom-most rule application in $\mathcal{D}_{i_1,\ldots,i_{n-1}}$, if $n > 1$, or in $\mathcal{D}$, if $n = 1$; if $\mathcal{D}$ for example terminates with a one-premise rule, then only the immediate subderivation $\mathcal{D}_0$ of $\mathcal{D}$ is defined in this way, not also $\mathcal{D}_1, \mathcal{D}_2, \ldots$, and similar for more-premise rules. This notation was extended here to derivation-terms for devising and checking the above rewrite-rules and it was thought that this origin should not get concealed in the result.)

It is apparent from the above rewrite-rules, that during the process of cut-elimination in a derivation according to the stepwise and local procedure used here implicitly new applications of weakening or of contraction or of both appear in a derivation, that previously may only have contained applications of logical rules and of cut. Therefore these rewrite-rules for cut-elimination on derivation-terms will have to be supplemented by further rules for the reduction of derivation-terms containing subterms $\mathbf{W}^{\Delta}t$ for weakening- and $\mathbf{C}^{\Delta}t$ for contraction-applications. These additional rules correspond to the places in this proof, where Lemma 1.7 and Lemma 1.8 are used (so to say as subroutines) and will be given together with the proofs for these statements.

Cut-elimination on a $\boldsymbol{\mathcal{G}}^{+}$-derivation-term t can then be seen as a finite sequence of reductions according to the system of rewrite-rules partly given here and being completed below: A reduction-process, in which always an innermost occurrence of a cut-term $t_1[\![x := t_0]\!]$ in t is considered, where (a) first t_0 and t_1 are transformed by a finite subsequence of reductions (involving applications of rewrite-rules for the treatment of weakening-, contraction- and inversion-subterms) to terms $\tilde{t_0}$ and $\tilde{t_1}$, that only contain applications of logical rules (or that are axioms), such that then (b) the cut $\tilde{t_1}[\![x := \tilde{t_0}]\!]$ can get reduced by one of the above cut-reduction rules to a term $\tilde{t}$ (by arguments used above it is clear that one of these rules is then always applicable). The proof guarantees the termination of this process for every given $\boldsymbol{\mathcal{G}}^{+}$-derivation-term, provided the termination of the subprocesses, which will be justified separately in Lemma 1.7 and Lemma 1.8. $\square$

Lemma 1.7. *Every derivation $\mathcal{D}$ in $\boldsymbol{\mathcal{G}}^{+}$, that contains only applications of weakenings mW and logical rules, can be transformed by a finite number of local transformation-steps to a derivation $\mathcal{D}'$ in $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$ (i.e. $\mathcal{D}'$ contains no application of mW any more, nor such of mC or Inv); this elimination process of weakenings can take place on the corresponding derivation-terms as the successive application of appropriate rewrite-rules.*

Proof. It suffices to show that weakening can be effectively eliminated from a derivation $\mathcal{D}$ in $\boldsymbol{\mathcal{G}}^{+}$ terminating with an application of mW, that for the rest contains only applications of logical rules (the lemma then follows by induction on the number of weakenings in a given derivation). This in turn can be shown by induction on the depth $|\mathcal{D}|$ of $\mathcal{D}$:

If $|\mathcal{D}| = 1$ and mW therefore is applied directly to an axiom in $\mathcal{D}$, then an easy reduction of $\mathcal{D}$ to a new axiom, which now incorporates the weakening, can take place. If $|\mathcal{D}| > 1$, then the rule mW is permuted upwards one step over a rule R$\rightarrow$ or L$\rightarrow$ and then the induction hypothesis is applicable. In the case of R$\rightarrow$ the following reduction is possible:

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0\\ [x:A],\Gamma \Rightarrow t_0 : B\end{matrix}}{\Gamma \Rightarrow \boldsymbol{\lambda} x^A.t_0^B : A \to B}\,\mathrm{R}{\to}}{\Gamma\Delta \Rightarrow \mathbf{W}^{\Delta}(\boldsymbol{\lambda} x^A.t_0^B)^{A\to B} : A \to B}\,\mathrm{mW}$$

reduces to

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0[x^A/(x')^A] \\ [x' : A], \Gamma \Rightarrow t_0[x^A/(x')^A] : B\end{matrix}}{[x' : A], \Gamma\Delta \Rightarrow \mathbf{W}^{\Delta}\big(t_0^B[x^A/(x')^A]\big) : B}\ \text{mW}}{\Gamma\Delta \Rightarrow \boldsymbol{\lambda}(x')^A.\mathbf{W}^{\Delta}\big(t_0^B[x^A/(x')^A]\big) : A \to B}\ \text{R}\!\to \quad .$$

The substitution of $(x')^A$ for x^A in $\mathcal{D}_0$ is possible by Lemma 1.4, if $(x')^A$ does not occur as a bound variable in t_0^B; to keep $x' : A$ distinct from elements of Γ furthermore the stronger condition, that $(x')^A$ does not occur in t_0^B at all, is necessary. The application of R$\to$ after mW in the displayed way in the reduced derivation is possible, if $(x')^A \notin \Delta$. – The case with mW following an application of L$\to$ at the bottom of $\mathcal{D}$ is more costly to write down, but quite analogously to treat.

Since the derivation-term t for a $\boldsymbol{\mathcal{G}}^{+}$-derivation $\mathcal{D}$ allow to represent (and to reconstruct) $\mathcal{D}$ completely, the reductions referred to (and in the case of R$\to$ explicitly given) in this proof can be gathered for the following list D of rewrite-rules for weakening-reductions:

D. Weakening Rewrite-Rules

a. $\mathbf{W}^{\Delta}(\mathbf{ax}_{x^P;\Gamma}) \ \longrightarrow_{(\text{Weak})}\ \mathbf{ax}_{x^P;\Gamma\Delta}$.

b. $\mathbf{W}^{\Delta}(\boldsymbol{\lambda} x^A.t^B) \ \longrightarrow_{(\text{Weak})}\ \boldsymbol{\lambda}(x')^A.\mathbf{W}^{\Delta}\big((t[x^A/(x')^A])^B\big)$,
where x' is such that $(x')^A$ does not occur in Δ nor in the term t^B.

c. $\mathbf{W}^{\Delta}(\mathbf{let}_{(y')^B}(t_1^C, x^{A\to B}t_0^A)) \ \longrightarrow_{(\text{Weak})}$
$\mathbf{let}_{\,(y')^B}\big(\mathbf{W}^{\Delta}\big((t_1[y^B/(y')^B])^C\big), x^{A\to B}\mathbf{W}^{\Delta}(t_0^A)\big)$,
where y' is such that $(y')^B$ does neither occur in Δ nor in t_1^C.

Weakening-elimination on derivation-terms takes the form of successive reductions of innermost occurrences of weakening in a term t according to the above rewrite-rules, i.e. of occurrences $\mathbf{W}^{\Delta}(t_0)$ in t, such that t_0 does not contain further subterms of the form $\mathbf{W}^{\Delta'}(t_{00})$. □

Lemma 1.8. *Every derivation $\mathcal{D}$ in $\boldsymbol{\mathcal{G}}^{+}$, that contains only applications of the rules multiple-contraction mC, inversion Inv and logical rules, can be transformed effectively by a finite number of transformation-steps to a derivation $\mathcal{D}'$ in $\boldsymbol{\mathcal{G}}_0^{+}$ (i.e. one, that possesses only applications of logical rules). This elimination process can furthermore be carried through completely on the derivation-terms corresponding to $\boldsymbol{\mathcal{G}}_0^{+}$-derivations in the form of applications of rules of a rewrite-rule system.*

Proof. It again suffices to show that an application of inversion or multiple-contraction can effectively be removed from a $\boldsymbol{\mathcal{G}}^{+}$-derivation $\mathcal{D}$ terminating with either Inv or mC, but otherwise containing only applications of logical rules (the lemma then follows by induction on the number of applications of inversion and multiple-contraction in $\mathcal{D}$). This in turn can be shown by an induction on either the logical complexity $|A \to B|$ of the annotated inversion-formula $x : A \to B$ in an application $\mathbf{I}_{x^{A\to B}, y^B}(...)$ of inversion[8], if the bottom-most rule in $\mathcal{D}$ is an inversion, or on the sum of the logical complexities of the contraction-formulas in Δ of an application $\mathbf{C}^{\Delta}(\dots)$ of mC at the bottom of $\mathcal{D}$, together with—in both cases—a subinduction on the depth $|\mathcal{D}|$ of $\mathcal{D}$.

Applications of mC or Inv, that follow axioms in $\mathcal{D}$, can be directly reduced to other axioms. Applications of these rules following logical rules can be permuted upwards over L$\to$ and R$\to$ in most cases, much in the same way as in the case for mW in Lemma 1.7; the subinduction hypothesis can then always be applied.

There are two less obvious cases:

(1) If $\mathcal{D}$ terminates with the inversion of the principal formula of an immediately preceding application of L$\to$, then $\mathcal{D}$ is of the form

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0 \\ x : A \to B, \Gamma \Rightarrow t_0 : A\end{matrix} \qquad \begin{matrix}\mathcal{D}_1 \\ [(z : B)^n], \Gamma \Rightarrow t_1 : C\end{matrix}}{x : A \to B, \Gamma \Rightarrow \mathbf{let}_{z^B}(t_1^C, x^{A\to B} t_0^A) : C}\,\mathrm{L}{\to}}{y : B, \Gamma \Rightarrow \mathbf{I}_{x^{A\to B}, y^B}\big((\mathbf{let}_{z^B}(t_1^C, x^{A\to B} t_0^A))^C\big) : C}\,\mathrm{Inv} \quad ,$$

where $n \in \mathbb{N}$.

The inversion can here be removed by taking the right subdeduction $\mathcal{D}_1$ of L$\to$, carrying out a renaming y^B for z^B with the help of Lemma 1.6 and by then using a number of contractions for the annotated formulas $y : B$. The induction hypothesis can then be applied to all of these contractions, since the logical complexity $|B|$ of B is smaller than that of $A \to B$. The reduced derivation then has the form

$$\dfrac{\dfrac{\dfrac{\begin{matrix}(\mathcal{D}_1)_{(y^P)}[z^P/y^P] \\ (y : B)^n, \Gamma \Rightarrow (t_1^C)_{(y^P)}[z^P/y^P] : C\end{matrix}}{(y : B)^n, \Gamma \Rightarrow \mathbf{C}^{\{y^P\}}\big((t_1^C)_{(y^P)}[z^P/y^P]\big) : C}\,\mathrm{mC}}{\vdots}\,\mathrm{mC}}{y : B, \Gamma \Rightarrow \underbrace{\mathbf{C}^{\{y^P\}}\big(\dots \mathbf{C}^{\{y^P\}}}_{n-1}\big((t_1^C)_{(y^P)}[z^P/y^P]\big) \dots\big) : C}\,\mathrm{mC} \quad .$$

(2) If $\mathcal{D}$ terminates with a contraction mC, that involves the principal annotated formula $x : A \to B$ of an application of L$\to$ immediately preceding the inversion, then $\mathcal{D}$ is of the

[8] Such an application was here indicated as an operation on derivation-terms.

form

$$\dfrac{\dfrac{\dfrac{\mathcal{D}_0}{x:A\to B, x:A\to B,\Gamma \Rightarrow t_0 : A} \qquad \dfrac{\mathcal{D}_1}{[y:B], x:A\to B,\Gamma \Rightarrow t_1 : C}}{x:A\to B, x:A\to B,\Gamma \Rightarrow \mathbf{let}_{y^B}(t_1^C, x^{A\to B}t_0^A) : C}\,\mathrm{L}{\to}}{x:A\to B, \Gamma\ominus(\Delta\ominus\{x:A\to B\}) \Rightarrow \mathbf{C}^{\Delta}(\dots) : C}\,\mathrm{mC} \quad .$$

$\mathcal{D}$ can here be transformed to the derivation

$$\dfrac{\dfrac{\dfrac{\mathcal{D}_0}{x:A\to B, x:A\to B,\Gamma \Rightarrow t_0 : A}}{x:A\to B,\Pi \Rightarrow \mathbf{C}^{\Delta}(t_0) : C}\,\mathrm{mC} \qquad \dfrac{\dfrac{\dfrac{\mathcal{D}_1}{[y:B], x:A\to B,\Gamma \Rightarrow t_1 : C}}{[y:B],\Gamma \Rightarrow \mathbf{I}_{x^{A\to B},y^B}(t_1) : C}\,\mathrm{Inv}}{[y:B],\Pi \Rightarrow \mathbf{C}^{\Delta\ominus\{x^{A\to B}\}}\big(\mathbf{I}_{x^\cdot,y^\cdot}(t_1)\big) : C}\,\mathrm{mC}}{x:A\to B,\Pi \Rightarrow \mathbf{let}_{y^B}\big(\mathbf{C}^{\Delta\ominus\{x^{A\to B}\}}\big(\mathbf{I}_{x^{A\to B},y^B}(t_1)\big), x^{A\to B}\mathbf{C}^{\Delta}(t_0)\big) : C}\,\mathrm{L}{\to} \quad ,$$

where $\Pi :\equiv \Gamma\ominus(\Delta\ominus\{x:A\to B\})$.

Here, since $|\mathcal{D}_0|, |\mathcal{D}_1| < |\mathcal{D}|$ holds, the subinduction-hypothesis can be applied to see that the inversion immediately below $\mathcal{D}_1$ and the contraction succeeding the end-sequent of $\mathcal{D}_0$ can be eliminated (by a required stepwise local process) with the results $\mathcal{D}_0'$ and $\tilde{\mathcal{D}}_1'$, that only contain logical rules. The induction hypothesis can then be applied to $\tilde{\mathcal{D}}_1'$ to see that a contraction $\mathbf{C}^{\Delta\ominus\{x^{A\to B}\}}$ immediately succeeding $\tilde{\mathcal{D}}_1'$ can be eliminated as a desired stepwise local process with result $\mathcal{D}_1'$ in $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$. Linking together $\mathcal{D}_0'$ and $\mathcal{D}_1'$ by L$\to$ then leads to an inversion- and contraction-free derivation $\mathcal{D}'$ in $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$.

On derivation-terms the reductions needed in the proof of this lemma can be presented as rewrite-rules of the following lists E and F.

E. Contraction Rewrite-Rules

a. $\mathbf{C}^{\Delta}(\mathbf{ax}_{x^P;\Gamma}) \;\longrightarrow_{(\mathrm{Cont})}\; \mathbf{ax}_{x^P;\Gamma\ominus\Delta}$.

b. $\mathbf{C}^{\Delta}(\boldsymbol{\lambda}x^A.t^B) \;\longrightarrow_{(\mathrm{Cont})}\; \boldsymbol{\lambda}x^A.\mathbf{C}^{\Delta}(t^B)$.

c. $\mathbf{C}^{\Delta}\big(\mathbf{let}_{y^B}(t_1^C, x^{A\to B}t_0^A)\,\big) \;\longrightarrow_{(\mathrm{Cont})}$

$$\begin{cases} \mathbf{let}_{(y^B)}\big(\mathbf{C}^{\Delta}(t_1^C), x^{A\to B}\mathbf{C}^{\Delta}(t_0^A)\big) & \dots\ x^{A\to B}\notin\Delta \\ \mathbf{let}_{(y^B)}\big(\mathbf{C}^{\Delta\ominus\{x^{A\to B}\}}\big((\mathbf{I}_{x^{A\to B},y^B}(t_1^C)\big), x^{A\to B}\mathbf{C}^{\Delta}(t_0^A)\big) & \dots\ x^{A\to B}\in\Delta \end{cases} .$$

F. Inversion Rewrite-Rules

a. $\mathbf{I}_{x^{A\to B},y^B}(\mathbf{ax}_{z^P;x^{A\to B},\Gamma}) \;\longrightarrow_{(\mathrm{Inv})}\; \mathbf{ax}_{z^P;y^B,\Gamma}$.

b. $\mathbf{I}_{x^{A\to B},y^B}(\boldsymbol{\lambda}z^C.t^D) \;\longrightarrow_{(\mathrm{Inv})}\; \boldsymbol{\lambda}(z')^C.\mathbf{I}_{x^{A\to B},y^B}(t^D[z^C/(z')^C])$,
where z' is such that $(z')^C \not\equiv y^B$ and $(z')^C$ does not occur in t^D.

c. $\mathbf{I}_{x^{A\to B},y^B}\big(\mathbf{let}_{w^D}(t_1^E, z^{C\to D}t_0^C)\big) \quad \longrightarrow_{(\mathrm{Inv})}$

$\mathbf{let}_{\,(w')^D}\big(\mathbf{I}_{x^{A\to B},y^B}\big((t_1[w^D/(w')^D])^E\big), z^{C\to D}\mathbf{I}_{x^{A\to B},y^B}(t_0^C)\big)$,
where w' is such that $(w')^D \not\equiv y^B$ and $(w')^D$ does not occur in the term t_1^E.

d. $\mathbf{I}_{x^{A\to B},y^B}\big(\mathbf{let}_{z^B}(t_1^C, x^{A\to B}t_0^A)\big) \quad \longrightarrow_{(\mathrm{Inv})}$

$\underbrace{\mathbf{C}^{\{y^B\}}\big(\ldots\mathbf{C}^{\{y^B\}}\big(}_{n-1}\,(t_1^C)_{(y^B)}[z^B/y^B]\big)\ldots\big)$,
where $n := \mathrm{mult}(z : B, \mathrm{ant}(t_1^C))$ and[9] $(t_1^C)_{(y^B)}$ is defined according to Lemma 1.5.

Contraction- and inversion-elimination on a $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$-derivation-term t, that contains no weakening-subterms, can then be looked upon as a process consisting of successive reductions of innermost occurrences of contraction- or inversion-subterms in t according to the above rewrite-rules in the lists E and F; the termination of this process is guaranteed by the arguments above (that apply also for derivation-terms, since by Lemma 1.3 $\boldsymbol{\mathcal{G}}^{+}$-derivations and $\boldsymbol{\mathcal{G}}^{+}$-derivation-terms correspond to each other uniquely). □

1.4 Vestergaard's "Anomaly"

In introductory remarks at the begin of section 5 in [Vest99], where the "computational anomaly" in his system is presented, Vestergaard starts from the observation, that with respect to his system $\boldsymbol{\mathcal{G}}_{\boldsymbol{v}}^{+}$ there are two inversion-reduction rewrite-rules, for which it is apparent from their shape, that they do not preserve the identity of derivation-terms. That is, syntactically different derivation-terms t_1 and t_2 can get reduced to the same derivation-term t. Only one of these two rules corresponds to a respective rewrite-rule for inversion-reduction in the setting of the system $\boldsymbol{\mathcal{G}}^{+}$ considered here instead; the other would correspond to the reduction of a derivation, which consists of the application of an inversion $\mathbf{I}_{x^{A\to B},y^B}(\ldots)$ (here indicated as an operation on derivation-terms) to an axiom $\mathbf{ax}_{x^{A\to B};\Gamma}$ with a non-atomic annotated principal formula $A \to B$ (such axioms have but been excluded in the formulation of $\boldsymbol{\mathcal{G}}^{+}$ similar as in the systems **G3[mi]** in [TS96]). The one remaining inversion-reduction rewrite-rule with this noticeable property is the rewrite-rule F.d given in the proof of Lemma 1.8:

$$\mathbf{I}_{x^{A\to B},y^B}\big(\mathbf{let}_{z^B}(t_1^C, x^{A\to B}t_0^A)\big) \quad \longrightarrow_{(\mathrm{Inv})} \quad \underbrace{\mathbf{C}^{\{y^B\}}\big(\ldots\mathbf{C}^{\{y^B\}}\big(}_{n-1}(t_1^C)_{(y^B)}[z^B/y^B]\big)\ldots\big) \;,$$

(where $n := \mathrm{mult}(z : B, \mathrm{ant}(t_1^C))$ and $(t_1^C)_{(y^B)}$ is defined according to Lemma 1.5). Here the subterm t_0^A occuring in the derivation-term on the left side of the reduction is obviously

[9] See footnote 7 for an explanation of multiplicities mult.

lost in the reduction, since it does not appear in the reduced term. Therefore different terms can get reduced to the same result by applications of this rule.

Vestergaard then asks whether it is possible that during the process of cut-elimination (represented as a finite reduction-sequence on derivation-terms) applications of the two rewrite-rules appearing with respect to his system, that do not preserve identity of terms, can actually change the "computational meaning" of derivation-terms in an unexpected way. Or whether an "unfortunate use of the inversion-principle" does never happen during a cut-elimination process—perhaps due to some very special features of this process as a whole. Unstated though (and only hinted at other places in the article), he seems to ask here, whether it is possible that the "computational meaning" of a derivation-term can be changed during the execution of the usual cut-elimination procedure *substantially different* from the way, how the "meaning" of a derivation is affected by normalization (which in the case of derivation-terms for derivations of an intuitionistic of minimal calculus simply corresponds to λ-reduction).

He then immediately proceeds by giving his example of a "computational anomaly", which is intended to provide an answer insofar, as it shows that the computational meaning of a derivation containing cut in his system can indeed be changed during the cut-elimination process in an unexpected way, that is, in a way not corresponding to normalization (λ-contraction) on derivation-terms.

Before looking closer at the "anomaly", let us briefly note this: Vestergaard does not mention the parallel case of a rewrite-rule among those necessary for dealing with axiomatic cut-reductions, which does obviously also not preserve the identity of derivation-terms and which appears in his system as well as in $\boldsymbol{\mathcal{G}}^{+}$: Here it is the rule A.c,

$$\mathbf{ax}_{x^P;y^D,\Pi_0}[\![y^D := t_0^D]\!] \quad \longrightarrow_{(\mathrm{Cut})} \quad \mathbf{ax}_{x^P;\mathrm{ant}(t_0^D)\Pi_0} \quad , \tag{1.4}$$

given in the proof of Theorem 1.1. Apparently the term t_0^D disappears during this reduction (only the annotated formulas in the antecedent of the derivation $\mathcal{D}_0$ represented by t_0 remain as side formulas of the resulting axiom, which but do not tell anything about the derivation $\mathcal{D}_0$ that leads to the conclusion $\mathrm{ant}(t_0) \Rightarrow t_0 : D$). – Were it then also possible that the "meaning" of a derivation-term could be (unexpectedly) changed due to an application of this rule during a process of cut-elimination executed on derivation-terms?

If $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$+Cut -derivation terms t are given the "meaning" of their natural-deduction derivation image $\Phi(\mathcal{D})$ of the derivation $\mathcal{D}$ represented by t, this is not possible, as can be seen in the following way: On the related $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{+}$+Cut-derivations the reduction (1.4) corresponds to:

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \mathrm{ant}(t_0^D) \Rightarrow t_0 : D\end{array} \qquad [y:D], x:P, \Pi_0 \Rightarrow \mathbf{ax}_{x^P;y^D,\Pi_0} : P}{x:P, \mathrm{ant}(t_0^D), \Pi_0 \Rightarrow \mathbf{ax}_{x^P;y^D,\Pi_0}[\![y^D := t_0^D]\!] : P}\,\mathrm{Cut}$$

reduces to

$$x : P, \mathrm{ant}(t_0^D), \Pi_0 \Rightarrow \mathbf{ax}_{x^P;\mathrm{ant}(t_0^D),\Pi_0} : P \ .$$

If now the images under the map Φ from section 2 are formed from the derivations on the left and on the right side of this axiomatic cut-reduction, it easily turns out, that both natural-deduction derivation images are equal to P^x. Thus this "meaning" of derivation-terms is certainly not changed by applications of the rule A.c .

It is therefore possible, that the "computational meaning" of a derivation-term (if this is interpreted as the natural-deduction image of the corresponding $\boldsymbol{\mathcal{G}_0^+}$+Cut-derivation) is not affected by applications of rewrite-rules with the seemingly very bad property that the identity of the terms, on which they act, is not preserved. On the other hand it can easily be checked for the rewrite-rules for upwards-permutation of Cut of the list B in the proof of Theorem 1.1, that while "identity of terms" is preserved under applications of these rules, they *do* nevertheless actually change the natural-deduction images of the corresponding $\boldsymbol{\mathcal{G}^+}$-derivations (if only in a way that corresponds to the execution of normalization steps on these natural-deduction derivations). Zucker in [Zu74] calls cut-elimination steps of these kind in list B *permutative conversions* and they play a vital rule for his result of a close correspondence between cut-elimination in his intuitionistic sequent-calculus $\mathcal{S}$ and normalization on $\mathcal{N}$, his intuitionistic system of natural-deduction.

For the efficient treatment of weakening in the exposition of Vestergaard's "anomaly" the following definition, which introduces an abbreviation for weakened derivation-terms, and a lemma about the relation between this new notation and weakening-reduction $\longrightarrow^*_{(Weak)}$ on derivation-terms will be used.

Definition 1.5. Let t be a derivation-term of a derivation $\mathcal{D}$ in $\boldsymbol{\mathcal{G}_0^+}$.
Then for typed variables $x_1^{A_1}, \dots, x_n^{A_n}$ the term $t\{x_1^{A_1}, \dots, x_n^{A_n}\}$ denotes the derivation-term of that derivation, which results from $\mathcal{D}$ by adding $x_1 : A_1, \dots, x_n : A_n$ in the antecedent of every sequent in $\mathcal{D}$ as well as to the side-formulas Γ in every axiom-subterm $\mathbf{ax}_{y^P;\Gamma}$ (for arbitrary y^P) of a derivation-term occuring in $\mathcal{D}$. ⊠

Lemma 1.9. *Every derivation $\mathcal{D}$ in $\boldsymbol{\mathcal{G}^+}$ of the form*

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma \Rightarrow t : C\end{array}}{x_1 : A_1, \dots, x_n : A_n, \Gamma \Rightarrow t : C}\, mW \quad ,$$

where $\mathcal{D}_0$ is a derivation in $\boldsymbol{\mathcal{G}_0^+}$ (hence it is an axiom or contains only applications of logical rules) and the typed variables $x_1^{A_1}, \dots, x_n^{A_n}$ do not occur as bound variables in t^C, can be effectively transformed to a derivation $\mathcal{D}'$ in $\boldsymbol{\mathcal{G}_0^+}$ of the form

$$\begin{array}{c}\mathcal{D}_0'\\ x_1 : A_1, \dots, x_n : A_n, \Gamma \Rightarrow t\{x_1^{A_1}, \dots, x_n^{A_n}\} : C\end{array} \quad .$$

Moreover, with respect to the weakening-reduction rewrite-rules in the list D given in the proof of Lemma 1.7 for every derivation-term t of a $\mathcal{G}_0^+$-derivation $\mathcal{D}$ and arbitrary typed variables $x_1^{A_1}, \ldots, x_n^{A_n}$, that do not occur as bound variables in t,

$$\mathbf{W}^{\bigoplus_{i=1}^{n}\{x_i^{A_i}\}}(t) \quad \longrightarrow^*_{(Weak)} \quad t\{x_1^{A_1}, \ldots, x_n^{A_n}\}$$

holds.

Proof. Can be seen to be implicit in the proof of Lemma 1.7. $\square$

The example of a "computational anomaly" given by Vestergaard uses the following derivation-terms $\imath_n$ and $\jmath_n$ (for $n \in \mathbb{N}_0$):

$$\imath_n :\equiv \begin{cases} \mathbf{ax}_{z^P;s^{P\to P}} & \ldots\ n = 0 \\ \mathbf{let}_{(x_n)^P}\big(\mathbf{ax}_{x_n^P;z^P}, s^{P\to P}\imath_{n-1}^P\big) & \ldots\ n > 0 \end{cases} \tag{1.5}$$

and

$$\jmath_n :\equiv \begin{cases} \mathbf{ax}_{z^P;(s^{P\to P})^2} & \ldots n = 0 \\ \mathbf{let}_{y^P}\big(\imath_n\{y^P\}, s^{P\to P}\imath_1^P\{s^{P\to P}\}\big) & \ldots\ n > 0\ . \end{cases}$$

(Here the same designations of these terms and of the involved variables in the type-expressions (more precisely, the names of the typed variables here, if the formula-types were dropped) have been kept to make comparisons with the reduction-sequence in [Vest99] easier.)

Vestergaard's example of a "computational anomaly" in his system $\mathcal{G}_{\boldsymbol{v}}^+$ can now be rewritten as sequences σ_n of reductions ($n \in \mathbb{N}$ refers to the subterm $\jmath_n$ in the first term t_n in all of these sequences) by applications of rewrite-rules from the lists A–F in section 1.3 for a cut-elimination process according to Theorem 1.1. For all $n \in \mathbb{N}$ the following holds:

$$\begin{aligned}
&\mathbf{let}_{(y')^P}\big(\jmath_n\{(y')^P\}, g^{P\to P}\imath_1\{g^{P\to P}, s^{P\to P}\}\big)[\![g^{P\to P} := \boldsymbol{\lambda} z^P.\mathbf{ax}_{(z')^P;z^P}]\!] \\
&\longrightarrow_{(Cut)} \quad \mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\Big(\jmath_n\{(y')^P\}[\![(y')^P := \mathbf{ax}_{(z')^P;z^P} \\
&\qquad\qquad [\![z^P := \imath_1\{g^{P\to P}, s^{P\to P}\}[\![g^{P\to P} := \boldsymbol{\lambda} z^P.\mathbf{ax}_{(z')^P;z^P}]\!]]\!]]\!]\Big) \\
&\longrightarrow_{(Cut)} \quad \mathbf{C}^{\{\ldots\}}\Big(\ldots[\![\ldots[\![z^P := \mathbf{let}_{x_1^P}\big(\mathbf{ax}_{x_1^P;z^P,s^{P\to P},g^{P\to P}}[\![g^{P\to P} := \boldsymbol{\lambda} z^P.\mathbf{ax}_{(z')^P;z^P}]\!], \\
&\qquad\qquad s^{P\to P}(\imath_0\{g^{P\to P}, s^{P\to P}\}[\![g^{P\to P} := \boldsymbol{\lambda} z^P.\mathbf{ax}_{(z')^P;z^P}]\!])\big)]\!]]\!]\Big) \\
&\longrightarrow_{(Cut)} \quad \mathbf{C}^{\{\ldots\}}\Big(\ldots\ldots[\![\ldots[\![z^P := \mathbf{let}_{x_1^P}\big(\mathbf{ax}_{x_1^P;z^P,(z')^P,s^{P\to P}}\ , \\
&\qquad\qquad s^{P\to P}(\imath_0\{g^{P\to P}, s^{P\to P}\}[\![g^{P\to P} := \boldsymbol{\lambda} z^P.\mathbf{ax}_{(z')^P;z^P}]\!])\big)]\!]]\!]\Big)
\end{aligned}$$

$$
\begin{aligned}
\longrightarrow_{(Cut)} \quad & \mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\Big(\jmath_n\{(y')^P\}[\![(y')^P := \mathbf{ax}_{(z')^P;z^P} \\
& \qquad [\![z^P := \mathbf{let}_{x_1^P}\big(\mathbf{ax}_{x_1^P;(z')^P,z^P,s^{P\to P}}, s^{P\to P}\mathbf{ax}_{z^P;(s^{P\to P})^2,(z')^P}\big)]\!]]\!]\Big) \\
\longrightarrow_{(Cut)} \quad & \mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\jmath_n\{(y')^P\}[\![(y')^P := \mathbf{ax}_{(z')^P;(z')^P,(s^{P\to P})^2,z^P}]\!]\big) \\
\longrightarrow_{(Cut)} \quad & \mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\mathbf{W}^{\{((z')^P)^2, z^P, (s^{P\to P})^2\}}(\jmath_n)\big) \\
\longrightarrow^*_{(Weak)} \quad & \mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\jmath_n\{((z')^P)^2, z^P, (s^{P\to P})^2\}\big) \\
& \equiv \mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\mathbf{let}_{y^P}\big(\imath_n\{y^P\}, s^{P\to P}\imath_1^P\{s^{P\to P}\}\big)\{((z')^P)^2, z^P, (s^{P\to P})^2\}\big) \\
& \equiv \mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\mathbf{let}_{y^P}\big(\imath_n\{y^P, ((z')^P)^2, z^P, (s^{P\to P})^2\}, \\
& \qquad s^{P\to P}\imath_1^P\{((z')^P)^2, z^P, (s^{P\to P})^3\}\big)\big) \\
\longrightarrow_{(Cont)} \quad & \mathbf{let}_{y^P}\big(\mathbf{C}^{\{z^P,(z')^P\}}\mathbf{I}_{s^{P\to P},y^P}(\imath_n\{y^P, ((z')^P)^2, z^P, (s^{P\to P})^2\}), \\
& \qquad s^{P\to P}\mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}(\imath_1\{((z')^P)^2, z^P, (s^{P\to P})^3\})\big) \\
\longrightarrow_{(Inv)} \quad & \mathbf{let}_{y^P}\big(\mathbf{C}^{\{z^P,(z')^P\}}\big(\mathbf{ax}_{y^P;z^P}\{((z')^P)^2, z^P, s^{P\to P}\}\big), \\
& \qquad s^{P\to P}\mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}(\imath_1\{((z')^P)^2, z^P, (s^{P\to P})^3\})\big) \\
\longrightarrow_{(Cont)} \quad & \mathbf{let}_{y^P}\big(\mathbf{ax}_{y^P;z^P}\{(z')^P, s^{P\to P}\} \\
& \qquad s^{P\to P}\mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}(\imath_1\{((z')^P)^2, z^P, (s^{P\to P})^3\})\big) \\
\longrightarrow^{(3)}_{(Cont)} \quad & \mathbf{let}_{y^P}\big(\mathbf{ax}_{y^P;z^P}\{(z')^P, (s^{P\to P})\}, s^{P\to P}((\imath_1\{s^{P\to P}\})\{(z')^P, z^P, s^{P\to P}\})\big) \\
& \equiv \mathbf{let}_{y^P}(\mathbf{ax}_{y^P;z^P}, s^{P\to P}\imath_1\{s^{P\to P}\})\{(z')^P, s^{P\to P}\}\ .
\end{aligned}
$$

(The $\longrightarrow^*_{(Weak)}$-reduction step above has its justification in Lemma 1.9, in the step $\longrightarrow^{(3)}_{(Cont)}$ three applications of contraction-reduction rewrite-rules have been gathered.)

Let for $n \in \mathbb{N}$ the term t_n be the first derivation-term in the above reduction-sequence σ_n and let t' denote the resulting derivation-term ($n \in \mathbb{N}$ refers to the free occurrence of n in $\jmath_n$ in the derivation-term at the beginning). If one looks at the natural-deduction derivation images $\Phi(\mathcal{D}_n)$ and $\Phi(\mathcal{D}')$ for the ($\boldsymbol{\mathcal{G}}^+$+Cut)-derivations $\mathcal{D}_n$ and $\mathcal{D}'$ corresponding to the derivation-terms t_n and t', it can easily be seen, that $\Phi(\mathcal{D}_n)$ is the typed natural-deduction derivation

$$
\dfrac{(P\to P)^s \qquad \dfrac{\dfrac{(P\to P)^s \qquad \dfrac{(P\to P)^s \qquad P^z}{sz : P}\,{\to}\mathrm{E}}{s(sz) : P}\,{\to}\mathrm{E}}{\vdots \atop s^{n-1}z : P}}{s^n z : P}\,{\to}\mathrm{E} \tag{1.6}
$$

,

whereas $\Phi(\mathcal{D}')$ is just

$$\frac{(P \to P)^s \qquad \dfrac{(P \to P)^s \qquad P^z}{sz : P} \to\mathrm{E}}{s(sz) : P} \to\mathrm{E} \tag{1.7}$$

Since for $n > 2$ the derivations $\Phi_0(\mathcal{D}_n)$ do not reduce to $\Phi_0(\mathcal{D}')$ by normalization ($\Phi_0(\mathcal{D}_n)$ and $\Phi_0(\mathcal{D}')$ being the derivations $\Phi(\mathcal{D}_n)$ and $\Phi(\mathcal{D}')$ after dropping the term-labels in all conclusions of rule-applications), this means that "something more" than what would have corresponded to normalization must have happened during cut-elimination here. It is straightforward to check that the "jump" for $n > 2$ of $\Phi(\mathcal{D}_n^{(i)})$ (with $\mathcal{D}_n^{(i)}$ being the derivation corresponding to the i-th derivation-term in σ_n) from (1.6) to (1.7) occurs just in the single inversion-reduction step in σ_n.

Informally it is clear, that an unnecessarily complicated proof, the natural-deduction derivation in (1.6), has obviously been reduced to a more compact form, the derivation in (1.7). But this reduction on proofs (as informal objects that are thought to underly formal derivations (cf. for example G. Kreisel in [Kr71], p. 111)) is not reflected by the way how normalization on natural-deduction simplifies proofs. (It also has to be noted that the "proof-reduction" between (1.6) and (1.7) is certainly not optimal and therefore seemingly indeed of a special kind.)

It is interesting, that in the setting of a system $\boldsymbol{\mathcal{G}}^{+}$ comparable to Vestergaard's (in the notation here:) $\boldsymbol{\mathcal{G}}^{+}_{\boldsymbol{v}}$, but containing an explicit weakening-rule, Vestergaard's example does not look convincing any more: This is because the contraction-reduction rewrite-rule E.c in the proof of Lemma 1.8 can certainly be modified to the following more careful form:

$$\mathbf{C}^{\Delta}\big(\mathbf{let}_{y^B}(t_1^C, x^{A\to B}t_0^A)\ \big) \quad \longrightarrow_{(\mathrm{Cont})}$$

$$\begin{cases} \mathbf{let}_{(y^B)}\big(\mathbf{C}^{\Delta}(t_1^C), x^{A\to B}\mathbf{C}^{\Delta}(t_0^A)\big) \\ \ldots\ x^{A\to B} \notin \Delta \text{ or } \big(x^{A\to B} \in \Delta \text{ and } \ 2.\mathrm{mult}(x^{A\to B}, \Delta) < \mathrm{mult}(x^{A\to B}, \mathrm{ant}(t_0))\big) \\ \mathbf{let}_{(y^B)}\big(\mathbf{C}^{\Delta\ominus\{x^{A\to B}\}}\big((\mathbf{I}_{x^{A\to B}, y^B}(t_1^C)\big), x^{A\to B}\mathbf{C}^{\Delta}(t_0^A)\big) \\ \ \ldots\ x^{A\to B} \in \Delta \text{ and } \ 2.\mathrm{mult}(x^{A\to B}, \Delta) = \mathrm{mult}(x^{A\to B}, \mathrm{ant}(t_0))) \end{cases}$$

On the corresponding $\boldsymbol{\mathcal{G}}^{+}_{\mathbf{0}}$-derivations this means: Multiple contraction is always allowed to permute upwards over L$\to$ directly (without the use of inversion) also in case that the principal formula of L$\to$ is active in the succeeding contraction, whenever it is possible to do this. That is, whenever there are enough occurrences of the principal formula of L$\to$ also in the right premise of the involved application of this rule to carry out the required multiple contraction also there: If $x : A \to B \in \Delta$, then

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0 \\ x : A \to B, x : A \to B, \Gamma \Rightarrow t_0 : A\end{matrix} \qquad \begin{matrix}\mathcal{D}_1 \\ [y : B], x : A \to B, \Gamma \Rightarrow t_1 : C\end{matrix}}{x : A \to B, x : A \to B, \Gamma \Rightarrow \mathbf{let}_{y^B}(t_1^C, x^{A\to B} t_0^A) : C}\,\mathrm{L}\!\to}{x : A \to B, \Gamma \ominus (\Delta \ominus \{x : A \to B\}) \Rightarrow \mathbf{C}^{\Delta}((\mathbf{let}_{y^B}(t_1^C, x^{A\to B} t_0^A))^C) : C}\,\mathrm{mC}$$

is now allowed to reduce (as in the first case of the rule E.c) to

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0 \\ x : A \to B, x : A \to B, \Gamma \Rightarrow t_0 : A\end{matrix}}{x : A \to B, \Pi \Rightarrow \mathbf{C}^{\Delta}(t_0^A) : A}\,\mathrm{mC} \qquad \dfrac{\begin{matrix}\mathcal{D}_1 \\ [y : B], x : A \to B, \Gamma \Rightarrow t_1 : C\end{matrix}}{[y : B], \Pi \Rightarrow \mathbf{C}^{\Delta}(t_1^C) : C}\,\mathrm{mC}}{x : A \to B, \Pi \Rightarrow \mathbf{let}_{y^B}\big(\mathbf{C}^{\Delta}(t_1^C), x^{A\to B}\mathbf{C}^{\Delta}(t_0^A)\big) : C}\,\mathrm{L}\!\to \quad ,$$

(where $\Pi :\equiv \Gamma \ominus (\Delta \ominus \{x : A \to B\})$) whenever $(A \to B)^{2.\mathrm{mult}(A\to B,\Delta)} \subseteq \Gamma \oplus \{A \to B\}$, or equivalently $(A \to B)^{2.\mathrm{mult}(A\to B,\Delta)} \subsetneq \Gamma \oplus \{A \to B, A \to B\} = \mathrm{ant}(t_0)$ holds.

Given this formulation of the inversion-reduction rewrite-rule E.c, the reduction-sequence for a cut-elimination process in Vestergaard's example would then continue after the weakening reduction-steps succeeding the six cut-elimination reductions as follows:

$$\begin{aligned}
\ldots \quad \longrightarrow^*_{(Weak)} \quad & \mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}(\jmath_n\{((z')^P)^2, z^P, (s^{P\to P})^2\}) \\
& \equiv \mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\mathbf{let}_{y^P}\big(\imath_n\{y^P\}, s^{P\to P}\imath_1^P\{s^{P\to P}\}\big)\{((z')^P)^2, z^P, (s^{P\to P})^2\}\big) \\
\longrightarrow_{(Cont)} \quad & \mathbf{let}_{y^P}\Big(\mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\imath_n\{y^P, ((z')^P)^2, z^P, (s^{P\to P})^2\}\big), \\
& \qquad s^{P\to P}\mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\imath_1^P\{((z')^P)^2, z^P, (s^{P\to P})^3\}\big)\Big) \\
\longrightarrow^{(3)}_{(Cont)} \quad & \mathbf{let}_{y^P}\big(\mathbf{let}_{y^P}(\ldots), s^{P\to P}\imath_1\{(z')^P, (s^{P\to P})^2\}\big) \\
\longrightarrow_{(Cont)} \quad & \mathbf{let}_{y^P}\big(\mathbf{let}_{x_n^P}\big(\mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\mathbf{ax}_{x_n^P;z^P}\{y^P, ((z')^P)^2, (z')^P, (s^{P\to P})^2\}\big), \\
& \qquad s^{P\to P}\mathbf{C}^{\{(z')^P, z^P, s^{P\to P}\}}\big(\imath_{n-1}\{y^P, ((z')^P)^2, z^P, (s^{P\to P})^2\}\big), \ldots\big) \\
\longrightarrow^*_{(Cont)} \quad & \mathbf{let}_{y^P}\big(\mathbf{let}_{x_n^P}\big(\mathbf{ax}_{x_n^P;z^P}\{y^P, (z')^P, s^{P\to P}\}, s^{P\to P}\imath_{n-1}\{y^P, (z')^P, s^{P\to P}\}\big), \ldots\big) \\
& \equiv \mathbf{let}_{y^P}\big(\imath_n\{y^D, (z')^P, s^{P\to P}\}, s^{P\to P}((\imath_1\{s^{P\to P}\})\{(z')^P, s^{P\to P}\})\big) \\
& \equiv \jmath_n\{(z')^P, s^{P\to P}\}.
\end{aligned}$$

(The $\longrightarrow^*_{(Weak)}$-reduction-step in this reduction-sequence is again justified by Lemma 1.9, in the reduction-step $\longrightarrow^{(3)}_{(Cont)}$ means three successive applications of contraction-reduction rewrite-rules. In the last reduction-step, that consists of a gathered number of single contraction-reduction steps, it was used that the contraction involved on the left can then be permuted upwards inductively over all of the $n-1$ applications of L$\to$ in $\imath_{n-1}$ similarly as in the previous step over the bottom-most application L$\to$ in $\imath_n$.)

Here an inversion-reduction step is no longer needed during this process and the form of the resulting term $\tilde{t'_n}$ is not independent of n any more. This means, that if the reduction-sequences $\tilde{\sigma_n}$ for $n \in \mathbb{N}$ of this kind were considered as applications of an operation cf (to make terms cut-free) on derivation-terms t_n (the left-most terms in the sequences), then the identity of derivation-terms from the family $\{t_n\}_{n\in\mathbb{N}}$ would be preserved under this operation cf. Moreover, if the natural-deduction image $\Phi(\tilde{\mathcal{D}'_n})$ under Φ of the derivation $\tilde{\mathcal{D}_n}$ in $\boldsymbol{\mathcal{G}_0^+}$, which corresponds to $\tilde{t'_n}$, is looked at, the derivation $\Phi(\mathcal{D}_n)$ in (1.6) is encountered again. The "meanings" of the terms t_n and $\tilde{t'_n}$ at the start and at the end of $\tilde{\sigma_n}$ are therefore identical.

Closer inspection shows that the whole process of cut-elimination for $\mathcal{D}_n$ is here—under the just slightly modified conditions of a variant of the contraction-reduction rewrite-rule E.c used here instead—undetectable from the natural-deduction images of the derivation-terms (there are also no changes in this natural-deduction images during the course of the derivation, as is easy to check). – The "anomaly" has disappeared here.

Yet, it is possible to give another example of a "computational anomaly" also in the system $\boldsymbol{\mathcal{G}^+}$ with an explicit weakening rule. Here essentially the idea behind the example of Vestergaard is used, but in a slightly different way. The example below was initially found by considering cut-elimination in a rather easy derivation of the untyped **G3[mi]**-system (cf. this example in Appendix A on p. 77), for which the first step in the cut-elimination procedure necessitates the application of a fork-reduction of list C in section 1.3.

The derivation-terms $\imath_n$ (for $n \in \mathbb{N}$) in (1.5) from Vestergaard's example will be used again in the new example. Furthermore let t_{00}^B be the derivation-term

$$t_{00}^B := \mathbf{let}_{(z')^P}\left(\mathbf{ax}_{y^B;x^A,z^P,(z')^P}, s^{P\to P}\mathbf{ax}_{z^P;x^A,y^B,s^{P\to P}}\right) .$$

Then for all $n \in \mathbb{N}$ the following reduction-sequence $\bar{\sigma_n}$ on derivation-terms represents a process of cut-elimination according to Theorem 1.1:

$$\begin{aligned}
&\mathbf{let}_{y^B}\left(\imath_n\{x^A, y^B\}, w^{A\to B}\mathbf{ax}_{x^A;w^{A\to B},z^P,s^{P\to P}}\right)[\![w^{A\to B} := \boldsymbol{\lambda} x^A.t_{00}^B]\!]\\
&\longrightarrow_{(Cut)}\quad \mathbf{C}^{\{x^A,y^B,(z^P)^2,(s^{P\to P})^2\}}\imath_n\{x^A, y^B\}[\![y^B := t_{00}^B[\![z^P := \mathbf{ax}_{x^A;w^{A\to B},z^P,s^{P\to P}}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad [\![w^{A\to B} := \boldsymbol{\lambda} x^A.t_{00}^B]\!]]\!]]\!]\\
&\longrightarrow_{(Cut)}\quad \mathbf{C}^{\{x^A,y^B,(z^P)^2,(s^{P\to P})^2\}}\imath_n\{x^A, y^B\}[\![y^B := t_{00}^B[\![z^P := \mathbf{ax}_{x^A;y^B,(z^P)^2,(s^{P\to P})^2}]\!]]\!]\\
&\longrightarrow_{(Cut)}\quad \mathbf{C}^{\{x^A,y^B,(z^P)^2,(s^{P\to P})^2\}}\imath_n\{x^A, y^B\}[\![y^B := \mathbf{W}^{\{y^B,(z^P)^2,(s^{P\to P})^2\}}(t_{00}^B)]\!]\\
&\longrightarrow^*_{(Weak)}\quad \mathbf{C}^{\{x^A,y^B,(z^P)^2,(s^{P\to P})^2\}}\imath_n\{x^A, y^B\}[\![y^B := t_{00}^B\{y^B, (z^P)^2, (s^{P\to P})^2\}]\!]
\end{aligned}$$

$$
\begin{aligned}
\longrightarrow_{(Cut)} \quad & \mathbf{C}^{\{x^A,y^B,(z^P)^2,(s^{P\to P})^2\}}\mathbf{let}_{(z')^P}\big(\imath_n\{x^A,y^B\} \\
& \qquad\qquad [\![y^B := \mathbf{ax}_{y^B;x^A,y^B,(z^P)^3,(z')^P,(s^{P\to P})^2}]\!], \\
& \qquad\qquad\qquad s^{P\to P}\mathbf{W}^{\{x^A,z^P,s^{P\to P}\}}(\mathbf{ax}_{z^P;x^A,(y^B)^2,(z^P)^2,(s^{P\to P})^3})\big) \\
\longrightarrow_{(Cut)} \quad & \mathbf{C}^{\{x^A,y^B,(z^P)^2,(s^{P\to P})^2\}}\Big(\mathbf{let}_{(z')^P}\big(\mathbf{W}^{\{x^A,y^B,(z^P)^3,(z')^P,(s^{P\to P})^2\}}(\imath_n\{x^A,y^B\}), \\
& \qquad\qquad\qquad s^{P\to P}\mathbf{W}^{\{x^A,z^P,s^{P\to P}\}}(\mathbf{ax}_{z^P;x^A,(y^B)^2,(z^P)^2,(s^{P\to P})^3})\big)\Big) \\
\longrightarrow^*_{(Weak)} \quad & \mathbf{C}^{\{x^A,y^B,(z^P)^2,(s^{P\to P})^2\}}\Big(\mathbf{let}_{(z')^P}\big(\imath_n\{(x^A)^2,(y^B)^2,(z^P)^3,(z')^P,(s^{P\to P})^2\}, \\
& \qquad\qquad\qquad s^{P\to P}(\mathbf{ax}_{z^P;(x^A)^2,(y^B)^2,(z^P)^3,(s^{P\to P})^4})\big)\Big) \\
\longrightarrow_{(Cont)} \quad & \mathbf{let}_{(z')^P}\Big(\mathbf{C}^{\{x^A,y^B,(z^P)^2,s^{P\to P}\}}\big(\\
& \qquad \mathbf{I}_{s^{P\to P},z^P}(\imath_n\{(x^A)^2,(y^B)^2,(z^P)^3,(z')^P,(s^{P\to P})^2\})\big), \\
& \qquad\qquad s^{P\to P}\mathbf{C}^{\{x^A,y^B,(z^P)^2,(s^{P\to P})^2\}}(\mathbf{ax}_{z^P;(x^A)^2,(y^B)^2,(z^P)^3,(s^{P\to P})^4})\Big) \\
\longrightarrow_{(Cont)} \quad & \mathbf{let}_{(z')^P}\big(\ldots, s^{P\to P}\mathbf{ax}_{z^P;x^A,y^B,z^P,(s^{P\to P})^2}\big) \\
\longrightarrow_{(Inv)} \quad & \mathbf{let}_{(z')^P}\big(\mathbf{C}^{\{x^A,y^B,(z^P)^2,s^{P\to P}\}} \\
& \qquad (\mathbf{ax}_{z^P;z^P}\{(x^A)^2,(y^B)^2,(z^P)^3,(z')^P,(s^{P\to P})^2\}),\ldots\big) \\
\longrightarrow_{(Cont)} \quad & \mathbf{let}_{(z')^P}\big(\mathbf{ax}_{z^P;(z^P)^2,s^{P\to P},(z')^P}, s^{P\to P}\mathbf{ax}_{z^P;(z^P)^2,(s^{P\to P})^2}\big)\{x^A,y^B\} \quad .
\end{aligned}
$$

(The $\longrightarrow^*_{(Weak)}$-steps here are again justified as applications of Lemma 1.9.)

Let $\bar{t_n}$ and $\bar{t'}$ denote the topmost and bottom-most derivation-terms in the above derivation-sequence $\bar{\sigma_n}$ respectively. It is straightforward to check that for the $\boldsymbol{\mathcal{G}}_\mathbf{0}^+$+Cut-derivation $\bar{\mathcal{D}_n}$ corresponding to $\bar{t_n}$ the image $\Phi(\bar{\mathcal{D}_n})$ under Φ is equal to $\Phi(\mathcal{D}_{\imath_n})$ ($\mathcal{D}_{\imath_n}$ the $\boldsymbol{\mathcal{G}}_\mathbf{0}^+$-derivation corresponding to $\imath_n$), which for $n \geq 1$ in turn equals (1.6). But the derivation $\bar{\mathcal{D}'}$ corresponding to $\bar{t'}$ has simply P^z as its natural-deduction image $\Phi(\bar{\mathcal{D}'})$.

This means that for $n > 1$ something special has happened here again to the natural-deduction images of derivation-terms in a reduction-sequence $\bar{\sigma_n}$ of the above kind, something, that is not explainable by normalization-steps on the natural-deduction images of the derivation-terms occuring in σ_n. – Closer inspection shows that the one and only jump takes place—as in Vestergaard's example—in the single inversion-reduction step occurring in the reduction-sequence.

1.5 Closer analysis of the "problematic" cut-elimination step in $\mathcal{G}^+$

In the introduction to [Vest99] R. Vestergaard refers to an article by G. Kreisel in [Kr71] of 1971, which was intended by the author to supplement an earlier article with aspects and thoughts springing up from the consequences of the then recent discovery of normalization for natural-deduction derivations by D. Prawitz in the early and mid-1960ies. Kreisel was interested in what exactly the new concept could tell about properties of and relations between *proofs* (as informal objects that mathematicians are familiar with):

> "Here I wish to emphasize formal results and and problems concerning *relations* between *proofs*, for example the *identity relation between proofs described by formal derivations of a given system*."

With respect to "normal derivations and conversions" Kreisel lists the following desirable properties of normal derivations (in a not directly specified formal system, that for example can be both a sequent- or a natural-deduction calculus) such that they can "serve as *canonical representations* of all proofs represented in the system considered, the way the numerals are canonical notations for the natural numbers":

> "A minimum requirement is then that *any derivation can be normalized*, that is transformed into a unique normal form by a series of steps, so-called "conversions", each of which preserves the proof described by the derivation. This requirement has a formal and an informal part:
>
> (α) The *formal* problem of establishing that the conversions terminate in a unique normal form (independent of the order in which they are applied).
>
> (βi) The *informal* recognition (by inspection) that the conversion steps considered preserve identity, and the informal problem of showing that
>
> (βii) distinct, that is incongruent normal derivations represent different proofs (in order to have unique, canonical, representations)."

Vestergaard draws from this a connection to his findings and specifically stresses the above requirement (βi) for this. He goes on to interpret the *proof* thought to be underlying a derivation $\mathcal{D}$ in his typed $\rightarrow$**G3mi**-like system $\mathcal{G}_{\boldsymbol{v}}^+$ as just the natural-deduction image $\Phi(\mathcal{D})$ of $\mathcal{D}$ (to be precise, only for a derivation $\mathcal{D}$ in $\mathcal{G}_{\mathbf{0}}^+$ not containing an application of inversion). If this is done, his example of a "computational anomaly" really shows that the cut-elimination procedure he uses similar to the one for the **G3[mi]**-systems given implicitly in [TS96] indeed changes "proofs", in Kreisel's use of this word, it thus does not preserve the identity of proofs.

But it has to be remembered at the same time that—as Zucker's result from 1974 shows—also the cut-elimination steps in the seemingly much better behaved LJ-near system $\mathcal{S}$ do *not* preserve the so-understood identity of proofs, insofar as the steps for upwards permutation of cut (Zucker calls these steps *permutative conversions*) change the natural-deduction images of derivations; again, if only in a way that can be simulated on the images by a finite sequence of normalization-steps[10].

It seems doubtful, that Kreisel had really wanted to have the meaning of the word "proof" in the quoted passages understood as just a natural-deduction derivation (since this is again a derivation in a strict formal system that can only model the mathematical practice of proving and therefore does not really contain the proofs mathematicians deal with and have in their minds). Nevertheless, in the absence of better suited candidates natural-deduction surely can serve as a very good approximation to the informal notion of proof.

Zucker in [Zu74] takes this as a starting point and provides arguments for the following: If one accepts the view that the performance of normalization-steps on a natural-deduction derivation *does not* change the underlying (informal) proof (a conjecture by D. Prawitz in [Pra71]), that is, that synonymity of derivations is equivalent to interreducibility by normalization-reductions, then the results in [Zu74] can be seen as a justification to interpret—at least for the negative fragment $\mathcal{S}^-$ of Zucker's sequent-system $\mathcal{S}$—the synonymity of derivations in the sequent-system $\mathcal{S}$ as the property of their interreducibility by (appropriately specified) cut-elimination steps in $\mathcal{S}$. (Pottinger in [Pott77] later indicated a way how to generalize Zucker's result to a full intuitionistic calculus.)

Vestergaard's result shows that "synonymity" of derivations in his typed $\rightarrow$**G3mi**-like system $\boldsymbol{\mathcal{G}}^+_v$ cannot be interpreted just as the interreducibility of their natural-deduction images by normalization, if this so-understood "meaning" of derivations shall be preserved by a cut-elimination procedure near to that for **G3[mi]** given implicitly in [TS96] (this result was transferred here to a system comparable to Vestergaard's but with an additional explicit weakening rule).

Prawitz above mentioned conjecture "Two [natural-deduction] derivations represent the same proof if and only if they are equivalent [i.e. interreducible by normalization-steps]" in [Pra71] was challenged by S. Feferman in [Fef75], particularly for the $\forall$-contraction step of normalization. He suggested instead:

> "Even if it does not settle the relation of identity between proofs, the work described by Prawitz may give simple syntactic explanations of other familiar relations and operations, for example, for the idea of *one proof specializing to another* or *of extracting from a proof just what is needed for its particular conclusion*."

[10] It is easy to check that the same is true in the here considered system $\boldsymbol{\mathcal{G}}^+$ for the derivation-reductions corresponding to the rewrite-rules of list B in section 3 for upwards-permutation of Cut.

Following Feferman's above words here an argument shall be given, why it can be thought, that even the problematic step during a cut-elimination process in the system $\boldsymbol{\mathcal{G}}^{+}$ as well as in $\boldsymbol{\mathcal{G}}_{v}^{+}$, namely upwards-permutation of contraction that needs an additional application of inversion (cf. the contraction-reduction rewrite-rule E.c, second case, that is the immediate cause of the anomaly), does in fact lead to the extraction from a given derivation $\mathcal{D}$ in $\boldsymbol{\mathcal{G}}_{0}^{+}$+Cont just what is needed to prove its conclusion. In the more precise sense that only such unnecessary subderivations are getting "axed out" as a consequence of this step during contraction-elimination, for which another subderivation leading to a stronger conclusion stays in the transformed derivation.

For simplicity let us consider the "problematic" step in the cut-elimination procedure for $\boldsymbol{\mathcal{G}}^{+}$ in the case of an analogous step that occurs during the cut-elimination procedure implicit in [TS96] for the untyped $\to$**G3mi**-system. There during a subprocess of contraction-elimination the following situation of a derivation $\mathcal{D}$

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_0\\ A \to B, A \to B, \Gamma \Rightarrow A\end{array} \qquad \begin{array}{c}\mathcal{D}_1\\ B, A \to B, \Gamma \Rightarrow C\end{array}}{A \to B, A \to B, \Gamma \Rightarrow C}\,\mathrm{L}{\to}}{A \to B, \Gamma \Rightarrow C}\,\mathrm{C} \tag{1.8}$$

can occur. Here the contraction at the bottom of $\mathcal{D}$ cannot be directly permuted upwards over L$\to$, at least not in the case, when $A \to B \notin \Gamma$. Here $\mathcal{D}$ gets transformed by the procedure in a first step to

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_0\\ A \to B, A \to B, \Gamma \Rightarrow\ A\end{array}}{A \to B, \Gamma \Rightarrow\ A}\,\mathrm{C} \qquad \dfrac{\dfrac{\begin{array}{c}\mathcal{D}_1\\ A \to B, B, \Gamma \Rightarrow\ C\end{array}}{B, B, \Gamma \Rightarrow\ C}\,\mathrm{Inv}}{B, \Gamma \Rightarrow\ C}\,\mathrm{C}}{A \to B, \Gamma \Rightarrow\ C}\,\mathrm{L}{\to} \tag{1.9}$$

by the use of an additional application of inversion. The removal of the inversion at the bottom of $\mathcal{D}_1$ is then responsible for the unwanted effects described by Vestergaard's anomaly, since during this operation of inversion-elimination some sub-derivations of $\mathcal{D}_1$ can disappear completely (in the setting of the typed calculus $\boldsymbol{\mathcal{G}}^{+}$ in an analogous situation some subderivations can equally get lost when going over to the respective natural-deduction images, an effect, which ultimately leads to the "anomalies"). It shall be tried to argue here, that while in a proof thought to be underlying (1.8) as an informal object indeed some sub-proofs, corresponding to subderivations of $\mathcal{D}_1$, can be lost as a consequence of inversion-reduction steps in $\mathcal{D}_1$ following after the situation (1.9), such lost subproofs lead only to weaker versions of the sequent $A \to B, \Gamma \Rightarrow A$, a proof of which must then still underly the transformed derivation $\mathcal{D}_0$ after the elimination of the succeeding contraction in (1.9) there.

To be able to demostrate this, two lemmas are necessary, the first analyzes the structure of $\mathcal{D}_1$ in (1.9) and the second describes the result of eliminating the newly appearing inversion there completely from the left immediate subderivation of L→.

Lemma 1.10. *Every derivation $\mathcal{D}_1$ in* **→G3mi** *with the conclusion $B, A \to B, \Gamma \Rightarrow C$ has the structure*

$$\begin{array}{ccc} \mathcal{D}_{11} & & \mathcal{D}_{1n} \\ (A \to B, \Gamma_1\Pi_1\Sigma_1 \Rightarrow C_1) & \dots & (A \to B, \Gamma_n\Pi_n\Sigma_n \Rightarrow C_1) \\ & \mathcal{D}_{10} & \\ & B, A \to B, \Gamma \Rightarrow C & \end{array} \tag{1.10}$$

(here a sequent in brackets $(\dots)$ means exactly one occurrence of this sequent as a leaf in the derivation-tree of the "partial derivation" $\mathcal{D}_{10}$), where

(i) each derivation $\mathcal{D}_{1i}$ $(i = 1, \dots, n)$ is either an axiom or terminates with an application of L→ and is hence of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_{1i0} \\ A \to B, \Gamma_i\Pi_i\Sigma_i \Rightarrow A\end{array} \qquad \begin{array}{c}\mathcal{D}_{1i1} \\ B, \Gamma_i\Pi_i\Sigma_i \Rightarrow C_i\end{array}}{A \to B, \Gamma_i\Pi_i\Sigma_i \Rightarrow C_i}\,L\!\to \quad , \tag{1.11}$$

where furthermore for all $i \in \{1, \dots, n\}$

$$\begin{array}{l} \Gamma_i \subseteq \Gamma \quad \textit{and} \\ \forall D \in \Gamma \cup \{B\}\big[\, D \in \Gamma_i \vee \\ \qquad\qquad \big(\exists D' \textit{ strictly positive subformula of } D\,\big)(\mathcal{D}' \in \Pi_i)\big] \end{array} \tag{1.12}$$

(but no property of the formulas in Σ_i is singled out here).

(ii) $\mathcal{D}_{10}$ is a partial derivation (i.e. it contains sequents $A \to B, \Gamma_i\Pi_i\Sigma_i \Rightarrow C_i$ as top-leafs, that are not necessarily axioms), which contains no axioms except ones that occur among the $\mathcal{D}_{11}, \dots, \mathcal{D}_{1n}$ and no L→-application with principal formula $A \to B$.

Lemma 1.11. *The result $\mathcal{D}_1'$ of eliminating a bottom-most application of inversion according to the cut-elimination procedure implicit in [TS96][11] in the* **→G3mi**-*derivation $\tilde{\mathcal{D}}_1$ of the form*

$$\begin{array}{ccc} \mathcal{D}_{11} & & \mathcal{D}_{1n} \\ (A \to B, \Gamma_1\Pi_1\Sigma_1 \Rightarrow C_1) & \dots & (A \to B, \Gamma_n\Pi_n\Sigma_n \Rightarrow C_n) \\ & \mathcal{D}_{10} & \\ & \dfrac{B, A \to B, \Gamma \Rightarrow C}{B, B, \Gamma \Rightarrow C}\,Inv & \end{array} \quad , \tag{1.13}$$

[11]Cf. the item (vii) of the inversion-lemma in the Proposition on p. 66, 67 and its proof in [TS96], on which the proof of the cut-elimination theorem for the **G3[mi]**-systems relies for the treatment of such a subproblem.

where the immediate subderivation $\mathcal{D}_1$ of the bottom-most inversion in $\tilde{\mathcal{D}}_1$ is of the form (1.10) *with the conditions on $\mathcal{D}_{10}, \mathcal{D}_{1i}, \Gamma_i, \Pi_i, \Sigma_i$ as in Lemma 1.10, is*

$$\begin{array}{ccc} \mathcal{D}'_{11} & & \mathcal{D}'_{1n} \\ (B, \Gamma_1\Pi_1\Sigma_1 \Rightarrow C_1) & \dots & (B, \Gamma_n\Pi_n\Sigma_n \Rightarrow C_n) \\ & \mathcal{D}_{10}\{A \to B/B\} & \\ & B, B, \Gamma \Rightarrow C & \end{array} \tag{1.14}$$

(where $\mathcal{D}_{10}\{A \to B/B\}$ means the result of replacing exactly one occurrence of $A \to B$ in the antecedent of every sequent by B) and where for all $i = 1, \dots, n$

$$\mathcal{D}'_{1i} := \begin{cases} \mathcal{D}_{1i1} & \dots \textit{ if } \mathcal{D}_{1i} \textit{ is not an axiom} \\ \textit{(the axiom:) } B, \Gamma_i\Pi_i\Sigma_i \Rightarrow C & \dots \textit{ if } \mathcal{D}_{1i} \textit{ is an axiom} \end{cases} \tag{1.15}$$

($\mathcal{D}_{1i0}$ and $\mathcal{D}_{1i1}$ mean the subderivations of $\mathcal{D}_1$ with these denotations from Lemma 1.10).

The *Proofs* of these two lemmas consist just of—appropriately formulated—inductions on the depth of the derivation $\mathcal{D}_1$.

The derivation $\mathcal{D}_1$ in (1.9) can by Lemma 1.10 be seen to be of the form (1.10) with the derivations $\mathcal{D}_{1i}$ of the form (1.11) and the conditions on $\mathcal{D}_{10}, \mathcal{D}_{1i}, \Gamma_i, \Pi_i, \Sigma_i$ as in Lemma 1.10.

By Lemma 1.11 the result of eliminating the inversion appearing in (1.9) is then

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_0 \\ A \to B, A \to B, \Gamma \Rightarrow A\end{array}}{A \to B, \Gamma \Rightarrow A}\,\mathrm{C} \qquad \dfrac{\begin{array}{c}\mathcal{D}'_1 \\ B, B, \Gamma \Rightarrow C\end{array}}{B, \Gamma \Rightarrow C}\,\mathrm{C}}{A \to B, \Gamma \Rightarrow C}\,\mathrm{L}{\to} \tag{1.16}$$

where $\mathcal{D}'_1$ is of the form (1.14) (with subderivations $\mathcal{D}'_{1i}$ for $i = 1, \dots n$ as defined in Lemma 1.11). As described by Lemma 1.11 $\mathcal{D}'_1$ is the result of dropping subderivations

$$\begin{array}{c} \mathcal{D}_{1i0} \\ A \to B, \Gamma_i\Pi_i\Sigma_i \Rightarrow A \end{array} \tag{1.17}$$

from $\mathcal{D}_1$ and of replacing a single passive occurrence of $A \to B$ in every sequent throughout $\mathcal{D}_{10}$ by B.

Due to (1.12) the conclusion of every derivation (1.17) is obviously weaker than the conclusion $A \to B, \Gamma \Rightarrow A$ of the derivation leading to the left premise of L$\to$ in (1.16). It seems therefore justifiable to say that while removing the additional occuring application of inversion in (1.9) leads in effect to the loss of subderivations in the result (1.16) of this subprocess, these lost derivations would correspond only to weaker versions of a sub-proof that stays in the transformed derivation (that one that underlies the proof of the left immediate subderivation of L$\rightarrow$ in (1.16)).

In other words, although the subderivations $\mathcal{D}_{1i0}$ in (1.17) of $\mathcal{D}_1$ disappear as a consequence of the upwards-permutation of contraction over L$\rightarrow$ in (1.8) via the first step (1.9) and the following inversion-reduction steps, the process as a whole may be seen here as keeping back only one essential copy of a derivation for $A \rightarrow B, \Gamma \Rightarrow A$ (underlying the further transformed derivation $\mathcal{D}_0$), while possibly many proofs of "special cases" $A \rightarrow B, \Gamma_i \Pi_i \Sigma_i \Rightarrow A$ (where Γ_i, Π_i satisfy (1.12)) of $A \rightarrow B, \Gamma \Rightarrow A$ get "axed out" from $\mathcal{D}_1$. The process of cut-elimination in this contraction-elimination step could then be thought of as dropping applications of unnecessary lemmas from the proof underlying $\mathcal{D}$, that turned out to be special cases of a statement, for which a proof is retained. Or, as extracting from the proof, that is thought to be formalized by $\mathcal{D}$, essential parts for a derivation of its conclusion, using the distinctive combinatorial properties of the L$\rightarrow$-rule in an intuitionistic or minimal **G3[mi]**-system.

1.6 An alternative system $\rightarrow$G2$'$mi^{e*}

It is not apparent from the outset, why Vestergaard chose to present his result in the setting of a typed system, where the antecedents of sequents are considered to be *multisets* of variable-annotated formulas (instead of as respective such *sets*). He only stresses that he is interested in the "computational meaning" of single rules and of derivations in the (untyped) **G3[mi]**-sytems (and whether such a precisely definable "computational meaning" of a derivation is affected or not during the execution of cut-elimination steps), where the antecedents of sequents in fact are multisets of (not annotated) formulas.

On the other hand the logical rules in Vestergaard's system (as in the very similar system $\mathcal{G}^+$ above) *do* act on their premises by treating multiple occurrences $[x : A]$ of annotated active formulas $x : A$ in a set-like way as one object for the rule-applications (with the obvious unstated motivation of considering such multiple occurrences as referring to the same assumption class in a corresponding natural-deduction derivation).

Also J. Zucker in his ground-breaking paper [Zu74] about the exact relationship between cut-elimination in a sequent-calculus and normalization in a related natural-deduction system took a sequent-calculus $\mathcal{S}$ as the basis for his investigation, in which the antecedents of sequents consist of *sets* of (precisely defined:) "indexed" formulas and where the rules were formed appripriately for this notion of sequents. Zucker's system $\mathcal{S}$ is (as to the logical shape of its rules, not with respect to the special indexing conventions used in it) close to Gentzen's LJ and could be easily transformed into a typed system, such that his results of a close correspondence between cut-elimination steps in $\mathcal{S}$ and normalization steps in $\mathcal{N}$ (his slightly modified system for natural deduction) would carry over to the typed system (if variable annotations and indexes of formulas were corresponding to each other bijectively, the typed system therefore then being only the result of rewriting $\mathcal{S}$).

In the conclusion of [Vest99] Vestergaard states that the "anomaly" could have been avoided by using a variant-system instead of (the here called system) $\boldsymbol{\mathcal{G}}_{\boldsymbol{v}}^{+}$ with the antecedents of sequents consisting of sets instead of multisets:

> "The computational anomaly could [...] have been avoided if we instead had considered (a variant) of **G3i** incorporating a notion of assumption classes. This can be accomplished, e.g. by defining antecedents to be sets of variables with a proposition annotated. In such a setup we could have utilized the implicit contraction which is expressed in the idempotency of set union to take place of the trouble instigator in **G3i**: the explicit contraction rule."

It could be argued in more detail, why we think that this is not, at least not directly possible: (1) In an axiomatic cut-elimination step comparable to the case covered by the rewrite-rule A.a a lemma for substitution (comparable to Lemma 1.4 and Lemma 1.6) has to be relied on also in a new system, but such lemmas do not hold any more in the case, that the axioms are not restricted to such that have only atomic active formulas (as in Vestergaard's system $\boldsymbol{\mathcal{G}}_{\boldsymbol{v}}^{+}$); else completely similar problems as are the cause of the "anomaly" could come in from the seemingly harmless case of an axiomatic cut-reduction in an new system comparable to the reduction A.a on derivation-terms in $\boldsymbol{\mathcal{G}}^{+}$. (2) One encounters difficulties with the treatment of upwards permutation of Cut in the case of non-principal cut-formulas (reductions on derivations similar to those given in list B for reductions on derivation-terms). Difficulties, that may indeed let Vestergaard's system $\boldsymbol{\mathcal{G}}_{\boldsymbol{v}}^{+}$ (and the here defined system $\boldsymbol{\mathcal{G}}^{+}$) with the antecedents of sequents consisting of multisets look as a quite natural choice for a typed $\rightarrow$**G3mi**-like calculus, that relates derivations in it quite naturally to natural-deduction derivations and at the same time allows to describe cut-elimination for derivations in it (to be precise, cut-elimination done close to the usual way for a **G3[mi]**-system) as a stepwise and local process.

Vestergaard's suggestion cited above can be carried through for the following typed system $\rightarrow$**G2′mi**^{e*}, which has a very $\rightarrow$**G3mi**-like formulation of its L$\rightarrow$-rule and is itself a **G3**-system (this means, contraction is an admissible rule for it), but which perhaps derives more from the type-annotation of a **G2**-system $\rightarrow$**G2′mi** (in which contraction is not an admissible rule any more—due to the fact that the context in the premises of L$\rightarrow$ is not the same any more as in the respective rule for the **G3[mi]**-systems). The explicit structural rules in $\rightarrow$**G2′mi**^{e*} have also (as in $\boldsymbol{\mathcal{G}}_{\boldsymbol{v}}^{+}$ and $\boldsymbol{\mathcal{G}}^{+}$) been defined just in such a way so as to make cut-elimination possible as a stepwise process of locally applied transformations (the role of contraction is taken over by a renaming-rule Ren, where two successive renamings always suffice to mimick an arbitrary given contraction).

Definition 1.6 (The derivation-term annotated system $\rightarrow$G2′mi^{e*}). The system $\rightarrow$**G2′mi**^{e*} is defined as follows: The antecedent of a sequent in this system is a *set* of variables of formula-type (written as variable-annotated formulas), the succedent consists of a (rigidly) typed derivation-term, whose free type-variables occur in the antecedent. $\rightarrow\mathbf{G2'mi}^{e*}$ has the following axioms and rules:

$$\text{Ax}\quad x : P;\, \Gamma \Rightarrow \mathbf{ax}_{x^P;\Gamma} : P \quad (P \text{ atomic})$$

$$\text{R}\!\rightarrow \frac{x : A;\, \Gamma \Rightarrow t : B}{\Gamma \Rightarrow \boldsymbol{\lambda} x^A.t^B : A \rightarrow B}$$

$$\text{L}\!\rightarrow \frac{x : A \rightarrow B;\, \Gamma_0 \Rightarrow t_0 : A \qquad y : B;\, \Gamma_1 \Rightarrow t_1 : C}{x : A \rightarrow B, \Gamma_0, \Gamma_1 \Rightarrow \mathbf{let}_{y^B}(t_1^C, x^{A\rightarrow B} t_0^A) : C}$$

$$\text{Ren}\, \frac{u : A;\, \Gamma \Rightarrow t : C}{v : A, \Gamma \Rightarrow \mathbf{Ren}_{u^A, v^A}(t^C) : C}$$

$$\text{mW}\, \frac{\Gamma \Rightarrow t : C}{\Delta, \Gamma \Rightarrow \mathbf{W}^{\Delta}(t^C) : C}$$

Here the following notations are used:

- Expressions Γ_0, Γ_1 or $\Gamma_0\Gamma_1$ mean the set $\Gamma_0 \cup \Gamma_1$, expressions $x : A, \Gamma$ denote the set $\{x : A\} \cup \Gamma$, and $x : A; \Gamma$ is also to be understood as $\{x : A\} \cup \Gamma$, but thereby $x : A$ is understood to be no element of Γ.
- As in Definition 1.2 typed variables x^A are also written as $x : A$ (when they occur in the antecedent and thereby informally refer to marked assumptions of a corresponding natural-deduction derivation). Terms t^C in the succedent of a sequent are written in the form $t : C$ (since they are often informally thought of as derivation-terms describing a natural-deduction derivation with conclusion C).

The term in the succedent of the end-sequent of a $\rightarrow\mathbf{G2'mi}^{e*}$-derivation $\mathcal{D}$ will be called the *derivation-term* of $\mathcal{D}$.

The systems $\rightarrow\mathbf{G2'mi}^{e*}$+Cut are the systems $\rightarrow\mathbf{G2'mi}^{e*}$ enriched with the cut-rule

$$\text{Cut}\, \frac{\Gamma \Rightarrow t_0 : D \qquad x : D; \Pi \Rightarrow t_1 : C}{\Gamma\Pi \Rightarrow t_1^C [\![x^D := t_0^D]\!] : C}$$

as an additional inference rule. ⊠

Derivation-terms in this system again represent derivations uniquely (due to the rigid typing in these terms). For cut-elimination in this system it is necessary that it is possible to reconstruct the antecedents $\text{ant}(t)$ of the conclusion-sequent $\Gamma \Rightarrow t : C$ of a derivation from the derivation-term t. Like in $\boldsymbol{\mathcal{G}}^+$ this can be done inductively using the following definition:

Definition 1.7. The operation ant on $\to\mathbf{G2'mi}^{e*}$-derivation is defined as follows

$$\begin{aligned}
\mathrm{ant}(\mathbf{ax}_{x^P;\Gamma}) &:= \{x : P\} \cup \Gamma \ \ ; \\
\mathrm{ant}(\boldsymbol{\lambda} x^A.t^B) &:= \mathrm{ant}(t^B) \setminus \{x : A\} \ \ ; \\
\mathrm{ant}(\mathbf{let}_{y^B}(t_1^C, x^{A\to B}t_0^A)) &:= \mathrm{ant}(t_0^A) \cup (\mathrm{ant}(t_1^C) \setminus \{y : B\}) \ \ ; \\
\mathrm{ant}(\mathbf{Ren}_{u^A,v^A}(t^C)) &:= (\mathrm{ant}(t^C) \setminus \{u : A\}) \cup \{v : A\} \ \ ; \\
\mathrm{ant}(\mathbf{W}^\Delta(t^C)) &:= \mathrm{ant}(t^C) \cup \Delta \ \ ; \\
\mathrm{ant}(t_1^C[\![x^D := t_0^D]\!]) &:= \mathrm{ant}(t_0^D) \cup (\mathrm{ant}(t_1^C) \setminus \{x : D\}) \ \ .
\end{aligned}$$

Again, outermost types of terms on the left sides of the definition have been dropped here, whenever it is possible to reconstruct them. ⊠

Then the following lemma holds:

Lemma 1.12. *For every ($\to\mathbf{G2'm}^{e*}$+Cut)-derivation-term t^C there is exactly one derivation $\mathcal{D}$ in ($\to\mathbf{G2'm}^{e*}$+Cut) such that $\mathcal{D}$ is of the form*

$$\begin{array}{c} \mathcal{D} \\ \Gamma \Rightarrow t : C \end{array}$$

(Γ a set of formulas); for this derivation $\mathcal{D}$ moreover $\Gamma = ant(t^C)$ holds.

Proof. Again (as for Lemma 1.3) by induction on the syntactical depth of t^C, thereby inspecting all rules of $\to\mathbf{G2'mi}^{e*}$ for the induction-step. □

Theorem 1.2. *Cut-elimination holds for $\to\mathbf{G2'mi}^{e*}$.*

More precisely, every derivation $\mathcal{D}$ in ($\to\mathbf{G2'mi}^{e*}$+Cut) can be transformed by a finite sequence of successively applied local reduction-steps with the result of a cut-free derivation in $\to\mathbf{G2'mi}^{e*}$ containing no applications of structural rules mW or Ren.

Furthermore the process of cut-elimination for a derivation $\mathcal{D}$ in $\to\mathbf{G2'mi}^{e*}$ can be completely simulated on derivation-terms by applications of rules from an appropriate rewrite-rule system starting at the derivation-term t of $\mathcal{D}$; these rule-applications have to respect a certain order, in which single rewrite-rule steps are successively executed.

The *Proof* of this theorem is very similar to that of Theorem 1.1. The rule Ren turns out to be eliminable from the bottom of a derivation containing only logical rules by upwards-permuation over logical rules without the need to introduce other non-logical rules of $\to\mathbf{G2'mi}^{e*}$. Nearly the same applies for mW, since here applications of Ren are needed to make upwards-permutation of mW possible[12]. Cut-elimination can then be done as a process of local transformation-steps referring to the subprocesses of eliminating applications of Ren and mW from cut-free derivations in $\to\mathbf{G2'mi}^{e*}$.

[12](in some cases of upwards-permutation of mW over R$\to$)

It seems that similar problems as with inversion, that is the obvious direct cause of the "computational anomalies", in $\mathcal{G}^+$ are completely avoided in $\to\mathbf{G2'mi}^{e*}$. This is on the other hand but not so astonishing, since the system $\to\mathbf{G2'mi}^{e*}$ is actually very near to Zucker's system $\mathcal{S}$ and would even completely[13] match with it, if a variant of it were considered with similar restrictions on the indices as in $\mathcal{S}$ (the closeness of $\to\mathbf{G2'mi}^{e*}$ to $\mathcal{S}$ can be seen very good from a slight reformulation of $\mathcal{S}$ by A.M. Ungar in [Ung92], Appendix A, starting on p. 186); for such a variant-system the correspondence between cut-elimination and normalization on normal-deduction derivations (with respect to an appropriately defined map Φ as in section 2) holds again as in Zucker's system.

[13] The **G3[mi]**-like formulation of the L$\to$-rule still stands out a bit then, but [I think, C.G.] this does not cause similar problems as in $\mathcal{G}^+$ and $\mathcal{G}_v^+$.

Chapter 2

Strong Cut-Elimination

G. Gentzen devised sequent-calculi for classical and intuitionistic predicate logic as systems that are equivalent to related[1] natural deduction calculi and that allowed him to obtain important formal results about the possible structure of proofs for arbitrary given provable conclusions; results that could then be "exported back" to the natural deduction systems (which were also mainly developed by Gentzen and can be considered as precise formal proof systems very near to the actual mathematical practice and therefore were and are of great foundational interest) to gain deep metamathematical importance.

For establishing the equivalence of sequent- and natural deduction calculi formalizing the same logic (equivalence in the sense that the same theorems are provable in these systems) a certain rule in sequent calculi, the cut rule, suggested itself as being useful and necessary and was introduced for that purpose by Gentzen. His main formal result for sequent-calculi, the "Hauptsatz", states that applications of the cut-rule in derivations either of his sequent-calculi LK for classical or LJ for intuitionistic predicate logic can be effectively removed and the derivation itself can be transformed into a cut-free form (i.e. one in which the cut-rule does not occur any more). This can be carried through with the help of an effective cut-elimination procedure that proceeds by the stepwise execution of local simplifications (i.e. reductions) to a given derivation containing applications of cut and arrives at a cut-free proof of the same conclusion after a finite number of such reduction-steps. The proof-reductions used in this procedure can be completely specified as to the exact conditions of their applicability and to the result produced by them and they can be gathered and listed into a finite catalogue of such steps.

Cut-elimination procedures such as the one implicit in Gentzen's proof of the "Hauptsatz" for LJ and LK prescribe a certain order (as well as many other similar procedures for related and different sequent calculi do the same) in which these reduction-steps have to be applied to a given derivation containing cut such as to then guarantee termination.

[1](this means: formalizing the same logics)

Usually first topmost occurrences of cut[2] in a derivation are considered and treated, are either removed right away whenever this is possible (e.g. in the case when an axiom is involved as a premise) or are permuted upwards over logical or structural rules or split into two or more applications of cut (of a somehow simpler kind); ("somehow simpler" means that:) in all cases a parameter associated with the number, places and forms of the applications of cut in a derivation $\mathcal{D}$ has to be seen to have decreased strictly while performing such a reduction-step, so that termination of the whole procedure can then be seen directly given the domain of this derivation-associated parameter concerning the applications of cut in $\mathcal{D}$ is well-founded.

There is often some indeterminateness left in the process, but this is mostly narrowed down considerably to either the choice of an arbitrary topmost cut or to the possibility that at some place in the derivation perhaps two or more reduction-steps can be chosen and is generally very far from allowing the execution of arbitrary applicable steps to a given derivation (at perhaps even arbitrary places therein) from the procedure's catalogue of reductions. That is, proofs for cut-elimination usually do not also show termination of a related procedure P_1, in which the possible reduction-steps implicit in the procedure P can be applied to a given derivation containing cut in an arbitrary, only by the general conditions of their applicability restricted order.

While Gentzen had found himself lead to the introduction of sequent-calculi for the purpose of proving his outstanding foundational results, D. Prawitz ([Pra65]) discovered a more direct possibility of arriving at basically the same metamathematical results by considering natural deduction calculi alone and by giving a structural proof-theory of these systems without (from the outset:) reference to sequent-calculi. He gave—in some ways—a similar procedure to cut-elimination in sequent-calculi that allowed to construct "direct", then called *normal* natural deduction derivations (i.e. derivations that can roughly be described as ones that avoid to go unnecessary "detours") when starting out from given arbitrary such derivations (in one of Gentzen's natural deduction systems NK for classical and NJ for intuitionistic logic). This procedure for the *normalization* of natural deduction derivations can (like a cut-elimination procedure) also be considered as consisting of the executions of atomic reduction-steps, steps that again can be completely described as to their exact outlook and the very precise circumstances of their applicability; they also can be gathered to form a short list of different types of reductions. Prawitz' original normalization-procedure demanded that these possible reduction-steps have to be applied to a given natural deduction derivation in a completely specified order that is determined by the procedure (namely always treating the rightmost, topmost and longest "detour" in the given derivation first and either removing it completely or decreasing it in its length). – Using the close connection between sequent- and natural-deduction-calculi

[2](in the case of the procedure implicit in Gentzen's proof a generalization of Cut, the mix-rule Mix, comes in)

Prawitz was then also able to give an alternative, though now more indirect proof for Gentzen's "Hauptsatz" in the sequent calculi LK and LJ and he did state this result as a corollary to his Normalization Theorem.

Later Prawitz recognized that the stepwise execution of arbitrary but applicable reduction-steps from his first found normalization-procedure to a given natural-deduction-derivation ultimately always leads to a normal proof after the execution of finitely many of such steps; the normal proof thereby constructed by the normalization-procedure is also unique (if some additional easy simplifications and transformations are observed). Prawitz thereby obtained what was then called a *Strong Normalization Theorem* ([Pra71]).

Although Prawitz' normalization result for natural deduction calculi had allowed him to arrive also at an alternative way of performing cut-elimination in the related sequent-calculi and although there appeared to be "obvious similarities" between normalization and cut-elimination as methods to obtain normal-forms for proofs, the question as to their exact connection or correspondence, whether they "really are the same thing" (all cited words here are from [Pott77]), was still unanswered. Cut-elimination had been seen to admit simulation through normalization by Prawitz, but this did only show that cut-elimination as a whole completed process could be done quite differently and did not tie these two concepts for constructing normal-forms of proofs together closely enough by giving a precise correspondence between reduction steps in either of these methods with each other.

A thorough investigation of the exact relationship between cut-elimination and normalization with respect to intuitionistic calculi was presented in [Zu74]. J. Zucker took a variant $\mathcal{S}$ of Gentzen's LJ, namely a version with the antecedents of sequents consisting of indexed formulas, as the starting point of his investigations. He defined a many-to-one map ϕ from his intuitionistic sequent-calculus $\mathcal{S}$ to $\mathbf{Ni}$ and was then able to prove that there exists a mutual correspondence under ϕ between "natural" cut-elimination steps (as such Zucker saw the ones also used by Gentzen) in $\mathcal{S}^-$, the negative fragment of $\mathcal{S}$, and normalization steps in $\mathbf{Ni}^-$, $\mathbf{Ni}$'s negative fragment. – His detailed analysis made it possible for Zucker to show that every *cut-elimination* or *strong cut-elimination theorem* for $\mathcal{S}^-$ implies a *normalization* or respectively a *strong normalization theorem* for $\mathbf{Ni}^-$ and vice versa.

Zucker was also able to extend his results to the full calculi $\mathcal{S}$ and $\mathbf{Ni}$, but only at the expense of having to deviate from some of what he saw are Gentzen's altogether very natural cut-elimination steps (and of having instead to employ somewhat "unnatural" ones[3]). G. Pottinger in [Pott77] gave an alternative approach to Zucker's results and also extended these to the full respective proof-systems for intuitionistic predicate logic by again giving some new cut-elimination steps not previously used (it is meant: not

[3]This concerns such cut-eliminations steps that deal with the permutation of cut upwards over introduction-rules for $\vee$ and $\exists$.

presented) by Gentzen. But he insisted that at least the fact that these few alternative cut-elimination steps facilitate a direct correspondence between cut-elimination in Zucker's $\mathcal{S}$ and normalization in **Ni** makes them appear perhaps even more natural than the respective ones utilized by Gentzen (for proof-transformation in similar situations).

In 1977 A.G. Dragalin (for an exposition see [Drag79]) gave a quite different and self-contained proof of the fact that a strong form of the cut-elimination theorem also holds for Gentzen's calculi LK and LJ, and namely in the interesting sense that this is even true w.r.t. exclusively such cut-elimination steps as had already been used by Gentzen.

The results of Zucker and Pottinger (here stated without proof:) *do* carry over to the minimal and intuitionistic sequent-calculi **GK3[mi]** with implicit structural rules as originally developed by S.C. Kleene (but here used in the notation as well as in the presentation of these calculi from [TS96]) and thereby also allow to establish an analogous correspondence between cut-elimination in **GK3[mi]** and normalization in **N[mi]**.

But as was explained earlier in Chapter 2, a Zucker-type correspondence does not exist between the **N[mi]**- and the **G3[mi]**-systems, again calculi without explicit structural rules and presented in [TS96] (*these* **G3**-systems[4] in a somewhat different presentation are mainly due to A.G. Dragalin, but were reformulated with only one formula in the succedent—as this is a more common formulation of intuitionistic systems than Dragalin's—by A.S. Troelstra). Hence strong normalization for **N[mi]** does not—at least not in an obvious way—carry over to yield a strong cut-elimination theorem for **G3[mi]**. – On the other hand Dragalin's proof for strong cut-elimination in LJ and LK does not directly apply to the **G3**-systems (since it was intentionally specified to cover LJ and LK—with weakening and contraction rules present there—and only cut-elimination steps already used by Gentzen).

2.1 A Strong Cut-Elimination Theorem for →G3mi and →⊥G3i

In Gentzen's procedure for cut-elimination in the sequent-calculi LK and LJ the reduction steps applied to a derivation $\mathcal{D}$ containing Mix[5] w.r.t. a topmost occurrence S of Mix are essentially *local*; this means they do not involve operations to be applied to whole subderivations in $\mathcal{D}$ (more precisely such subderivations ending more than one rule application above the premise of S) but do only combine immediate subderivations of S and other such subderivations ending not more than one rule application above S in a new way (which can mean some such subderivations are being dropped altogether) and with some few rule applications being added at the bottom of an appropriate combination of

[4]The designation $G3$ for a Gentzen-system without explicit structural rules originated with S.C. Kleene's system of this name in [Kl52].

[5]Which essentially takes over the role of Cut in his proof.

subderivations such that S 's conclusion-sequent is reached again. In the added rule applications extensive use is made of LJ and LK 's structural rules (weakening, contraction and—since Gentzen used lists as antecedents and succedents of sequents—also exchange).

In the notation and formulation of sequent-calculi for minimal, intuitionistic and classical predicate logic according to [TS96] the structural rules weakening and contraction that appear in the basic **G1**-systems have been completely absorbed into the calculi in the **G3**-systems. This means they are not longer present as explicit derivation-rules but can be proven to be derived (or "admissible") rules of the calculi, i.e. lemmas about derivability valid in the systems **G3[mic]** .

Definition 2.1 (The Gentzen systems $\to$G3mi^e, $\to\bot$G3i^e). The variant $\to\bot$**G3i**^{e} of **G3i** 's absurdity-containing implicative fragment $\to\bot$**G3i** with explicit structural and inversion rules is specified by the following axioms and rules:

$$\text{Ax}\quad P,\Gamma \Rightarrow P \quad (P \text{ atomic}) \qquad\qquad \text{L}\bot\quad \bot,\Gamma \Rightarrow A$$

$$\text{L}\!\to \frac{A \to B,\Gamma \Rightarrow A \qquad B,\Gamma \Rightarrow C}{A \to B,\Gamma \Rightarrow C} \qquad\qquad \text{R}\!\to \frac{A,\Gamma \Rightarrow B}{\Gamma \Rightarrow A \to B}$$

$$\text{W}\,\frac{\Gamma \Rightarrow C}{A,\Gamma \Rightarrow C}$$

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

$$\text{Inv}\,\frac{A \to B,\Gamma \Rightarrow C}{B,\Gamma \Rightarrow C}$$

The variants $\to$**G3m**^{e} and $\to$**G3i**^{e} of the implicative fragments $\to$**G3m** and $\to$**G3i** of **G3m** and **G3i** with explicit structural and inversion rules are defined just as $\to\bot$**G3i**^{e}, but with all axioms L$\bot$ left out. (Since $\to$**G3m** and $\to$**G3i** mean the same formal system, also $\to$**G3m**^{e} and $\to$**G3i**^{e} are identical calculi and will be together referred to as $\to$**G3mi**^{e}.

L$\to$ and R$\to$ will be called the *logical rules*, weakening W and contraction C the *structural rules* of the systems defined here.

The systems $\to$**G3mi**^{e} and $\to\bot$**G3i**^{e} will sometimes be enlarged by the adding the cut-rule

$$\text{Cut}\,\frac{\Gamma \Rightarrow D \qquad D,\Pi \Rightarrow C}{\Gamma\Pi \Rightarrow C}\ .$$

The resulting systems will be denoted by $\to$**G3mi**^{e}+Cut and $\to\bot$**G3i**^{e}+Cut. ⊠

When attempting to construct a stepwise local cut-elimination procedure for $\to\bot$**G3i**, which operates in the usual way of always treating a topmost occurrence of cut first, by either removing it completely (whenever this is possible if an axiom is involved) or by permuting it upwards over logical rules, or by splitting it into a few cuts of "simpler" kind (if the cut-formula is principal in both inferences of the immediate subderivations), one is for example led to the following cut-elimination reductions as in the lists A, B and C below. In the reduced derivations essential use is made of $\rightarrow\!\perp\!\mathbf{G3i}^e$'s structural rules weakening and contraction and therefore these reductions take ($\rightarrow\!\perp\!\mathbf{G3i}$+Cut)-derivations over to ($\rightarrow\!\perp\!\mathbf{G3i}^e$+Cut)-derivations.

Cut-elimination for $\rightarrow\!\mathbf{G3mi}$ can thereby be treated as a special case, in which fewer reduction-steps for transformations involving rule-applications immediately succeeding axioms have to be devised (due to the fact that $\rightarrow\!\mathbf{G3mi}$ has the same rules but less axioms than $\rightarrow\!\perp\!\mathbf{G3i}$). For the purpose of motivating the below reductions it will therefore only be spoken of $\rightarrow\!\perp\!\mathbf{G3i}$ and the just slightly more general situations occuring for derivations in this system.

A. Reductions by elimination or simplification of cuts with axioms:

(1) Either of the premises of Cut is an axiom Ax :

a. $$\dfrac{P,\Gamma_0 \Rightarrow P \qquad \overset{\mathcal{D}_1}{P,\Pi \Rightarrow C}}{P,\Gamma_0\Pi \Rightarrow C}\,\text{Cut} \qquad >_{red} \qquad \dfrac{\overset{\mathcal{D}_1}{P,\Pi \Rightarrow C}}{\overline{P,\Gamma_0\Pi \Rightarrow C}}\,\text{W}$$

b. $$\dfrac{\overset{\mathcal{D}_0}{\Gamma \Rightarrow P} \qquad P,\Pi \Rightarrow P}{\Gamma\Pi \Rightarrow P}\,\text{Cut} \qquad >_{red} \qquad \dfrac{\overset{\mathcal{D}_0}{\Gamma \Rightarrow P}}{\overline{\Gamma\Pi \Rightarrow P}}\,\text{W}$$

c. $$\dfrac{\overset{\mathcal{D}_0}{\Gamma \Rightarrow D} \qquad P,D,\Pi \Rightarrow P}{P,\Gamma\Pi \Rightarrow P}\,\text{Cut} \qquad >_{red} \qquad P,\Gamma\Pi \Rightarrow P$$

(2) Either of the premises of Cut is an axiom L$\perp$:

d. $$\dfrac{\perp,\Gamma \Rightarrow D \qquad \overset{\mathcal{D}_1}{D,\Pi \Rightarrow C}}{\perp,\Gamma\Pi \Rightarrow C}\,\text{Cut} \qquad >_{red} \qquad \perp,\Gamma\Pi \Rightarrow C$$

e. $$\dfrac{\overset{\mathcal{D}_0}{\Gamma \Rightarrow D} \qquad \perp,D,\Pi_0 \Rightarrow C}{\perp,\Gamma\Pi_0 \Rightarrow C}\,\text{Cut} \qquad >_{red} \qquad \perp,\Gamma\Pi_0 \Rightarrow C$$

f. $$\dfrac{\perp,\Gamma_0 \Rightarrow \perp \qquad \perp,\Pi_0 \Rightarrow C}{\perp,\Gamma_0\Pi_0 \Rightarrow C}\,\text{Cut} \qquad >_{red} \qquad \perp,\Gamma_0\Pi_0 \Rightarrow C$$

g.
$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A\to B,\Gamma_0\Rightarrow A\end{matrix}\qquad \begin{matrix}\mathcal{D}_{01}\\ B,\Gamma_0\Rightarrow\bot\end{matrix}}{A\to B,\Gamma_0\Rightarrow\bot}\,\text{L}{\to}\qquad \bot,\Pi\Rightarrow C}{A\to B,\Gamma_0\Pi\Rightarrow C}\,\text{Cut}\qquad >_{red}$$

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A\to B,\Gamma_0\Rightarrow A\end{matrix}}{A\to B,\Gamma_0\Pi\Rightarrow A}\,\text{W}\qquad \dfrac{\begin{matrix}\mathcal{D}_{01}\\ B,\Gamma_0\Rightarrow\bot\end{matrix}\qquad \bot,\Pi\Rightarrow C}{B,\Gamma_0\Pi\Rightarrow C}\,\text{Cut}}{A\to B,\Gamma_0\Pi\Rightarrow C}\,\text{R}{\to}$$

B. Reductions by permuting cuts upwards over logical rules:

a.
$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A\to B,\Gamma_0\Rightarrow A\end{matrix}\qquad \begin{matrix}\mathcal{D}_{01}\\ B,\Gamma_0\Rightarrow D\end{matrix}}{A\to B,\Gamma_0\Rightarrow D}\,\text{L}{\to}\qquad \begin{matrix}\mathcal{D}_{1}\\ D,\Pi\Rightarrow C\end{matrix}}{A\to B,\Gamma_0\Pi\Rightarrow C}\,\text{Cut}\qquad >_{red}$$

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A\to B,\Gamma_0\Rightarrow A\end{matrix}}{A\to B,\Gamma_0\Pi\Rightarrow A}\,\text{W}\qquad \dfrac{\begin{matrix}\mathcal{D}_{01}\\ B,\Gamma_0\Rightarrow D\end{matrix}\qquad \begin{matrix}\mathcal{D}_{1}\\ D,\Pi\Rightarrow C\end{matrix}}{B,\Gamma_0\Pi\Rightarrow C}\,\text{Cut}}{A\to B,\Gamma_0\Pi\Rightarrow C}\,\text{L}{\to}$$

b.
$$\dfrac{\begin{matrix}\mathcal{D}_{0}\\ \Gamma\Rightarrow D\end{matrix}\qquad \dfrac{\begin{matrix}\mathcal{D}_{10}\\ A,D,\Pi\Rightarrow B\end{matrix}}{D,\Pi\Rightarrow A\to B}\,\text{R}{\to}}{\Gamma\Pi\Rightarrow A\to B}\,\text{Cut}\qquad >_{red}$$

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{0}\\ \Gamma\Rightarrow D\end{matrix}\qquad \begin{matrix}\mathcal{D}_{10}\\ D,A,\Pi\Rightarrow B\end{matrix}}{A,\Gamma\Pi\Rightarrow B}\,\text{Cut}}{\Gamma\Pi\Rightarrow A\to B}\,\text{R}{\to}$$

c.
$$\dfrac{\begin{matrix}\mathcal{D}_{0}\\ \Gamma\Rightarrow D\end{matrix}\qquad \dfrac{\begin{matrix}\mathcal{D}_{10}\\ A\to B,D,\Pi_0\Rightarrow A\end{matrix}\qquad \begin{matrix}\mathcal{D}_{11}\\ B,D,\Pi_0\Rightarrow C\end{matrix}}{D,A\to B,\Pi_0\Rightarrow C}\,\text{L}{\to}}{A\to B,\Gamma\Pi_0\Rightarrow C}\,\text{Cut}\qquad >_{red}$$

$$\dfrac{\dfrac{\dfrac{\mathcal{D}_0}{\Gamma \Rightarrow D} \quad \dfrac{\mathcal{D}_{10}}{A \to B, D, \Pi_0 \Rightarrow A}}{A \to B, \Gamma\Pi_0 \Rightarrow A}\,\mathrm{Cut} \qquad \dfrac{\dfrac{\mathcal{D}_0}{\Gamma \Rightarrow D} \quad \dfrac{\mathcal{D}_{11}}{D, B, \Pi_0 \Rightarrow C}}{B, \Gamma\Pi_0 \Rightarrow C}\,\mathrm{Cut}}{A \to B, \Gamma\Pi_0 \Rightarrow C}\,\mathrm{L}\!\to$$

C. Fork Cut-Reduction[6]:

$$\dfrac{\dfrac{\dfrac{\mathcal{D}_{00}}{A, \Gamma \Rightarrow B}}{\Gamma \Rightarrow A \to B}\,\mathrm{R}\!\to \qquad \dfrac{\dfrac{\mathcal{D}_{10}}{A \to B, \Pi \Rightarrow A} \quad \dfrac{\mathcal{D}_{11}}{B, \Pi \Rightarrow C}}{A \to B, \Pi \Rightarrow C}\,\mathrm{L}\!\to}{\Gamma\Pi \Rightarrow C}\,\mathrm{Cut} \qquad >_{red}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\mathcal{D}_{00}}{A, \Gamma \Rightarrow B}}{\Gamma \Rightarrow A \to B}\,\mathrm{R}\!\to \quad \dfrac{\mathcal{D}_{10}}{A \to B, \Pi \Rightarrow A}}{\Gamma\Pi \Rightarrow A}\,\mathrm{Cut} \quad \dfrac{\mathcal{D}_{00}}{A, \Gamma \Rightarrow B}}{\Gamma^2\Pi \Rightarrow B}\,\mathrm{Cut} \quad \dfrac{\mathcal{D}_{11}}{B, \Pi \Rightarrow C}}{\dfrac{\Gamma^2\Pi^2 \Rightarrow C}{\Gamma\Pi \Rightarrow C}\,\mathrm{C}}\,\mathrm{Cut}$$

A process of cut-elimination based on the above reductions also requires the effective removal of the contraction and weakening rules, not present in $\boldsymbol{\to\bot}$**G3i**. For this purpose it will suffice to give a list of local transformation steps which allow to build a procedure for eliminating a single weakening or respectively a single contraction rule as the last rule application S from the bottom of a derivation $\mathcal{D}$, where S 's immediate subderivation is in fact a $\boldsymbol{\to\bot}$**G3i**-derivation, i.e. does not contain cut nor any of $\boldsymbol{\to\bot}$**G3i**^{e} 's structural rules.

Upwards permutation of weakening turns out to be straightforward, while that of contraction needs another of $\boldsymbol{\to\bot}$**G3i**^{e} 's structural rules, namely inversion of L$\to$.

D. Weakening Reductions:

(1) Involving an Axiom:

a. $$\dfrac{P, \Gamma \Rightarrow P}{D, P, \Gamma \Rightarrow P}\,\mathrm{W} \quad >_{red} \quad P, D, \Gamma \Rightarrow P$$

b. $$\dfrac{\bot, \Gamma \Rightarrow A}{D, \bot, \Gamma \Rightarrow A}\,\mathrm{W} \quad >_{red} \quad \bot, D, \Gamma \Rightarrow A$$

[6]The name “fork-reduction” follows Dragalin [Drag79].
[6]The name “fork-reduction” follows Dragalin [Drag79].

(2) Permuting weakening upwards over logical rules:

c.
$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ A, \Gamma \Rightarrow B\end{array}}{\Gamma \Rightarrow A \to B}\,\mathrm{R}{\to}}{D, \Gamma \Rightarrow A \to B}\,\mathrm{W} \quad >_{red} \quad \dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ A, \Gamma \Rightarrow B\end{array}}{A, D, \Gamma \Rightarrow B}\,\mathrm{W}}{D, \Gamma \Rightarrow A \to B}\,\mathrm{R}{\to}$$

d.
$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ A \to B, \Gamma \Rightarrow A\end{array} \qquad \begin{array}{c}\mathcal{D}_{01}\\ B, \Gamma \Rightarrow C\end{array}}{A \to B, \Gamma \Rightarrow C}\,\mathrm{L}{\to}}{D, A \to B, \Gamma \Rightarrow C}\,\mathrm{W} \quad >_{red}$$

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ A \to B, \Gamma \Rightarrow A\end{array}}{A \to B, D, \Gamma \Rightarrow A}\,\mathrm{W} \qquad \dfrac{\begin{array}{c}\mathcal{D}_{01}\\ B, \Gamma \Rightarrow C\end{array}}{B, D, \Gamma \Rightarrow C}\,\mathrm{W}}{A \to B, D, \Gamma \Rightarrow C}\,\mathrm{L}{\to}$$

E. Contraction Reductions:

(1) Involving an Axiom:

a. $$\dfrac{P, D, D, \Gamma_0 \Rightarrow P}{P, D, \Gamma_0 \Rightarrow P}\,\mathrm{C} \quad >_{red} \quad P, D, \Gamma_0 \Rightarrow P$$

b. $$\dfrac{P, P, \Gamma_0 \Rightarrow P}{P, \Gamma_0 \Rightarrow P}\,\mathrm{C} \quad >_{red} \quad P, \Gamma_0 \Rightarrow P$$

c. $$\dfrac{\bot, D, D, \Gamma_0 \Rightarrow A}{\bot, D, \Gamma_0 \Rightarrow A}\,\mathrm{C} \quad >_{red} \quad \bot, D, \Gamma_0 \Rightarrow A$$

d. $$\dfrac{\bot, \bot, \Gamma_0 \Rightarrow A}{\bot, \Gamma_0 \Rightarrow A}\,\mathrm{C} \quad >_{red} \quad \bot, \Gamma_0 \Rightarrow A$$

(2) Permuting contraction upwards over logical rules:

e.
$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ A, D, D, \Gamma_0 \Rightarrow B\end{array}}{D, D, \Gamma_0 \Rightarrow A \to B}\,\mathrm{R}{\to}}{D, \Gamma_0 \Rightarrow A \to B}\,\mathrm{C} \quad >_{red} \quad \dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ D, D, A, \Gamma_0 \Rightarrow B\end{array}}{A, D, \Gamma_0 \Rightarrow B}\,\mathrm{C}}{D, \Gamma_0 \Rightarrow A \to B}\,\mathrm{R}{\to}$$

f. Contracted formula is not principal in L$\to$:

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A\to B,D,D,\Gamma_0\Rightarrow A\end{matrix}\qquad\begin{matrix}\mathcal{D}_{01}\\ B,D,D,\Gamma_0\Rightarrow C\end{matrix}}{D,D,A\to B,\Gamma_0\Rightarrow C}\,\mathrm{L}{\to}}{D,A\to B,\Gamma_0\Rightarrow C}\,\mathrm{C}\qquad >_{red}$$

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ D,D,A\to B,\Gamma_0\Rightarrow A\end{matrix}}{A\to B,D,\Gamma_0\Rightarrow A}\,\mathrm{C}\qquad\dfrac{\begin{matrix}\mathcal{D}_{01}\\ D,D,B,\Gamma_0\Rightarrow C\end{matrix}}{B,D,\Gamma_0\Rightarrow C}\,\mathrm{C}}{A\to B,D,\Gamma_0\Rightarrow C}\,\mathrm{L}{\to}$$

g. Contracted formula is also principal in L$\to$:

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A\to B,A\to B,\Gamma\Rightarrow A\end{matrix}\qquad\begin{matrix}\mathcal{D}_{01}\\ B,A\to B,\Gamma\Rightarrow C\end{matrix}}{A\to B,A\to B,\Gamma\Rightarrow C}\,\mathrm{L}{\to}}{A\to B,\Gamma\Rightarrow C}\,\mathrm{C}\qquad >_{red}$$

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A\to B,A\to B,\Gamma\Rightarrow A\end{matrix}}{A\to B,\Gamma\Rightarrow A}\,\mathrm{C}\qquad\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{01}\\ A\to B,B,\Gamma\Rightarrow C\end{matrix}}{B,B,\Gamma\Rightarrow C}\,\mathrm{Inv}}{B,\Gamma\Rightarrow C}\,\mathrm{C}}{A\to B,\Gamma\Rightarrow C}\,\mathrm{L}{\to}$$

Now also reduction-steps for the systematic removal of inversion are needed to build a cut-elimination procedure for $\to\perp$**G3i**. It will again suffice to give such reduction-steps that permit the removal of a bottom-most application of inversion in a derivation that is otherwise a $\to\perp$**G3i**-derivation (i.e. one containing neither Cut nor one of $\to\perp\mathbf{G3i}^e$'s structural rules). No other structural rule (let alone a new one) is needed for upwards permutation of inversion.

F. Inversion Reductions:

(1) Involving an Axiom:

a. $\dfrac{P,A\to B,\Gamma\Rightarrow P}{B,P,\Gamma\Rightarrow P}\,\mathrm{Inv}\qquad >_{red}\qquad P,B,\Gamma\Rightarrow P$

b. $\dfrac{\perp,A\to B,\Gamma\Rightarrow C}{\perp,B,\Gamma\Rightarrow C}\,\mathrm{Inv}\qquad >_{red}\qquad \perp,B,\Gamma\Rightarrow C$

(2) Permuting inversion upwards over logical rules:

c.
$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ C, A \to B, \Gamma \Rightarrow D\end{matrix}}{A \to B, \Gamma \Rightarrow C \to D}\,\mathrm{R}\!\to}{B, \Gamma \Rightarrow C \to D}\,\mathrm{Inv} \quad >_{red} \quad \dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ C, A \to B, \Gamma \Rightarrow D\end{matrix}}{C, B, \Gamma \Rightarrow D}\,\mathrm{Inv}}{B, \Gamma \Rightarrow C \to D}\,\mathrm{R}\!\to$$

d. Active formula of inversion is not also principal formula of L→ :

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ C \to D, A \to B, \Gamma \Rightarrow C\end{matrix} \qquad \begin{matrix}\mathcal{D}_{01}\\ D, A \to B, \Gamma \Rightarrow E\end{matrix}}{A \to B, C \to D, \Gamma \Rightarrow E}\,\mathrm{L}\!\to}{B, C \to D, \Gamma \Rightarrow E}\,\mathrm{Inv} \quad >_{red}$$

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A \to B, C \to D, \Gamma \Rightarrow C\end{matrix}}{C \to D, B, \Gamma \Rightarrow C}\,\mathrm{Inv} \qquad \dfrac{\begin{matrix}\mathcal{D}_{01}\\ A \to B, D, \Gamma \Rightarrow E\end{matrix}}{D, B, \Gamma \Rightarrow E}\,\mathrm{Inv}}{C \to D, B, \Gamma \Rightarrow E}\,\mathrm{L}\!\to$$

e. Active formula of inversion is also principal formula of L→ :

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A \to B, \Gamma_0 \Rightarrow A\end{matrix} \qquad \begin{matrix}\mathcal{D}_{01}\\ B, \Gamma_0 \Rightarrow E\end{matrix}}{A \to B, \Gamma_0 \Rightarrow E}\,\mathrm{L}\!\to}{B, \Gamma_0 \Rightarrow E}\,\mathrm{Inv} \quad >_{red} \quad \begin{matrix}\mathcal{D}_{01}\\ B, \Gamma_0 \Rightarrow E\end{matrix}$$

It will be possible to show (by following and varying Dragalin's proof of strong cut-elimination for LJ and LK) that the reduction-steps from the above lists A–F can be applied to a given (→⊥**G3i**+Cut)-derivation $\mathcal{D}$ stepwise in an arbitrary order and at arbitrary places within $\mathcal{D}$ or within the meanwhile already transformed derivation (and where single reductions are only subject to the restrictions of their applicability as apparent from their description in A–F), such that for every sufficiently long sequence of reduction applications a cut-free form of $\mathcal{D}$ is reached. In short, *strong cut-elimination holds* for →⊥**G3i** *with respect to the set of reductions in the lists A–F.*

As also indicated above the reductions in A–F derive from analyzing closely the cut-elimination procedure implicit in the Cut-Elimination Theorem for the **G3**-systems in [TS96] (for the special case considered here of **G3[mi]**s' absurdity-containing implicative fragment →⊥**G3i**) and allow to rebuild and at the same time further specify this procedure as a stepwise process of locally applied transformations.

In order to consider a strong form of a cut-elimination theorem for →⊥**G3i** and for →**G3mi** w.r.t. above listed (types) of reduction rules now a clarifying definition is necessary.

Definition 2.2 (Reduction, normal derivations, strong cut-elimination). Let L be one of the calculi $\to$**G3mi**^{e} or $\to\perp$**G3i**^{e}.

1. Let $\mathcal{D}$, $\mathcal{D}'$ be L-derivations.

   $\mathcal{D}$ *L-1-reduces to* $\mathcal{D}'$ (in signs: $\mathcal{D} > \mathcal{D}'$) iff there exists a subderivation $\mathcal{D}_0$ of $\mathcal{D}$ such that $\mathcal{D}_0 >_{red} \mathcal{D}'_0$ by one of the reductions of the list A–F, if L is $\to\perp$**G3i**^{e}, [of A.a–c, B–C, D.a, D.c, D.d, E.a., E.b, E.e–g, F.a, F.c–e, if L is $\to$**G3mi**^{e}], and $\mathcal{D}'$ is the result of the replacement of $\mathcal{D}_0$ by $\mathcal{D}'_0$ in $\mathcal{D}$.

2. A derivation is said to be *L-normal* iff it does not L-1-reduce to any other derivation, i.e. if no $\mathcal{D}'$ exists such that $\mathcal{D} > \mathcal{D}'$.

3. *Strong cut-elimination holds* in L *w.r.t. reductions of the lists* A–F iff for all L-derivations $SN_{>}(\mathcal{D})$ holds (this means in the notation[7] of [TS96] that $\mathcal{D}$ is *strongly normalizing* w.r.t. $>$, i.e. that the reduction-tree of $\mathcal{D}$ w.r.t. L-1-reduction $>$ is finite). ⊠

The following theorem is the main result of this section.

Theorem 2.1 (Strong Cut-Elimination for $\to$G3mi and $\to\perp$G3i). *Strong cut-elimination holds for the calculi* $\to$**G3mi** *and* $\to\perp$**G3i** *with respect to reduction steps in the lists A–F.*

The *Proof* of this theorem is split into several lemmas and will be concluded later in this section on page 69.

Lemma 2.1. *A derivation $\mathcal{D}$ in one of the calculi* $\to$**G3mi**^{e} *or* $\to\perp$**G3i**^{e} *is normal iff it does neither contain weakenings, inversions, contractions nor cuts as rule-applications, i.e. iff $\mathcal{D}$ is a* $\to$**G3mi**-, *or respectively, a* $\to\perp$**G3i**-*derivation.*

Proof. It is clear that a derivation $\mathcal{D}$ which does not contain weakenings, inversions, contractions or cuts as rule-applications is normal (since all reductions $>_{red}$ of the types listed in A–F presuppose the existence of at least one weakening-, inversion-, contraction- or cut-rule in the derivation; therefore $\mathcal{D} > \mathcal{D}'$ for some derivation $\mathcal{D}'$ is not possible).

On the other hand any derivation $\mathcal{D}$ containing at least one rule application that is a weakening, an inversion, a contraction or a cut cannot be normal:

To see this choose a top-most such rule-application S and let $\mathcal{D}_0$ be the subderivation of $\mathcal{D}$ with S as its bottom-most rule application. Then all rule applications in immediate subderivations of $\mathcal{D}_0$ above S are applications of logical rules of L (i.e. of L$\to$- or R$\to$-rules).

[7](here slightly expanded with the additional used explicit sign $>$)

If S is a weakening it is easy to see from the list D of reductions that at least one of the reductions $>_{red}$ from this list is applicable to the immediate subderivation $\mathcal{D}_{00}$ of S in $\mathcal{D}_0$ (the reductions in this list have exactly been chosen such as to exhaust all possible cases). – The same can be checked for the lists F and E, if S is an inversion or a contraction. Hence $\mathcal{D}_0 >_{red} \mathcal{D}_0'$ for some $\mathcal{D}_0'$ in all these cases.

If S is an application of cut and $\mathcal{D}_{00}$ and $\mathcal{D}_{01}$ are its immediate subderivations in $\mathcal{D}$ then either (1) one of $\mathcal{D}_{00}$ or $\mathcal{D}_{01}$ is an axiom, or (2) both of $\mathcal{D}_{00}$ and $\mathcal{D}_{01}$ are not axioms and furthermore are of such form that the cut-formula of S is not principal in at least one of the two bottom-most rule-applications in $\mathcal{D}_{00}$ and respectively in $\mathcal{D}_{01}$ immediately above S, or (3) both of $\mathcal{D}_{00}$ and $\mathcal{D}_{01}$ are not axioms and the cut-formula of S is principal in both of the bottom-most rule-applications in $\mathcal{D}_{00}$ and $\mathcal{D}_{01}$. It can easily be checked that in case (1) one of the "axiomatic"-reductions $>_{red}$ from list A is applicable to $\mathcal{D}_0$, in case (2) one of the reductions from the list B of cut-permutation reductions (that deal with upwards permutation of cut over logical rules), and in case (3) a fork-reduction is applicable to $\mathcal{D}_0$. This means that then again $\mathcal{D}_0 >_{red} \mathcal{D}_0'$ holds.

Thus whatever rule application out of W, Inv, C or Cut the inference S happens to be, always $\mathcal{D}_0 >_{red} \mathcal{D}_0'$ holds. Hence $\mathcal{D} > \mathcal{D}'$ follows for that derivation $\mathcal{D}'$ which is the result of the replacement of $\mathcal{D}_0$ by $\mathcal{D}_0'$ in $\mathcal{D}$. Thus $\mathcal{D}$ is not normal. $\square$

Definition 2.3 (Reductive Derivations). Let L be either of the calculi $\rightarrow$**G3mi**^{e} or $\rightarrow\perp$**G3i**^{e}.

A derivation $\mathcal{D}$ is called *L-reductive*, iff it has a finite reduction-tree with repect to the L-1-reduction $>$ (i.e. iff $\mathcal{D}$ is strongly normalizing with respect to $>$, which we abbreviate symbolically to $SN_>(\mathcal{D})$). The *reductive complexity* $red(\mathcal{D})$ of a L-reductive derivation $\mathcal{D}$ is the size of the reduction-tree of $\mathcal{D}$ with respect to $>$. $\boxtimes$

Some simple properties of reductive derivations are stated in the following two lemmas.

Lemma 2.2. *Let $\mathcal{D}$ be a derivation that terminates with a basic logical rule S, i.e. $\mathcal{D}$ is of the form*

$$\frac{\mathcal{D}_0 \qquad (\mathcal{D}_1)}{\Gamma \Rightarrow C}\, S$$

and suppose $\mathcal{D} > \mathcal{D}'$.

Then $\mathcal{D}'$ terminates with the same rule and is of the form

$$\frac{\mathcal{D}_0' \qquad (\mathcal{D}_1')}{\Gamma \Rightarrow C}\, S$$

where for exactly one of the immediate subderivations $\mathcal{D}_0, \mathcal{D}_1$ of S it holds that $\mathcal{D}_i > \mathcal{D}_i'$, while for the other one (if S is an application of the two-premise rule $L\rightarrow$ at all) $\mathcal{D}_i \equiv \mathcal{D}_i'$ $(i = 0, 1)$ is true.

Proof. This clearly follows since no $>_{red}$-reduction of the types listed in A–F is applicable to $\mathcal{D}$ itself (since in all these reductions the bottom-most rule has to be different from a logical rule) and such reductions can therefore only be applicable to proper subderivations of $\mathcal{D}$, thus to subderivations of $\mathcal{D}_0$ or $\mathcal{D}_1$. $\square$

Lemma 2.3. *Suppose a derivation $\mathcal{D}$ terminates with a logical rule S and has immediate subderivation(s) $\mathcal{D}_0$ ($\mathcal{D}_0$ or $\mathcal{D}_1$), where $\mathcal{D}_0$ is ($\mathcal{D}_0$ and $\mathcal{D}_1$ are) reductive. Then $\mathcal{D}$ is also reductive.*

Proof. By induction on the sum of the sizes of the reduction-trees of $\mathcal{D}_0$ and of $\mathcal{D}_1$ with respect to $>$, i.e. on the sum $red(\mathcal{D}_0) + red(\mathcal{D}_1)$ using Lemma 2.2 as well as $red(\mathcal{D}) > red(\mathcal{D}')$ for all $\mathcal{D}'$ such that $\mathcal{D} > \mathcal{D}'$ (which is obvious from the definition of the reductive complexity) in the induction step. $\square$

Definition 2.4 (Inductive derivations, inductive complexity).
Let L be either of the calculi $\rightarrow$**G3mi**^{e} or $\rightarrow\perp$**G3i**^{e}.

The class of L-*inductive* derivations in L is given by an inductive definition with clauses (1), (2) and (3) below[8]. At the same time a derivation-associated number $red(\mathcal{D})$, the *inductive complexity* of $\mathcal{D}$ is defined in parallel.

(1) Every L-derivation $\mathcal{D}'$ consisting only of an axiom of L is L-*inductive*. The inductive complexity of $\mathcal{D}'$ is then defined by $ind(\mathcal{D}') := 1$.

(2) If $\mathcal{D}'$ terminates with a logical rule S of L and has immediate subdeduction(s) $\mathcal{D}'_0$ ($\mathcal{D}'_0$ and $\mathcal{D}'_1$), then $\mathcal{D}'$ is L-*inductive* iff $\mathcal{D}'_0$ is ($\mathcal{D}'_0$ and $\mathcal{D}'_1$ are) L-*inductive*. The inductive complexity of $\mathcal{D}'$ is then defined by $ind(\mathcal{D}') := ind(\mathcal{D}'_0) + 1$ (respectively by $ind(\mathcal{D}') := ind(\mathcal{D}'_0) + ind(\mathcal{D}'_1) + 1$).

(3) If $\mathcal{D}'$ terminates with an application of W, C, Inv or Cut and $\mathcal{D}'_1, \dots, \mathcal{D}'_n$ is a complete list of L-derivations such that $\mathcal{D}' > \mathcal{D}'_i$ (for $i = 1, \dots, n$), then $\mathcal{D}'$ is L-*inductive* iff all $\mathcal{D}'_1, \dots, \mathcal{D}'_n$ are L-*inductive*. The inductive complexity of $\mathcal{D}'$ in this situation is defined by $ind(\mathcal{D}') := 1 + \sum_{i=1}^{n} ind(\mathcal{D}'_i)$. $\boxtimes$

Since the further proof of Theorem 2.1 is in essence largely the same for $\rightarrow$**G3mi** and for $\rightarrow\perp$**G3i**, the explicit reference to either of this systems will be dropped in the following; this will also apply to notations like "L-reductive" (L meaning one of these calculi) and it will then be tacitly assumed that all statements given will be valid in each of these two cases respectively and accordingly. In cases and at places where differences occur this will be stated clearly.

[8] Dragalin prefers to state a very similar definition more exactly than above in the form of a formal calculus *Ind* with inductive L-derivations as its "theorems"; it was hoped here that this presentation of the definitions is clearer to understand.

Lemma 2.4. *Every inductive derivation is reductive.*

Proof. By induction on the size $ind(\mathcal{D})$ of a "proof" of $\mathcal{D}$ to be inductive.

If $ind(\mathcal{D}) = 1$, then $\mathcal{D}$ is an axiom, which is a normal and hence also a reductive derivation.

If $ind(\mathcal{D}) > 1$ and $\mathcal{D}$ terminates with a logical rule S, then the inductive complexities $ind(\mathcal{D}_0)$ and $ind(\mathcal{D}_1)$ of the immediate subderivations of S are by definition smaller than $ind(\mathcal{D})$ and $\mathcal{D}_0$ and $\mathcal{D}_1$ are then inductive by definition as well. By the induction hypothesis it follows that $\mathcal{D}_0$ and $\mathcal{D}_1$ are reductive; Lemma 2.3 now implies that $\mathcal{D}$ is reductive.

If $ind(\mathcal{D}) > 1$ and $\mathcal{D}$ terminates with a structural rule, an inversion or a cut and $\mathcal{D}_1, \dots, \mathcal{D}_n$ is a complete list of all derivations $\mathcal{D}'$ such that $\mathcal{D} > \mathcal{D}'$, then by Definition 2.4, (3), all $\mathcal{D}_i$ $(i = 1, \dots, n)$ are inductive and $ind(\mathcal{D}_1) < ind(\mathcal{D})$ $(i = 1, \dots, n)$. From this by induction hypothesis it follows that all $\mathcal{D}_i$ are reductive, which in turn implies that $\mathcal{D}$ is reductive (namely by clause (2) of Def. 2.3). □

Lemma 2.5. *Suppose that $\mathcal{D}$ terminates with a logical rule and has $\mathcal{D}_0$ and $\mathcal{D}_1$ as its immediate subderivations. Then $\mathcal{D}$ is inductive iff $\mathcal{D}_0$ and $\mathcal{D}_1$ are inductive; moreover it holds that $ind(\mathcal{D}_0), ind(\mathcal{D}_1) < ind(\mathcal{D})$.*

Proof. This is an immediate consequence of clause 2 in the definition of inductiveness in Definition 2.4. □

Lemma 2.6. *If $\mathcal{D}$ is inductive and $\mathcal{D} > \mathcal{D}'$, then $\mathcal{D}'$ is also inductive and $ind(\mathcal{D}') < ind(\mathcal{D})$.*

Proof. By induction on the depth of the derivation $\mathcal{D}$.

If $\mathcal{D}$ is an axiom, then $\mathcal{D} > \mathcal{D}'$ is not possible, hence there is nothing to show.

If $\mathcal{D}$ is inductive and terminates with a logical rule S, then it is of the form

$$\frac{\mathcal{D}_0 \qquad (\mathcal{D}_1)}{\Gamma \Rightarrow C}\,\mathrm{S} \quad .$$

By Def. 2.4, (2), $\mathcal{D}_0, \mathcal{D}_1$ are inductive. $\mathcal{D} > \mathcal{D}'$ implies that either $\mathcal{D}_0 > \mathcal{D}_0'$ or $\mathcal{D}_1 > \mathcal{D}_1'$. Suppose for once that $\mathcal{D}_1 > \mathcal{D}_1'$ holds. Since then by definition of the inductive complexity $ind(\mathcal{D}_1) < ind(\mathcal{D})$ holds, now by the induction hypothesis $\mathcal{D}_1'$ is implied to be inductive as well, and also $ind(\mathcal{D}_1) < ind(\mathcal{D}_1')$ follows. Then $\mathcal{D}'$ is of the form

$$\frac{\mathcal{D}_0 \qquad \mathcal{D}_1'}{\Gamma \Rightarrow C}\,\mathrm{S} \quad ,$$

and is again inductive by Def. 2.4, clause (2); moreover then

$$ind(\mathcal{D}') = ind(\mathcal{D}_0) + ind(\mathcal{D}_1') + 1 < ind(\mathcal{D}_0) + ind(\mathcal{D}_1) = 1 = ind(\mathcal{D})$$

holds. – If $\mathcal{D}_0 > \mathcal{D}'_0$, the argument is similar.

If $\mathcal{D}$ is inductive and terminates with a structural rule, an inversion or a cut, then the statement of the lemma is directly implied by clause (3) of Def. 2.4. □

Lemma 2.7. *A derivation $\mathcal{D}$ is inductive iff every $\mathcal{D}'$ such that $\mathcal{D} > \mathcal{D}'$ is inductive.*

Proof. "⇒": Is the main statement of Lemma 2.6.

"⇐": By induction on the depth of a derivation $\mathcal{D}$:

If $\mathcal{D}$ is an axiom, then $\mathcal{D}$ is inductive by clause (1) of Def. 2.4.

If $\mathcal{D}$ terminates with a logical rule S, then it is of the form $\dfrac{\mathcal{D}_0 \qquad (\mathcal{D}_1)}{\Gamma \Rightarrow C}\,\mathrm{S}$. Suppose now that $\mathcal{D}'$ is inductive for all $\mathcal{D}'$ such that $\mathcal{D} > \mathcal{D}'$.

Then $\mathcal{D}_0$ is inductive: Suppose $\mathcal{D}_0 > \mathcal{D}'_0$. Then also $\mathcal{D} > \mathcal{D}'$ with $\mathcal{D}'$ being the derivation $\dfrac{\mathcal{D}'_0 \qquad (\mathcal{D}_1)}{\Gamma \Rightarrow C}\,\mathrm{S}$ and by assumption $\mathcal{D}'$ is inductive as well. Thus (since $\mathcal{D}'_0$ has here been arbitrary with $\mathcal{D}_0 > \mathcal{D}'_0$) it has been shown that $\mathcal{D}'_0$ inductive holds for arbitrary $\mathcal{D}'_0$ such that $\mathcal{D}_0 > \mathcal{D}'_0$. Since the depth of $\mathcal{D}_0$ is smaller than that of $\mathcal{D}$, the induction hypothesis is applicable and gives that $\mathcal{D}_0$ is inductive. – In a completely analogous way it can be shown that $\mathcal{D}_1$ is also inductive.

Now that $\mathcal{D}_0$ and $\mathcal{D}_1$ have been recognized as being inductive, it follows that $\mathcal{D}$ is inductive as well (because $\mathcal{D}$ terminates with a logical rule, cf. Def. 2.4, (2)).

If $\mathcal{D}$ terminates with a structural rule, an inversion or a cut, then the statement "$\mathcal{D}$ is inductive" precisely amounts to the assumption of "⇐"; hence there is nothing else to show in this case.

□

Lemma 2.8. *Every normal derivation is inductive.*

Proof. A normal derivation $\mathcal{D}$ does not contain weakenings, inversions, contractions or cuts by Lemma 2.1. Then inductiveness of $\mathcal{D}$ follows by an obvious induction using only the clauses (1) and (2) of Definition 2.4. □

Lemma 2.9. *Every subderivation of an inductive derivation is inductive.*

Proof. It suffices to show that immediate subderivations of the bottom-most rule application S in an inductive derivation $\mathcal{D}$ are inductive themselves (the lemma then follows by stepwise induction). This will be shown by induction on $ind(\mathcal{D})$.

If $ind(\mathcal{D}) = 1$ then $\mathcal{D}$ is an axiom and has only itself as a subderivation.

If $ind(\mathcal{D}) > 1$ for an inductive derivation $\mathcal{D}$ that terminates with a logical rule that has $\mathcal{D}_0$ as well as possibly also $\mathcal{D}_1$ as immediate subderivations, then the inductiveness of $\mathcal{D}_0$ and $\mathcal{D}_1$ follows from clause (2) in Def. 2.4 of $\mathcal{D}$ to be inductive.

If $ind(\mathcal{D}) > 1$ for an inductive derivation $\mathcal{D}$ that terminates with a rule S that is a structural rule, an inversion or a cut, then $\mathcal{D}$ has the form

$$\frac{\mathcal{D}_0 \qquad (\mathcal{D}_1)}{\Gamma \Rightarrow C}\,\mathrm{S} \quad .$$

To prove that $\mathcal{D}_0$ and $\mathcal{D}_1$ are inductive it suffices by Lemma 2.7 to show that all derivations $\mathcal{D}_i'$ such that $\mathcal{D}_i > \mathcal{D}_i'$ are inductive $(i = 0, 1)$. Both cases are dealt with analogously, so we look for example at $\mathcal{D}_1$. Suppose $\mathcal{D}_1 > \mathcal{D}_1'$. Then $\mathcal{D} > \mathcal{D}'$ holds, where $\mathcal{D}'$ is the derivation

$$\frac{\mathcal{D}_0 \qquad (\mathcal{D}_1')}{\Gamma \Rightarrow C}\,\mathrm{S} \quad .$$

By Lemma 2.6 it can be seen that $ind(\mathcal{D}') < ind(\mathcal{D})$. From the induction hypothesis applied to $\mathcal{D}'$ it then follows that $\mathcal{D}_1'$ is inductive. Thus—$\mathcal{D}_1'$ was arbitrary such that $\mathcal{D}_1 > \mathcal{D}_1'$, and in view of Lemma 2.7 as mentioned above—$\mathcal{D}_1$ is then recognized as being inductive. – Similarly $\mathcal{D}_0$ can be seen to be inductive. □

The following 4 lemmas are the "heart" of the proof of Theorem 2.1 and together will state that the inductiveness of derivations is preserved under applications of weakening, inversion, contraction or cut, which take place at the bottom of inductive derivations. That is, a derivation terminating with one of these 4 rules, that has an inductive subderivation (in the case of Cut: ... that has inductive subderivations), is again inductive. (In the case of the logical rules R→ and L→ this is part of Definition 2.4.)

Lemma 2.10. *Every derivation $\mathcal{D}$ obtained by adding an application of weakening at the bottom of an inductive derivation $\mathcal{D}_0$ is inductive.*

Proof. It has to be shown that every derivation $\mathcal{D}$ of the form

$$\frac{\begin{array}{c}\mathcal{D}_0\\ \Gamma \Rightarrow C\end{array}}{A, \Gamma \Rightarrow C}\,\mathrm{W} \quad , \tag{2.1}$$

where $\mathcal{D}_0$ is inductive, is inductive.

In view of Lemma 2.7 it suffices to show for any such $\mathcal{D}$ that[9]

$$\forall \mathcal{D}' \,\big(\, \mathcal{D} > \mathcal{D}' \Rightarrow \mathcal{D}' \text{ is inductive } \big). \tag{2.2}$$

This will be shown by induction on $ind(\mathcal{D}_0)$.

[9]The notation in this and similar statements to come is to be understood as part of an informal meta-language dealing with properties of and relations between derivations.

Let $\mathcal{D}$ be of the form (2.1), with $\mathcal{D}_0$ inductive. Assume the induction hypothesis for $\mathcal{D}$.

Let $\mathcal{D}'$ be arbitrary such that $\mathcal{D} > \mathcal{D}'$. To prove (2.2) it needs to be shown that $\mathcal{D}'$ is inductive.

Case 1: $\mathcal{D} > \mathcal{D}'$ is due to a reduction $>_{red}$ that takes place within $\mathcal{D}_0$, i.e. which does not involve nor change the weakening at the bottom of $\mathcal{D}$:

Then $\mathcal{D}'$ is of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_0' \\ \Gamma \Rightarrow C\end{array}}{A, \Gamma \Rightarrow C}\,\mathrm{W}$$

with $\mathcal{D}_0 > \mathcal{D}_0'$. By Lemma 2.6 it follows from $\mathcal{D}_0 > \mathcal{D}_0'$ that $ind(\mathcal{D}_0') < ind(\mathcal{D}_0)$ holds; hence the induction hypothesis is applicable to $\mathcal{D}'$, which gives that $\mathcal{D}'$ is inductive.

Case 2: $\mathcal{D} > \mathcal{D}'$ is true because of $\mathcal{D} >_{red} \mathcal{D}'$, i.e. the reduction step $\mathcal{D} > \mathcal{D}'$ takes place at the last rule application of $\mathcal{D}$ and thus involves the bottom-most weakening in $\mathcal{D}$:

The reduction $\mathcal{D} >_{red} \mathcal{D}'$ must then be a weakening reduction from the list D of types of such reductions, since none of the other reductions of the lists A–F has a weakening at the bottom of the derivation to be reduced.

If $ind(\mathcal{D}_0) = 1$, then $\mathcal{D}_0$ consists only of an axiom and hence the reduction $\mathcal{D} >_{red} \mathcal{D}'$ must be one of the types D.a or D.b . But then $\mathcal{D}'$ is again an axiom, which is a normal and hence an inductive derivation; thus $\mathcal{D}'$ is inductive in these cases.

If $ind(\mathcal{D}_0) = 1$ and $\mathcal{D} >_{red} \mathcal{D}'$ holds because of and via a reduction of one of the types D.c or D.d, then the inductiveness of $\mathcal{D}'$ follows easily from the induction hypothesis: For example in the case of a reduction of type D.d the derivation $\mathcal{D}$ has the form

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00} \\ B \to D, \Gamma_0 \Rightarrow B\end{array} \qquad \begin{array}{c}\mathcal{D}_{01} \\ D, \Gamma_0 \Rightarrow C\end{array}}{B \to D, \Gamma_0 \Rightarrow C}\,\mathrm{L}\!\to}{A, B \to D, \Gamma_0 \Rightarrow C}\,\mathrm{W} \quad . \tag{2.3}$$

Since $\mathcal{D}_0$ (here the subderivation of $\mathcal{D}$ terminating with the application of L$\to$ above W) is inductive and ends with a logical rule, both $\mathcal{D}_{00}$ and $\mathcal{D}_{01}$ are inductive and $ind(\mathcal{D}_{00}), ind(\mathcal{D}_{01}) < ind(\mathcal{D}_0)$ by Lemma 2.5. Then by the induction hypothesis the derivations

$$\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ B \to D, \Gamma_0 \Rightarrow\ B\end{array}}{A, B \to D, \Gamma_0 \Rightarrow\ B}\,\mathrm{W} \qquad \text{as well as} \qquad \dfrac{\begin{array}{c}\mathcal{D}_{01}\\ D, \Gamma_0 \Rightarrow\ C\end{array}}{A, D, \Gamma_0 \Rightarrow\ C}\,\mathrm{W}$$

are inductive; from this by Definition 2.4, (2), it now follows that the derivation $\mathcal{D}'$, which—since $\mathcal{D} > \mathcal{D}'$ via a reduction of type D.d and $\mathcal{D}$ is of form (2.3)—must be of the form

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ B \to D, \Gamma_0 \Rightarrow\ B\end{array}}{B \to D, A, \Gamma_0 \Rightarrow\ B}\,\mathrm{W} \qquad \dfrac{\begin{array}{c}\mathcal{D}_{01}\\ D, \Gamma_0 \Rightarrow\ C\end{array}}{D, A, \Gamma_0 \Rightarrow\ C}\,\mathrm{W}}{B \to D, A, \Gamma_0 \Rightarrow\ C}\,\mathrm{L}{\to} \quad ,$$

is inductive. – The proof for the remaining case, in which $\mathcal{D} >_{red} \mathcal{D}'$ is due to a reduction of type D.c, is easier still.

Since for arbitrary $\mathcal{D}'$ such that $\mathcal{D} > \mathcal{D}'$ the inductiveness of $\mathcal{D}'$ has now been shown, (2.2) has been proved. As already said, from this the lemma follows. ☐

Lemma 2.11. *Every derivation $\mathcal{D}$ terminating with an application of inversion to the end-sequent of an inductive derivation $\mathcal{D}_0$ is itself inductive.*

Proof. It has to be established, that every derivation $\mathcal{D}$ of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ A \to B, \Gamma \Rightarrow\ C\end{array}}{B, \Gamma \Rightarrow\ C}\,\mathrm{Inv} \quad , \tag{2.4}$$

where $\mathcal{D}_0$ is inductive, is itself inductive. As before, on the basis of Lemma 2.7 only a proof of

$$\forall \mathcal{D}' \ \big(\mathcal{D} > \mathcal{D}' \Rightarrow \mathcal{D}' \text{ is inductive } \big)$$

for all $\mathcal{D}$ as above needs to be given.

This again can be shown by induction on $ind(\mathcal{D}_0)$. The proof proceeds analogously to that of (2.2) in Lemma 2.10; except that in case 2, when $\mathcal{D} > \mathcal{D}'$ is due to $\mathcal{D} >_{red} \mathcal{D}'$, a reduction of type F.e and of the form

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ A \to B, \Gamma_0 \Rightarrow\ A\end{array} \qquad \begin{array}{c}\mathcal{D}_{01}\\ B, \Gamma_0 \Rightarrow\ E\end{array}}{A \to B, \Gamma_0 \Rightarrow\ E}\,\mathrm{L}{\to}}{B, \Gamma_0 \Rightarrow\ E}\,\mathrm{Inv} \qquad >_{red} \qquad \begin{array}{c}\mathcal{D}_{01}\\ B, \Gamma_0 \Rightarrow\ E\end{array} \tag{2.5}$$

has to be considered additionally, for it has no counterpart in the list D of types of weakening reductions. The inductiveness of $\mathcal{D}'$ here follows directly: Firstly, $\mathcal{D}' \equiv \mathcal{D}_{01}$ holds (as can be seen from (2.4), (2.5)), and—on the other hand—$\mathcal{D}_{01}$ is inductive as a consequence of Lemma 2.5, because it is a subderivation of the inductive derivation $\mathcal{D}_0$ (the immediate subderivation of the bottom-most inversion in $\mathcal{D}$), which ends with a logical rule (namely with L→). □

Lemma 2.12. *Every derivation $\mathcal{D}$ obtained from an inductive derivation $\mathcal{D}_0$ by a single succeeding application of contraction is inductive.*

Proof. It has to be shown, that every derivation $\mathcal{D}$ of the form

$$\dfrac{\begin{matrix}\mathcal{D}_0\\ A, A, \Gamma \Rightarrow\ C\end{matrix}}{A, \Gamma \Rightarrow\ C}\,\mathrm{C} \quad , \tag{2.6}$$

where $\mathcal{D}_0$ is inductive, is inductive. By Lemma 2.7 again only

$$\forall\, \mathcal{D}' \;\big(\, \mathcal{D} > \mathcal{D}' \Rightarrow \mathcal{D}' \text{ is inductive } \big) \tag{2.7}$$

has to be proved for all $\mathcal{D}$ considered here.

This will be shown by induction on $\big(|A|+1, \mathit{ind}(\mathcal{D}_0)\big)$ with respect to the lexicographic order on $\mathbb{N} \times \mathbb{N}$, that is to say by induction on the depth of the formula contracted at the bottom of $\mathcal{D}$ together with a subinduction on the inductive complexity $\mathit{ind}(\mathcal{D}_0)$ of $\mathcal{D}_0$.

Let $\mathcal{D}$ be of the form (2.6), with $\mathcal{D}_0$ inductive. Assume the induction and subinduction hypothesis for $\mathcal{D}$.

Let $\mathcal{D}'$ be arbitrary such that $\mathcal{D} > \mathcal{D}'$. The aim now is to recognize $\mathcal{D}'$ as an inductive derivation.

Case 1: $\mathcal{D} > \mathcal{D}'$ is due to a reduction $>_{red}$, that takes place within $\mathcal{D}_0$, i.e. one, which does not involve the contraction at the bottom of $\mathcal{D}$.

Then $\mathcal{D}'$ is of the form

$$\dfrac{\begin{matrix}\mathcal{D}_0'\\ A, A, \Gamma \Rightarrow\ C\end{matrix}}{A, \Gamma \Rightarrow\ C}\,\mathrm{C}$$

with $\mathcal{D}_0'$ such that $\mathcal{D}_0 > \mathcal{D}_0'$. Since $\mathit{ind}(\mathcal{D}_0') < \mathit{ind}(\mathcal{D}_0)$ (by Lemma 2.6) and the contraction at the bottom of $\mathcal{D}$ was unchanged by the reduction, the subinduction hypothesis is applicable to $\mathcal{D}'$ and implies that $\mathcal{D}'$ is inductive.

Case 2: $\mathcal{D} > \mathcal{D}'$ is due to $\mathcal{D} >_{red} \mathcal{D}'$, i.e. the reduction step $\mathcal{D} > \mathcal{D}'$ takes place at the bottom of $\mathcal{D}$.

The reduction must then be of the type of a contraction reduction from the list E, since all *other* $>_{red}$-reductions in the lists A–F do not apply to derivations ending with a contraction rule.

If $\mathcal{D}_0$ consists only of an axiom, then the reduction $\mathcal{D} >_{red} \mathcal{D}'$ must be of one of the types E.a–d, which all reduce to axioms, hence to normal and as such also inductive derivations. Thus $\mathcal{D}'$ is inductive in these cases.

If $\mathcal{D}_0$ is not an axiom and $\mathcal{D} >_{red} \mathcal{D}'$ takes place via a reduction of one of the types E.e or E.f, then the inductiveness of $\mathcal{D}'$ easily follows from the subinduction hypothesis (and in part analogously to the more special case treated explicity below), noticing that the formula(s) contracted one step above the bottom of $\mathcal{D}'$ is (are) again A and that therefore the (syntactical) depth of the contracted formula(s) in the newly introduced contractions has (have) not increased (which is a necessary condition for applying the subinduction hypothesis).

If $\mathcal{D}_0$ is not an axiom and $\mathcal{D} >_{red} \mathcal{D}'$ is a reduction of the type E.g, then $\mathcal{D} >_{red} \mathcal{D}'$ has the form

$$\dfrac{\dfrac{\dfrac{\overset{\mathcal{D}_{00}}{B \to D, B \to D, \Gamma \Rightarrow B} \qquad \overset{\mathcal{D}_{01}}{D, B \to D, \Gamma \Rightarrow C}}{B \to D, B \to D, \Gamma \Rightarrow C}\,\mathrm{L}\!\to}{B \to D, \Gamma \Rightarrow C}\,\mathrm{C} \qquad >_{red} \tag{2.8}$$

$$\dfrac{\dfrac{\overset{\mathcal{D}_{00}}{B \to D, B \to D, \Gamma \Rightarrow B}}{B \to D, \Gamma \Rightarrow B}\,\mathrm{C} \qquad \dfrac{\dfrac{\overset{\mathcal{D}_{01}}{B \to D, D, \Gamma \Rightarrow C}}{D, D, \Gamma \Rightarrow C}\,\mathrm{Inv}}{D, \Gamma \Rightarrow C}\,\mathrm{C}}{B \to D, \Gamma \Rightarrow C}\,\mathrm{L}\!\to \quad ,$$

where—to make the correspondence to (2.6) clear—it holds that $A \equiv B \to D$ and that $\mathcal{D}_0$ is the immediate subderivation of the bottom-most contraction in the reduction to be reduced. Here by Lemma 2.5 $\mathcal{D}_{00}$ and $\mathcal{D}_{01}$ are inductive (as immediate subderivations of the derivation $\mathcal{D}_0$ ending with the logical rule L$\to$) and with $ind(\mathcal{D}_{00}), ind(\mathcal{D}_{01}) < ind(\mathcal{D}_0)$.

Then by the subinduction hypothesis the derivation $\tilde{\mathcal{D}_0}$ of the form

$$\dfrac{\overset{\mathcal{D}_{00}}{B \to D, B \to D, \Gamma \Rightarrow B}}{B \to D, \Gamma \Rightarrow B}\,\mathrm{C}$$

is inductive (as a precondition for using the subinduction hypothesis the contraction formula $B \to D$ herein is the same as the original one A at the bottom of $\mathcal{D}$).

By Lemma 2.11 now the derivation $\tilde{\mathcal{D}_{01}}$

$$\dfrac{\begin{matrix}\mathcal{D}_{01}\\ B \to D, D, \Gamma \Rightarrow\ C\end{matrix}}{D, D, \Gamma \Rightarrow\ C}\,\text{Inv}$$

is also inductive (because $\mathcal{D}_{01}$ is inductive); since $|D| = 1 < |B \to D| + 1 = |A| + 1$, the induction hypothesis is applicable for a contraction at the bottom of $\tilde{\mathcal{D}_{01}}$ and shows that the derivation $\tilde{\mathcal{D}_1}$ of the form

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{01}\\ B \to D, D, \Gamma \Rightarrow\ C\end{matrix}}{D, D, \Gamma \Rightarrow\ C}\,\text{Inv}}{D, \Gamma \Rightarrow\ C}\,\text{C}$$

is inductive as well.

Then clearly the inductiveness of $\mathcal{D}'$, the derivation at the right side of the reduction in (2.8), follows from Def. 2.4, (2), since it is of the form

$$\dfrac{\tilde{\mathcal{D}_0} \qquad \tilde{\mathcal{D}_1}}{B \to D, \Gamma \Rightarrow\ C}\,\text{L}{\to}$$

and $\tilde{\mathcal{D}_0}$, $\tilde{\mathcal{D}_1}$ have already been recognized as inductive derivations.

Since for arbitrary $\mathcal{D}'$ with $\mathcal{D} > \mathcal{D}'$ it has been shown that $\mathcal{D}'$ is inductive, (2.7) has been proved. This completed the proof of the lemma. $\square$

Lemma 2.13. *Every derivation $\mathcal{D}$, which ends with an application of cut, that has inductive immediate subderivations in $\mathcal{D}$, is inductive itself.*

Proof. The lemma states that every derivation $\mathcal{D}$ of the form

$$\dfrac{\begin{matrix}\mathcal{D}_0\\ \Gamma \Rightarrow D\end{matrix} \qquad \begin{matrix}\mathcal{D}_1\\ D, \Pi \Rightarrow C\end{matrix}}{\Gamma\Pi \Rightarrow C}\,\text{Cut}\quad, \tag{2.9}$$

where $\mathcal{D}_0$ and $\mathcal{D}_1$ are inductive derivations, is inductive.

The proof will use induction on $\big(|D|\,, ind(\mathcal{D}_0) + ind(\mathcal{D}_1)\big)$ with respect to the lexicographic order on $\mathbb{N}_0 \times \mathbb{N}$; phrased differently, this says that the proof will proceed by induction on the (syntactical) depth $|D|$ of the cut-formula together with a subinduction on the sum of the inductive complexities of the immediate subderivations $\mathcal{D}_0$ and $\mathcal{D}_1$ of the cut.

Let $\mathcal{D}$ be a derivation of the form (2.9) with $\mathcal{D}_0$, $\mathcal{D}_1$ inductive subderivations. In view of Lemma 2.7 it suffices to show for arbitrary given $\mathcal{D}$, that is of the form (2.9) and for which the induction hypothesis is assumed to be true, that

$$\forall \mathcal{D}' \ \big(\mathcal{D} > \mathcal{D}' \Rightarrow \mathcal{D}' \text{ is inductive } \big) \tag{2.10}$$

holds.

Thus we assume the induction hypothesis for $\mathcal{D}$ and we let $\mathcal{D}'$ be such that $\mathcal{D} > \mathcal{D}'$. It will be shown that $\mathcal{D}'$ is inductive.

Case 1: The reduction $>_{red}$ underlying $\mathcal{D} > \mathcal{D}'$ takes place in either $\mathcal{D}_0$ or $\mathcal{D}_1$.

Suppose for example $\mathcal{D}_1 > \mathcal{D}_1'$ and $\mathcal{D}'$ is the result

$$\dfrac{\begin{array}{c}\mathcal{D}_0\\ \Gamma \Rightarrow D\end{array} \qquad \begin{array}{c}\mathcal{D}_1'\\ D, \Pi \Rightarrow C\end{array}}{\Gamma, \Pi \Rightarrow C}\,\text{Cut}$$

of replacing $\mathcal{D}_1$ in $\mathcal{D}$ by $\mathcal{D}_1'$. Then by Lemma 2.6 $ind(\mathcal{D}_1') < ind(\mathcal{D}_1)$; hence $ind(\mathcal{D}_0) + ind(\mathcal{D}_1') < ind(\mathcal{D}_0) + ind(\mathcal{D}_1)$. By appeal to the subinduction hypothesis it then follows that $\mathcal{D}'$ is inductive.

The same argument can be carried out analogously, if $\mathcal{D}'$ is the result of a $>_{red}$-reduction that takes place within $\mathcal{D}_0$.

Case 2: The reduction $\mathcal{D} > \mathcal{D}'$ is due to a reduction involving the cut at the bottom of $\mathcal{D}$; this means that $\mathcal{D} > \mathcal{D}'$ is a consequence of $\mathcal{D} >_{red} \mathcal{D}'$.

The reduction $\mathcal{D} >_{red} \mathcal{D}'$ must then be one of the types A–C, since $>_{red}$-reductions of the types D–F are not applicable to a derivation that has a cut as its bottom-most rule application.

The derivation $\mathcal{D}'$ is directly recognizable to be inductive in the cases, where $\mathcal{D} >_{red} \mathcal{D}'$ holds because of an axiomatic cut-reduction of type A.c–f, since then $\mathcal{D}'$ consists only of an axiom (which—as a normal derivation—is inductive by Definition 2.4, (1)).

In case $\mathcal{D} >_{red} \mathcal{D}'$ holds because of an axiomatic cut-reduction of one of the types A.a or A.b, $\mathcal{D}'$ is formed from $\mathcal{D}$ by application of one or more weakenings at the bottom of either one of the immediate inductive subderivations of $\mathcal{D}$; since by Lemma 2.10 the additional application of weakening at the bottom of an inductive derivation again leads to an inductive derivation, $\mathcal{D}'$ can be seen to be inductive by one or more appeals (just as many as there are weakenings at the bottom of $\mathcal{D}'$ below $\mathcal{D}_0$ or $\mathcal{D}_1$ respectively) to Lemma 2.10.

The case of an axiomatic reduction $\mathcal{D} >_{red} \mathcal{D}'$ of type A.g is treated similarly to the case of a reduction of type B.a: In this latter case one notices firstly, that—with the notations of formulas as in the list B.b above—the subderivations $\mathcal{D}_{00}$ and $\mathcal{D}_{01}$ of $\mathcal{D}$ are inductive (since $\mathcal{D}_0$ is inductive and therefore Definition 2.4, (2) can be used); furthermore $ind(\mathcal{D}_{01}) < ind(\mathcal{D}_0)$ (again by Definition Definition 2.4). Thus the subinduction hypothesis can be applied to the derivation

$$\dfrac{\begin{array}{c}\mathcal{D}_{01}\\ B, \Gamma_0 \Rightarrow D\end{array} \qquad \begin{array}{c}\mathcal{D}_1\\ D, \Pi \Rightarrow C\end{array}}{B, \Gamma_0\Pi \Rightarrow C}\,\text{Cut} \quad , \tag{2.11}$$

since D is again the cut-formula of the original cut at the bottom of $\mathcal{D}$, but—because of $ind(\mathcal{D}_{01}) + ind(\mathcal{D}_1) < ind(\mathcal{D}_0) + ind(\mathcal{D}_1)$—the sum of the inductive complexities of this cut is now lower than that in the original cut at the bottom of $\mathcal{D}$.

Since $\mathcal{D}_{00}$ is inductive, so is

$$\begin{array}{c}\mathcal{D}_{00}\\ A \to B, \Gamma_0 \Rightarrow A\\ \hline\hline A \to B, \Gamma_0\Pi \Rightarrow A\end{array}\,\text{W}$$

(by a (finite) number of appeals to Lemma 2.10). Thus it then follows from this and the inductiveness of the derivation in (2.11) that $\mathcal{D}'$, which here has the form

$$\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ A \to B, \Gamma_0 \Rightarrow A\\ \hline\hline A \to B, \Gamma_0\Pi \Rightarrow A\end{array}\,\text{W} \qquad \dfrac{\begin{array}{c}\mathcal{D}_{01}\\ B, \Gamma_0 \Rightarrow D\end{array} \qquad \begin{array}{c}\mathcal{D}_1\\ D, \Pi \Rightarrow C\end{array}}{B, \Gamma_0\Pi \Rightarrow C}\,\text{Cut}}{A \to B, \Gamma_0\Pi \Rightarrow C}\,\text{L}\!\to \quad ,$$

is inductive.

The cases, in which $\mathcal{D} >_{red} \mathcal{D}'$ is due to a reduction of one of the types B.b or B.c, can be treated quite analogously and even easier.

In the case, where $\mathcal{D} >_{red} \mathcal{D}'$ is due to a fork cut-reduction of the type in list C above, $\mathcal{D}$ is of the form

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ A, \Gamma \Rightarrow B\end{array}}{\Gamma \Rightarrow A \to B}\,\text{R}\!\to \qquad \dfrac{\begin{array}{c}\mathcal{D}_{10}\\ A \to B, \Pi \Rightarrow A\end{array} \qquad \begin{array}{c}\mathcal{D}_{11}\\ B, \Pi \Rightarrow C\end{array}}{A \to B, \Pi \Rightarrow C}\,\text{L}\!\to}{\Gamma\Pi \Rightarrow C}\,\text{Cut} \quad .$$

By appeals to Definition 2.4, (2), $\mathcal{D}_{00}$, $\mathcal{D}_{01}$ and $\mathcal{D}_{11}$ are seen to be inductive; furthermore $ind(\mathcal{D}_{10}) < ind(\mathcal{D}_1)$ holds. Thus $ind(\mathcal{D}_0) + ind(\mathcal{D}_{10}) < ind(\mathcal{D}_0) + ind(\mathcal{D}_1)$ holds, and hence the subinduction hypothesis can be applied to the derivation $\tilde{\mathcal{D}}$, which is of the form

$$\mathrm{Cut}\ \dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A,\Gamma\Rightarrow B\end{matrix}}{\Gamma\Rightarrow A\to B}\,\mathrm{R}{\to} \qquad \begin{matrix}\mathcal{D}_{10}\\ A\to B,\Pi\Rightarrow A\end{matrix}}{\Gamma\Pi\Rightarrow A}\ ,$$

to see that this derivation is inductive. Then two consecutive uses of the induction hypothesis make clear that also the derivation $\tilde{\tilde{\mathcal{D}}}$

$$\mathrm{Cut}\ \dfrac{\mathrm{Cut}\ \dfrac{\begin{matrix}\tilde{\mathcal{D}}\\ \Gamma\Pi\Rightarrow\ A\end{matrix} \qquad \begin{matrix}\mathcal{D}_{00}\\ A,\Gamma\Rightarrow\ B\end{matrix}}{\Gamma^2\Pi\Rightarrow\ B} \qquad \begin{matrix}\mathcal{D}_{11}\\ B,\Pi\Rightarrow\ C\end{matrix}}{\Gamma^2\Pi^2\Rightarrow\ C}$$

is inductive (since $|A|, |B| < |A\to B|$ holds, the depths of the cut-formulas in the two cuts displayed within $\tilde{\tilde{D}}$ above are both smaller than the depth of the cut-formula $A\to B$ at the bottom of $\mathcal{D}$, which justifies the applicability of the induction hypothesis in both cases). A number of consecutive applications of Lemma 2.12 (as many as there are formulas in the multisets Γ and Π) then give that the derivation

$\mathcal{D}' \equiv \begin{matrix}\tilde{\tilde{D}}\\ \Gamma^2,\Pi^2\Rightarrow\ C\\ \hline\hline \Gamma\Pi\Rightarrow\ C\end{matrix}\ \mathrm{W}$, which is also identical to

$$\dfrac{\mathrm{Cut}\ \dfrac{\mathrm{Cut}\ \dfrac{\mathrm{Cut}\ \dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A,\Gamma\Rightarrow B\end{matrix}}{\Gamma\Rightarrow A\to B}\,\mathrm{R}{\to} \qquad \begin{matrix}\mathcal{D}_{10}\\ A\to B,\Pi\Rightarrow A\end{matrix}}{\Gamma\Pi\Rightarrow A} \qquad \begin{matrix}\mathcal{D}_{00}\\ A,\Gamma\Rightarrow B\end{matrix}}{\Gamma^2\Pi\Rightarrow B} \qquad \begin{matrix}\mathcal{D}_{11}\\ B,\Pi\Rightarrow C\end{matrix}}{\Gamma^2\Pi^2\Rightarrow\ C}}{\Gamma\Pi\Rightarrow\ C}\,\mathrm{C}\ ,$$

is inductive.

Thus in all cases of reductions $\mathcal{D} >_{red} \mathcal{D}'$ the inductiveness of $\mathcal{D}$ implies that one of $\mathcal{D}'$.

Now (2.10) has been show, which concludes the proof of the lemma. □

The Lemmas 2.10–2.13 now allow to conclude the Proof of Theorem 2.1.

Theorem 2.2. *Every derivation is reductive.*

Proof. It follows by an immediate induction on the depth of a derivation with the help of Lemma 2.5 and the Lemmas 2.10–2.13, that every derivation is inductive. Lemma 3.2.4 then implies the theorem. ☐

Proof. [Proof of Theorem 2.1] This is now immediate from Theorem 2.2 and the definition of "strong cut-elimination holds w.r.t. cut-elimination steps in the lists A–F" (appropriate for either of $\to$**G3mi** or $\to\bot$**G3i**). ☐

2.2 A more general version of a Strong Cut-Elimination Theorem for →G3mi and →⊥G3i

It has been pointed out before that the reduction rules in A–F come from a nearer analysis of the Cut-Elimination Theorem for the **G3**-systems in [TS96] and of the process implicit in the proof of this theorem. A proof for cut-elimination in **G3[mic]**, which proceeds by considering *topmost* occurrences of cuts with maximal cutrank[10] and by replacing subderivations ending with such cuts by derivations of either lower cutrank or by derivations containing cuts of lower cutrank together with one cut of again maximal cutrank, but now of smaller level[11] (all replacements involved in this process are locally applied transformations, if for once the necessary use of weakening- and contraction-operations is put aside, operations, that in the **G3**-calculi have global effects on the subderivations to which they are applied.)

Considering this basis for the cut-elimination reductions for $\to\bot$**G3i** as presented above, it could be argued that the strong cut-elimination result Theorem 2.1, which refers to the reduction-rules in the lists A–F is not really a very strong statement, because these reduction rules are actually too closely connected to the usual top-down cut-elimination procedure and do permit to little freedom in choosing appropriate next reduction steps for a possibly more efficient deterministic or non-deterministic alternative procedure. The reductions in A–F do not allow permutations of structural rules and cuts with each other, and so any cut-elimination procedure for $\to\bot$**G3i** that operates according to these rules cannot really gain much efficiency over the usual procedure: This is due to the fact that e.g. most work towards the completion of the elimination of a certain cut somewhere deeper down in a derivation $\mathcal{D}$ (in the sense that such an elimination—often consisting of the application of lemmas transforming whole subderivations—is treated as a single step in

[10] In [TS96] the *cutrank* of an application of cut is defined as the depth of the cut-formula plus one.

[11] The *level* of an application S of cut is in [TS96] defined as the sum of the depths of the two immediate subderivations of S.

the cut-elimination proof for **G3[mic]** in [TS96]) is likely to be often blocked from getting done in an upwards direction by unreduced cuts or by "residuals" of previous reductions higher up in $\mathcal{D}$ (by "residuals" still unremoved newly introduced cuts or structural rules of $\rightarrow\perp$**G3i**^{e} are meant); this is so because reductions A–F do not allow shifts or movements of structural rules or of cut over each other *at all*.

Under these circumstances it can be thought that the application of reductions to a given derivation $\mathcal{D}$ in an arbitrary and not top-down restricted order does not really make too much sense as actual progress with the elimination of cuts or structural rules deeper down in $\mathcal{D}$ more often than not will depend on the one with eliminations above it in $\mathcal{D}$ (so that a sequential top-down treatment of reductions might not be too much worse in its computational complexity).

But since some reductions in A–F lead to the removal of whole subderivations (Ac., Ad. and notably Fe.) there might nevertheless be *some* substantial gain thinkable w.r.t. the complexity-behaviour of a more general cut-elimination procedure operating on the basis of the reductions A–F. Still, this gain is possibly limited and more could be achieved by the introduction of reductions for the (limited) permutation of structural rules and cut.

In the following additional rules, listed in G below, for restricted permutation of $\rightarrow\perp$**G3i**^{e} 's structural rules and cut are adopted as the basis for the formulation of a strong cut-elimination-theorem for $\rightarrow$**G3mi** and for $\rightarrow\perp$**G3i**. Underlying the choice of these rules is the stipulation that upwards permutation[12] of weakening shall possess highest priority, to be followed in priority by upwards permutation of inversion, contraction and cut (in this order); this stipulation seems to be suggested by the way usual cut-elimination, using the rules A–F, is actually formulated as a deterministic procedure, yet some variations of it (with perhaps even better behavior) are still conceivable. It follows that permutations of weakening and inversion upwards over contraction and cut will be permitted, but not vice versa (as otherwise infinite reduction sequences clearly are possible); contraction will be allowed to permute upwards over cut in some cases and whenever this is possible, but this permutation is not possible in general.

G. Permutation-Reductions for Structural Rules:

(1) Weakening over inversion, contraction and cut:

a.
$$\cfrac{\cfrac{\begin{array}{c}\mathcal{D}_{00}\\ A\rightarrow B,\Gamma\Rightarrow\ C\end{array}}{B,\Gamma\Rightarrow\ C}\,\text{Inv}}{D,B,\Gamma\Rightarrow\ C}\,\text{W} \quad >_{red} \quad \cfrac{\cfrac{\begin{array}{c}\mathcal{D}_{00}\\ A\rightarrow B,\Gamma\Rightarrow\ C\end{array}}{A\rightarrow B,D,\Gamma\Rightarrow\ C}\,\text{W}}{B,D,\Gamma\Rightarrow\ C}\,\text{Inv}$$

[12]([Vest99] uses the very visual expression "upwards propagation" in this respect)

b. $$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{A, A, \Gamma \Rightarrow C}}{A, \Gamma \Rightarrow C}\,\mathrm{C}}{B, A, \Gamma \Rightarrow C}\,\mathrm{W} \quad >_{red} \quad \dfrac{\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{A, A, \Gamma \Rightarrow C}}{A, A, B, \Gamma \Rightarrow C}\,\mathrm{W}}{A, B, \Gamma \Rightarrow C}\,\mathrm{C}$$

c. $$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{\Gamma \Rightarrow D} \qquad \genfrac{}{}{0pt}{}{\mathcal{D}_{01}}{D, \Pi \Rightarrow C}}{\Gamma\Pi \Rightarrow C}\,\mathrm{Cut}}{A, \Gamma\Pi \Rightarrow C}\,\mathrm{W} \quad >_{red} \quad \dfrac{\mathrm{W}\,\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{\Gamma \Rightarrow D}}{A, \Gamma \Rightarrow D} \qquad \genfrac{}{}{0pt}{}{\mathcal{D}_{1}}{D, \Pi \Rightarrow C}}{A, \Gamma\Pi \Rightarrow C}\,\mathrm{Cut}$$

(2) Inversion over contraction and cut:

d. $$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{D, D, A \to B, \Gamma \Rightarrow C}}{A \to B, D, \Gamma \Rightarrow C}\,\mathrm{C}}{B, \Gamma\Pi \Rightarrow C}\,\mathrm{Inv} \quad >_{red} \quad \dfrac{\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{A \to B, D, D, \Gamma \Rightarrow C}}{D, D, B, \Gamma \Rightarrow C}\,\mathrm{Inv}}{D, B, \Gamma \Rightarrow C}\,\mathrm{C}$$

e. $$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{A \to B, A \to B, \Gamma \Rightarrow C}}{A \to B, \Gamma \Rightarrow C}\,\mathrm{C}}{B, \Gamma \Rightarrow C}\,\mathrm{Inv} \quad >_{red} \quad \dfrac{\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{A \to B, A \to B, \Gamma \Rightarrow C}}{B, A \to B, \Gamma \Rightarrow C}\,\mathrm{Inv}}{B, B, \Gamma \Rightarrow C}\,\mathrm{Inv}}{B, \Gamma \Rightarrow C}\,\mathrm{C}$$

f. $$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{A \to B, \Gamma_0 \Rightarrow D} \qquad \genfrac{}{}{0pt}{}{\mathcal{D}_{01}}{D, \Pi \Rightarrow C}}{A \to B, \Gamma_0\Pi \Rightarrow C}\,\mathrm{Cut}}{B, \Gamma_0\Pi \Rightarrow C}\,\mathrm{Inv} \quad >_{red}$$

$$\dfrac{\dfrac{\genfrac{}{}{0pt}{}{\mathcal{D}_{00}}{A \to B, \Gamma_0 \Rightarrow D}}{B, \Gamma_0 \Rightarrow D}\,\mathrm{Inv} \qquad \genfrac{}{}{0pt}{}{\mathcal{D}_{01}}{D, \Pi \Rightarrow C}}{B, \Gamma_0\Pi \Rightarrow C}\,\mathrm{Cut}$$

g. Similarly and symmetrically to case f, if $A \to B$ occurs in Π in the conclusion of $\mathcal{D}_{01}$ in the derivation to be reduced in case f (but not in the antecedent $\Gamma \equiv A \to B, \Gamma_0$ of the conclusion of $\mathcal{D}_{00}$ there).

(3) Contraction over cut:

h.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A,A,\Gamma_0 \Rightarrow D\end{matrix} \qquad \begin{matrix}\mathcal{D}_{01}\\ D,\Pi\Rightarrow\ C\end{matrix}}{A,A,\Phi_0\Rightarrow\ C}\text{Cut}}{A,A,\Phi_1\Rightarrow\ C}\text{C}}{\vdots}\text{C}}{A,A,\Phi_n\Rightarrow\ C}\text{C}}{A,\Phi_n\Rightarrow\ C}\text{C} \qquad >_{red}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ \text{C}\,\dfrac{A,A,\Gamma_0\Rightarrow D}{A,\Gamma_0\Rightarrow D}\end{matrix}\qquad \begin{matrix}\mathcal{D}_{01}\\ D,\Pi\Rightarrow\ C\end{matrix}}{A,\Phi_0\Rightarrow\ C}\text{Cut}}{A,\Phi_1\Rightarrow\ C}\text{C}}{\vdots}\text{C}}{A,\Phi_n\Rightarrow\ C}\text{C}$$

where $\Phi_0 \equiv \Gamma_0\Pi$ and the first n contractions below cut cannot be permuted upwards over cut, i.e. their respective principal formulas occur just once in Γ_0 and Π respectively.

i. Similarly and symmetrically on the right as in case h, if A occurs twice in Π in the derivation to be reduced in case h.

Theorem 2.3. (Strong Cut-Elimination for →G3mi and →⊥G3i, more general version)
Strong cut-elimination holds for the calculi **→G3mi** *and* **→⊥G3i** *with respect to the reduction steps in the lists A–G.*

The proof of this theorem resisted our [my, C.G.] attempts to give it in the framework of Dragalin's concepts and notations in his proof for strong cut-elimination for LK and LJ (these concepts and notations have been used above in the proof of Theorem 2.1). We can therefore at present give only a sketch of a very *ad hoc* version of the proof.

Proof. [Sketch of the Proof] The use of the notation $\mathcal{D} > \mathcal{D}'$ for two (→**G3mi**^{e}+Cut)- or (→⊥**G3i**^{e}+Cut)-derivations $\mathcal{D}$ and $\mathcal{D}'$ will here be understood as extending the meaning of "$\mathcal{D}$ reduces to $\mathcal{D}'$" as defined in Definition 2.2, (1), by including also the new reduction rules of the list G.

(1) Strong normalization holds for every derivation $\mathcal{D}$ in the systems (→**G3mi**^{e}+Cut) or in (→⊥**G3i**^{e}+Cut) *with respect to the reductions D–G* (i.e. w.r.t. all rules in

A–G *but* the cut-reductions A–C): This can be proved in a straightforward way by using and slightly adapting the concepts of the proof for Theorem 2.1.

Furthermore it can be easily checked, that applications of one of the reductions D–G do not increase the *logical size* (= the number of applications of logical rules) of a derivation $\mathcal{D}$, nor the logical depth, the logical level and the rank of a particular application S of a cut in $\mathcal{D}$ (the *logical depth* $\|\mathcal{D}'\|$ of a derivation $\mathcal{D}'$ is defined similarly to the (usual) depth of $\mathcal{D}'$, but by counting only rule-applications of logical rules are; the *logical level* $ll(S)$ of an application S of cut in a derivation $\mathcal{D}'$ is defined as the sum of the logical depths of the two immediate subderivations of S in $\mathcal{D}'$).

(2) The maximum number of cut-reduction steps from A–C for completely eliminating a cut S at the same time with all its residuals occuring during a reduction-sequence $\sigma \equiv \mathcal{D} > \mathcal{D}_1 > \mathcal{D}_2 > \mathcal{D}_3 > \ldots$ starting from a derivation $\mathcal{D}$, which does only contain the cut S, can be calculated from the logical level $ll(S)$ and the *rank* $rank(S)$ of S (= the depth of the cut-formula of S plus 1) alone. Let this maximum number of cut-reduction steps be bounded by a function $c_1(ll(S), rank(S))$. The logical depth of every resulting cut-free derivation $\mathcal{D}'$ may—in comparison with $\mathcal{D}$—have increased (due to applications of cut-reductions of the type B.c), but it can still be bounded by a function $l_1(\|\mathcal{D}\|, ll(S), rank(S))$.

(3) Considering a ($\rightarrow$**G3mi**^{e}+Cut)- or ($\rightarrow\perp$**G3i**^{e}+Cut)-derivation $\mathcal{D}$ containing exactly two cuts S_1 and S_2, it is possible—since the reductions A–G do not permit the permutation of two applications of the cut rule over the other—to find a bound for the maximal number of steps caused by cut-reductions in every reduction sequence $\sigma \equiv \mathcal{D} > \mathcal{D}_1 > \mathcal{D}_2 > \mathcal{D}_3 > \ldots$ starting from $\mathcal{D}$. This can be achieved by first looking at the steps necessary for the removal of S_1 and S_2 separately, if (a) S_1 and S_2 occur in subderivations of $\mathcal{D}$ apart from each other, or successively, if (b) an immediate subderivation of S_1 contains S_2 or (c) the opposite is true.

But for example in situation (b) it has to be taken into account, that (i) the complete removal of S_1 first, together with all its possible residuals, before dealing with S_2 may increase increase the logical level of S_2 in the resulting derivation to $d + l_1(d, ll(S_1), rank(S_1))$, where d is the logical depth of S_2 in $\mathcal{D}$; and furthermore that (ii) the reduction of S_2 and or of any residual of S_2 may—if the reduction happens to be a fork cut-reduction—almost double the the amount of steps that have previously been necessary for the complete elimination of S_1 alone.

Still, and over all, the amount of cut-elimination steps in σ stays finite and can be bounded by a function $c_2(ll(S_2), \max(rank(S_1) + rank(S_2)))$, where S_2 is here taken to be the bottom-most of the cuts S_1 and S_2. The logical depth of every resulting cut-free derivation $\mathcal{D}'$ may then also be bounded by a function

$$l_2(\|\mathcal{D}\|, ll(S_2), \max(rank(S_1) + rank(S_2))) \ .$$

(4) Carrying on in this way step by step it is then possible to find a bound $c_n(\|\mathcal{D}\|, r)$ for the maximum number of steps due to cut-elimination reductions of A–C in an arbitrary reduction sequence $\sigma \equiv \mathcal{D} > \mathcal{D}_1 > \mathcal{D}_2 > \mathcal{D}_3 > \dots$ starting with $\mathcal{D}$, where n is the number of applications of cut in $\mathcal{D}$, and r is the *cutrank* of $\mathcal{D}$, i.e the maximal rank of all applications of cut in $\mathcal{D}$.

At the same time a bound $l_n(\|\mathcal{D}\|, l, r))$, where l means the maximum logical level of all cuts in $\mathcal{D}$, for the logical depth of every resulting cut-free derivation $\mathcal{D}'$ can be given as well.

(5) Strong normalization for a ($\rightarrow$**G3mi**^{e}+Cut)- or ($\rightarrow\perp$**G3i**^{e}+Cut)-derivation $\mathcal{D}$ with respect to the rules A–G then follows from (1) and (4). This is true, since in an arbitrarily chosen reduction sequence $\sigma \equiv \mathcal{D} > \mathcal{D}_1 > \mathcal{D}_2 > \mathcal{D}_3 > \dots$ the number of consecutive steps caused by reductions of type E–G always has to be finite (due to (1)) and therefore after every sufficiently long subpart of σ consisting only of reductions of type D–G a cut-reduction has to follow. But then by (5) also the number of reductions in σ, that are due to cut-reductions of A–C, is bounded as well. As a consequence σ must be of finite length. □

Acknowledgement

I do want to thank at this place a number of people, to whom I felt and feel thankful for many things, often different and separate ones, during my stay in Amsterdam over the last year.

I do want to thank my advisor from the institute, Domenico Zambella (for helping me through the first few months of my study with frequent supervising and guiding conversations as well as explanations), the program manager Ingrid van Loon (for too many things than I could have kept track of them, let alone that I could record them here, I can only say that I could always go to Ingrid and ask her for help or information and she would certainly do her best to find assistence) and my advisor with the work at this thesis, Prof. A.S. Troelstra (above all for not giving up on me notwithstanding my particularities and "directions for use"). I want to thank the members of the exam committee, Dr. Kees Doets, Dr. Maarten de Rijke and Prof. Troelstra for their comments and questions about my thesis as referees. And I want to thank all of the lecturers and teachers, that I had at the Universiteit van Amsterdam over the last year for the very positive way how it was possible to interact with them (I should mention them all: Dr. Frank Veltman, Dr. Domenico Zambella, Dr. Kees Doets, Prof. A.S. Troelstra, Dr. Maarten Marx, Dr. Yde Venema), and also the Scientific Coordinator of the MoL-Program, Prof. Dick de Jongh.

Important thanks also go to my colleagues Walter Dean, Giorgi ("George") Goguadze, Rafael Accorsi, Sophia Vellisseratou and Nick Arevadze from the 1998-99 Master-of-Logic Program of the ILLC, with whom I had the good experience of sharing a nice office-room at Sarphatistraat 143, which we all liked really a lot. I am particularly thankful for the experience of many fruitful discussions with Walt and George (and at some occasions also with Annette Bleeker) about mathematical details on our blackboard in the office (it was there that I rediscovered mathematics as a field to think, debate and find common ground with others). I am also in other respect heavily indebted to Rafa for many hints, how to deal with problems on the computers, and to Sophia for several things (e.g. for providing me with as good advice as she could think of on how possibly to cope with the problem of finding housing in September).

I also have to thank the people on my floor at Cleydertweg 29 (it is the 7th floor there) for the exceptional good community, that they have formed for the time until the summer

(Thanks a lot, family!). I want to mention particularly Arndt Riester, (for sharing his informations with me and combining his efforts with mine in summer during our common problem of finding a new room) Dominga Rogolino (for frequently sending me wishes from ultimately as far away as Parma, Italy), and then also Martha Cecilia Ruiz-Muriel, Sebastian Brand and Elena Govberg-Afros. No, Boz, Tulio, Regina, Sofia, Richard, Alice, Claudia, Dörthe, Joyce, Perry, Jan, Anneli ... you are and will not be forgotten.

I do want to thank many of the new MoL-students (hey Katarina, Shai, Mariana, Clemens) and Walt and Sophia as well as Yde Venema, Eva Hoogland, Marc Pauly and Domenico Zambella for coming to my defense. I was really fond about that! (Greetings go also to Alex for our common night-time computer-experiences and to Mischa and Yure.)

I also want to thank my mother, Herta Grabmayer, for many things, that she did for me from Austria, amongst many others in helping me with formalities, planning and doing the necessary financial transactions for me and in keeping me informed about things going on at home. The driving-abilities of my father, Erwin Grabmayer, that initially managed to get me to Amsterdam safely are at the same time not forgotten.

Yes, and I do have to thank my Amsterdam landlord, Aad Wolsink, for many things, but most importantly for taking me as a tenant in his flat, for being nice to me and for coming to my defense.

Once again: Let me please thank you all very much for that! (Sorry to all, who I have forgotten to mention in their right explicitly!)

Clemens Grabmayer

Appendix A: Handout at the defense on 15th October 1999

Example underlying a "Computational Anomaly" of distinct derivations $\mathcal{D}_n$ in $\to$G3mi, that all reduce to the same derivation $\mathcal{D}'$ by cut-elimination[13]

Let for $n \in \mathbb{N}$ the derivation $\mathcal{D}_n$ be

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_{00}\\ A, B, C, C \to C \Rightarrow B\end{matrix}}{B, C, C \to C \Rightarrow A \to B}\,\mathrm{R}{\to} \qquad \dfrac{A \to B, A, C, C \to C \Rightarrow A \qquad \begin{matrix}\mathcal{D}_{11}^{(n)}\\ B, A, C, C \to C \Rightarrow C\end{matrix}}{A \to B, A, C, C \to C \Rightarrow C}\,\mathrm{L}{\to}}{A, B, C, C \to C \Rightarrow C}\,\mathrm{Cut} \quad ,$$

where A, B and C are atomic formulas and $\mathcal{D}_{00}$ is the derivation

$$\dfrac{C \to C, C, A, B \Rightarrow C \qquad C, C, A, B \Rightarrow B}{C \to C, C, A, B \Rightarrow B}\,\mathrm{L}{\to}$$

and[14] $\mathcal{D}_{11}^{(n)} \equiv \mathcal{D}^{(n)}[A, B]$ with $\mathcal{D}^{(n)}$ the derivation[15]

[13] Here (1) cut-elimination is performed similar as in [TS96] for the systems **G3[mi]**, but as a stepwise process of locally applied transformations, and (2) a multiple-contraction rule is used for doing this.

[14] ($\tilde{\mathcal{D}}[E_1, \ldots, E_n]$ for formulas $E_1, \ldots, E_n$ and a derivation $\tilde{\mathcal{D}}$ means the derivation that results from the addition of the formulas $E_1, \ldots, E_n$ to the antecedent of every sequent in $\tilde{\mathcal{D}}$.)

[15] The derivations $\mathcal{D}^{(n)}$, if P were read for C, correspond (in the case of the untyped system $\to$**G3mi** here it is better to say: relate) to the derivation-term $\imath_n$ used by Vestergaard, which in the setting of the system $\boldsymbol{\mathcal{G}}^+$ is defined on p. 28; more precisely, $\Phi_0(\mathcal{D}_{\imath_n})$ equals $\mathcal{D}^{(n)}$, if P in $\Phi_0(\mathcal{D}_{\imath_n})$ is exchanged by C and $\mathcal{D}_{\imath_n}$ is the $\boldsymbol{\mathcal{G}}_\mathbf{0}^+$-derivation corresponding to $\imath_n$.

$$
\begin{array}{c}
\text{L}\!\to\dfrac{\text{L}\!\to\dfrac{C\to C,C\Rightarrow C\qquad C,C\Rightarrow C}{C\to C,C\Rightarrow C}\qquad C,C\Rightarrow C}{\vdots}\\[1ex]
\text{L}\!\to\dfrac{\overline{C\to C,C\Rightarrow C}\qquad\qquad C,C\Rightarrow C}{C\to C,C\Rightarrow C}
\end{array}
$$

(with n applications of L$\to$). The first cut-elimination step in $\mathcal{D}_n$ is a fork-reduction step:

$$
\begin{array}{c}
\text{Cut}\dfrac{\text{Cut}\dfrac{\text{Cut}\dfrac{\text{R}\!\to\dfrac{\begin{array}{c}\mathcal{D}_{00}\\ A,B,C,C\to C\Rightarrow B\end{array}}{B,C,C\to C\Rightarrow A\to B}\qquad A\to B,A,C,C\to C\Rightarrow A}{A,B,C^2,(C\to C)^2\Rightarrow A}\qquad \begin{array}{c}\mathcal{D}_{00}\\ A,B,C,C\to C\Rightarrow B\end{array}}{A,B^2,C^3,(C\to C)^3\Rightarrow B}\qquad \begin{array}{c}\mathcal{D}_{11}^{(n)}\\ B,A,C,C\to C\Rightarrow C\end{array}}{A^2,B^2,C^4,(C\to C)^4\Rightarrow C}\\
\hline\hline
A,B,C^2,(C\to C)^2\Rightarrow C
\end{array}\text{C}
$$

Then an axiomatic cut-reduction step follows:

$$
\begin{array}{c}
\text{Cut}\dfrac{\text{Cut}\dfrac{A,B,C^2,(C\to C)^2\Rightarrow A\qquad \dfrac{C\to C,C,A,B\Rightarrow C\qquad C,C,A,B\Rightarrow B}{C\to C,C,A,B\Rightarrow B}\text{L}\!\to}{A,B^2,C^3,(C\to C)^3\Rightarrow B}\qquad \begin{array}{c}\mathcal{D}_{11}^{(n)}\\ B,A,C,C\to C\Rightarrow C\end{array}}{A^2,B^2,C^4,(C\to C)^4\Rightarrow C}\\
\hline\hline
A,B,C^2,(C\to C)^2\Rightarrow C
\end{array}\text{C}
$$

Next a permutation-reduction step of the topmost cut over L$\to$, followed by two axiomatic cut-reductions yield:

$$
\begin{array}{c}
\text{Cut}\dfrac{\text{L}\!\to\dfrac{C\to C,(C\to C)^2,C^3,A,B^2\Rightarrow C\qquad C,C^3,(C\to C)^2,A,B^2\Rightarrow B}{A,B^2,C^3,(C\to C)^3\Rightarrow B}\qquad \begin{array}{c}\mathcal{D}_{11}^{(n)}\\ B,A,C,C\to C\Rightarrow C\end{array}}{A^2,B^2,C^4,(C\to C)^4\Rightarrow C}\\
\hline\hline
A,B,C^2,(C\to C)^2\Rightarrow C
\end{array}\text{C}
$$

The following derivation is the result of a permutation-step of Cut over L$\to$:

$$
\begin{array}{c}
\dfrac{\text{W}\begin{array}{c}C\to C,(C\to C)^2,C^3,A,B^2\Rightarrow C\\ \hline\hline C\to C,(C\to C)^3,C^4,A^2,B^2\Rightarrow C\end{array}\qquad \dfrac{C,C^3,(C\to C)^2,A,B^2\Rightarrow B\qquad \begin{array}{c}\mathcal{D}_{11}^{(n)}\\ B,A,C,C\to C\Rightarrow C\end{array}}{C,C^4,(C\to C)^3,A^2,B^2\Rightarrow C}\text{Cut}}{A^2,B^2,C^4,(C\to C)^4\Rightarrow C}\text{L}\!\to\\
\hline\hline
A,B,C^2,(C\to C)^2\Rightarrow C
\end{array}\text{C}
$$

We then reach the following cut-free derivation through the application of an axiomatic cut-reduction (and by an axiomatic multiple-weakening reduction):

$$\dfrac{\dfrac{C \to C, (C \to C)^3, C^4, A^2, B^2 \Rightarrow C \qquad \begin{array}{c}\mathcal{D}^{(n)}[A^2, B^2, C^4, (C \to C)^2] \\ C, C \to C, C^4, (C \to C)^2, A^2, B^2 \Rightarrow C\end{array}}{A^2, B^2, C^4, (C \to C)^4 \Rightarrow C}\,\mathrm{L}\!\to}{A, B, C^2, (C \to C)^2 \Rightarrow C}\,\mathrm{C}^{\{A,B,C^2,(C\to C)^2\}} \quad .$$

If now all contractions are permuted upwards over L→ *simultaneously* (in the analogous sense as on derivation-terms multiple-contraction is permuted upwards in a $\boldsymbol{\mathcal{G}}^+$-derivation according to the contraction rewrite-rule E.c, second case), inversion has to be used for the treatment of the left premise. This leads to:

$$\dfrac{\mathrm{C}^{\{A,B,C^2,(C\to C)^2\}}\,\dfrac{C \to C, (C \to C)^3, C^4, A^2, B^2 \Rightarrow C}{C \to C, C \to C, C^2, A, B \Rightarrow C} \qquad \dfrac{\dfrac{\begin{array}{c}\mathcal{D}^{(n)}[A^2, B^2, C^4, (C \to C)^2] \\ C \to C, C, C^4, (C \to C)^2, A^2, B^2 \Rightarrow C\end{array}}{C^6, (C \to C)^2, A^2, B^2 \Rightarrow C}\,\mathrm{Inv}_{C\to C,C}}{C, C \to C, C^2, A, B \Rightarrow C}\,\mathrm{C}^{\{A,B,C^3,C\to C\}}}{(C \to C)^2, C^2, A, B \Rightarrow C}\,\mathrm{L}\!\to$$

If now inversion is permuted upwards, almost all of $\mathcal{D}^{(n)}[A^2, B^2, C^4, (C \to C)^2]$ gets lost (with the exception of the axiom in the bottom-most application of L→ in it):

$$\dfrac{\mathrm{C}^{\{A,B,C^2,(C\to C)^2\}}\,\dfrac{C \to C, (C \to C)^3, C^4, A^2, B^2 \Rightarrow C}{C \to C, C \to C, C^2, A, B \Rightarrow C} \qquad \dfrac{C, C^5, (C \to C)^2, A^2, B^2 \Rightarrow C}{C, C \to C, C^2, A, B \Rightarrow C}\,\mathrm{C}^{\{A,B,C^3,C\to C\}}}{(C \to C)^2, C^2, A, B \Rightarrow C}\,\mathrm{L}\!\to$$

Two axiomatic multiple-contraction reductions lead to:

$$\dfrac{C \to C, C \to C, C^2, A, B \Rightarrow C \qquad C, C \to C, C^2, A, B \Rightarrow C}{(C \to C)^2, C^2, A, B \Rightarrow C}\,\mathrm{L}\!\to \quad .$$

This result $\mathcal{D}'$ of the cut-elimination procedure performed at $\mathcal{D}_n$ is now clearly independent of n.

For all $n \in \mathbb{N}$ derivations $\bar{\mathcal{D}}_n$ (given on page 32 in the form of corresponding $\boldsymbol{\mathcal{G}}^+$-derivation-terms $\bar{t}_n$) paralleling $\mathcal{D}_n$ in the typed system $\boldsymbol{\mathcal{G}}^+$, $\bar{\mathcal{D}}_n$ correspond to the natural-deduction derivation $\Phi_0(\bar{\mathcal{D}}_n)$

$$
\cfrac{(C \to C)^s \qquad \cfrac{\vdots}{\cfrac{(C \to C)^s \qquad \cfrac{(C \to C)^s \qquad C^z}{C}\ {\to}\mathrm{E}}{C}\ {\to}\mathrm{E}}}{C}\ {\to}\mathrm{E}
$$

(with n applications of $\to$E), whereas the result $\bar{\mathcal{D}}'$ (relating to $\mathcal{D}'$) of the (usual) cut-elimination-procedure applied to $\bar{\mathcal{D}}_n$ (largely paralleling in $\boldsymbol{\mathcal{G}}^+$ the above reductions in untyped $\boldsymbol{\to}$**G3mi**) corresponds just to the natural-deduction image $\Phi_0(\bar{\mathcal{D}}')$ of trivial shape

$$C^z \ .$$

This is what constitutes an "anomaly" here.

The drawback at this my example is, that if contractions were not permuted upwards in the gathered form of a multiple-contraction but as single contractions, the example would not result in an "anomaly". Although it accounts for an a bit more careful formulation of one reduction-rule (for reductions on derivation-terms this is the contraction rewrite-rule E.c), the example does not cover the most general possible situation.

I do think that with a bit more effort an "anomaly" could also be constructed if the typed system allowed only single-contractions (the proof of cut-elimination is then still possible in the way I gave it). But I have no example for this most general situation, yet.

Bibliography

[Drag79] Dragalin, A.G.: *Mathematical Intuitionism* (Introduction to Proof Theory), Moscow, 1979; Translations of Mathematical Monographs, Volume 67, American Mathematical Society, Providence, Rhode Island, 1988.

[Fef75] Feferman, S.: Review of [Pra71] in: *Journal of Symbolic Logic 2*, Vol. 40, 1975, p. 232—234.

[Gen35] Gentzen, G.: "Untersuchungen über das logische Schließen I,II", *Mathematische Zeitschrift 39*, S. 176–210, 405–431.

[Kl52] Kleene, S.C.: *Introduction to Metamathematics*, North-Holland Publishing Company, Amsterdam, 1952.

[Kr71] Kreisel, G.: "A Survey of Proof-Theory II", p. 109–170 in: Fenstad, J.E.[ed.]: *Proceedings of the 2nd Scandinavian Logic Symposium*, North-Holland Publishing Company, Amsterdam–London, 1971.

[Pott77] Pottinger, G.: "Normalization as a homomorphic image of Cut-Elimination", *Annals of Mathematical Logic 12*, 1977, p. 323–357.

[Pra65] Prawitz, D.: *Natural Deduction* (A Proof-Theoretical Study), Almqvist& &Wiksell, Stockholm, Göteborg, Uppsala, 1965.

[Pra71] Prawitz, D.: "Ideas and Results in Proof-Theory", p. 235–307 in: Fenstad, J.E. [ed]: *Proceedings of the 2nd Scandinavian Logic Symposium*, North-Holland Publishing Company, Amsterdam, London, 1971.

[TS96] Troelstra, A.S., Schwichtenberg, H.: *Basic Proof Theory*, Cambridge Tracts in Theoretical Computer Science, Cambridge University Press, 1996.

[Ung92] Ungar, A.M.: *Normalization, Cut-Elimination and the Theory of Proofs*, CSLI Lecture Notes No. 28, Leland Stanford Junior University, 1992.

[Vest99] Vestergaard, R.: "Revisiting Kreisel: A Computational Anomaly in the Troelstra–Schwichtenberg G3i-System", March 1999, available from R. Vestergaard's Homepage: *http://www.cee.hw.ac.uk/~jrvest/* .

[Zu74] Zucker, J.: "The Correspondence between Cut-Elimination and Normalization", *Annals of Mathematical Logic 7*, 1974, p. 1–112, 156.