

# **A Calculus of Falsifiers**

Cameron Allett

A thesis presented for the degree of Doctor of Philosophy

Department of Computer Science  
University of Bath  
England, UK  
October 2025

### **Copyright Notice**

Attention is drawn to the fact that copyright of this thesis rests with the author and copyright of any previously published materials included may rest with third parties. This thesis has been supplied on condition that anyone who consults it understands that they must not copy it or use material from it except as licenced, permitted by law or with the consent of the author or other copyright owners, as applicable.

### **Declarations**

The material presented here for examination for the award of a research degree by research has not been incorporated into a submission for another degree.

I am the author of this thesis, and the work described therein was carried out by myself personally.

## Abstract

The falsifier calculus, a deep-inference proof system for first-order predicate logic in the language of Hilbert's epsilon-calculus, is introduced. It uses a novel inference rule, the falsifier rule, to introduce epsilon-terms into proofs, distinct from the critical axioms of the traditional epsilon-calculus. The falsifier rule is a generalisation of one of the quantifier-shifts, inference rules for shifting quantifiers inside and outside of formulae. Like the epsilon-calculus and proof systems which include quantifier-shifts, the falsifier calculus admits non-elementarily smaller cut-free proofs of certain classes of first-order theorems than the sequent calculus.

Analogous to the way in which Herbrand's Theorem decomposes a proof into a first-order and a propositional part, connected by a Herbrand disjunction as an intermediate formula, the falsifier calculus provides a new decomposition theorem for first-order proofs which gives rise to a new notion of intermediate formula in the epsilon-calculus, falsifier disjunctions. It is demonstrated that certain classes of first-order theorems admit non-elementarily smaller falsifier disjunctions than Herbrand disjunctions.

Through these properties, the falsifier calculus contributes new insights to our understanding of the structure and complexity of proofs in first-order predicate logic. The phenomenon of the non-elementary compression of cut-free proofs has long been of interest in structural proof theory, and the falsifier calculus provides a new perspective on what properties make a proof system admit this compression as well as the role of epsilon-terms in this compression. Furthermore, the decomposition theorem admitted by the falsifier calculus provides a novel perspective on the structure of Herbrand disjunctions and their complexity, yielding new insights into one of the fundamental theorems of classical proof theory.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Proof Theory . . . . .	6
1.2	Predicate Logic . . . . .	7
1.3	Herbrand's Theorem . . . . .	8
1.4	The Epsilon-Calculus . . . . .	11
1.5	Deep Inference . . . . .	12
1.6	Overview of Thesis . . . . .	15
<b>2</b>	<b>Preliminaries</b>	<b>17</b>
2.1	Language . . . . .	17
2.2	Semantics . . . . .	20
2.3	Open Deduction . . . . .	22
<b>3</b>	<b>The Structure of First-Order Proofs</b>	<b>31</b>
3.1	Proof Systems . . . . .	31
3.2	Rule Decomposition . . . . .	37
3.3	Quantifier-Shifts and Non-Elementary Compression . . . . .	44
3.4	Case Analysis Extraction and Herbrand's Theorem . . . . .	47
<b>4</b>	<b>The Falsifier Calculus</b>	<b>50</b>
4.1	The Falsifier Rule . . . . .	52
4.2	The Falsifier Decomposition Theorem . . . . .	60
4.3	Examples . . . . .	63
4.4	Comparison With Existing Proof Systems and Interpretations . . . . .	67
<b>5</b>	<b>Proof of The Falsifier Decomposition Theorem</b>	<b>72</b>
5.1	Rule Decomposition . . . . .	73
5.2	Phase 1 . . . . .	73
5.3	Phase 2 . . . . .	78
5.4	Phase 3 . . . . .	86
5.5	Final Proof . . . . .	93
5.6	Optimisations . . . . .	94
<b>6</b>	<b>The Restricted Falsifier Calculus</b>	<b>96</b>
6.1	Universal Instantiation Rules . . . . .	96
6.2	A Restricted Form of The Falsifier Decomposition Theorem . . . . .	99
<b>7</b>	<b>Conclusion</b>	<b>103</b>
<b>A</b>	<b>Appendix</b>	<b>110</b>
A.1	Tangential Remarks . . . . .	110

## List of Figures

1	The inference rules of $\text{SKSg}_P$ and $\text{SKSg1}$ . . . . .	33
2	The inference rules of $\text{SKS1}$ . . . . .	34
3	The inference rules of $\text{SKSg}_\varepsilon \cup \{\exists, \text{qc}\downarrow\}$ . . . . .	64
4	Transformation for case (1) of Lemma 5.7 . . . . .	89
5	Transformation for case (2) of Lemma 5.7 . . . . .	90
6	Transformation for case (3) of Lemma 5.7 . . . . .	90

# 1 Introduction

The relation of the sign to the thing signified is being destroyed, the game of exchanges between signs is being multiplied of itself and for itself. And the increasing complication demands that there should be signs for signs...

---

S. Weil

There is a long history of investigating the shape and structure of mathematical proofs. Since the inception of proof theory in the early twentieth century, a range of techniques have been developed to represent and manipulate formal proofs. While early investigations in proof theory were concerned primarily with establishing the completeness and consistency of formal proof systems for various logics, research has tended towards investigations into the complexity and normalisation theory of proofs as the field has matured. This thesis will contribute towards these efforts by providing a new normalisation result for proofs in first-order predicate logic and a new perspective on an old problem in the study of proof complexity: the compression of first-order proofs.

This thesis will introduce a new proof system for first-order predicate logic, called the *falsifier calculus*, and explore its proof-theoretic properties. The falsifier calculus is defined in the language of Hilbert's *epsilon-calculus* [57], which extends the language of first-order predicate logic by  $\varepsilon$ -terms  $\varepsilon_x A(x)$  for all variables  $x$  and formulae  $A(x)$  such that, for a given semantics  $\llbracket - \rrbracket_{\mathbb{D}}$  with domain  $\mathbb{D}$ , each  $\varepsilon$ -term  $\varepsilon_x A(x)$  is assigned an element in  $\mathbb{D}$  by

$$\llbracket \varepsilon_x A(x) \rrbracket_{\mathbb{D}} = \begin{cases} d & \text{if there exists some } d \in \mathbb{D} \text{ such that } \llbracket A(d) \rrbracket_{\mathbb{D}} \\ a & \text{for some arbitrary } a \in \mathbb{D}, \text{ otherwise} \end{cases}$$

where  $d$  is chosen by a choice function on  $\mathcal{P}(\mathbb{D})$  and  $a \in \mathbb{D}$  is fixed. The falsifier calculus uses a novel inference rule for introducing  $\varepsilon$ -terms into proofs, called the *falsifier rule*  $\varepsilon$ , which is given by

$$\varepsilon \frac{\forall x (A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))}$$

and is distinct from the critical axioms used to introduce  $\varepsilon$ -terms into proofs in the traditional epsilon-calculus. Through the use of this rule, the falsifier calculus admits cut-free proofs of certain classes of first-order theorems which are non-elementarily smaller than their corresponding smallest cut-free proofs in traditional proof systems such as the sequent calculus. This compression is also admitted by several other proof systems [2, 17] so that the falsifier calculus provides a novel insight into the properties that make a proof system admit this compression as well as the role of  $\varepsilon$ -terms in this compression. Furthermore, the falsifier calculus admits a new decomposition theorem for proofs in first-order predicate logic, the *Falsifier Decomposition Theorem*, providing a new perspective on their structure. The Falsifier Decomposition Theorem can be understood as an analogue to Herbrand's Theorem in which  $\varepsilon$ -terms may also occur in the upper propositional segment of the proof, resulting in a non-elementary compression over proofs in Herbrand normal form. The Falsifier Decomposition Theorem gives rise to the notion of *falsifier disjunctions*, which are analogues to Herbrand disjunctions in the language of the epsilon-calculus such that certain classes of first-order theorems admit non-elementarily smaller falsifier disjunctions than Herbrand disjunctions, yielding a new perspective on the structure of Herbrand disjunctions and their complexity.

The falsifier calculus is defined using a proof formalism that admits *deep inference*, which allows for greater freedom in applying inference rules than conventional proof systems by allowing them to apply at arbitrary depth inside of formulae. This yields a rich normalisation theory which is not observable in traditional proof formalisms, and which will be exploited

in this thesis to prove normalisation results for first-order predicate logic using the falsifier calculus.

## 1.1 Proof Theory

The field of proof theory is generally considered to have been initiated with the introduction of Hilbert's program in the 1920s. The discovery of several apparent paradoxes in foundational areas of mathematics in the late nineteenth and early twentieth centuries had led to a "foundational crisis of mathematics", to which Hilbert proposed a research program towards resolving. Hilbert's program sought to provide a formal axiomatic foundation for all areas of mathematics and demonstrate the consistency of these axioms, i.e., that they were free of contradiction. With this came the introduction of the first formal deductive systems in mathematics, such as *Hilbert-Frege systems* [32], which are comprised of a list of formula axioms and the inference rule of *modus ponens*, the principle that the formula  $B$  may be deduced from the formulae  $A$  and  $A \rightarrow B$ , which is used to derive valid formulae from the axioms. Hilbert's program is usually understood to have come to an end with the publication of Gödel's second incompleteness theorem in 1931 [44], which demonstrated that any deductive system strong enough to prove the axioms of Peano arithmetic is unable to prove its own consistency.

Although Hilbert-Frege systems and initial investigations in proof theory were able to provide complete proof systems for logics foundational to mathematics such as propositional classical logic and first-order predicate logic, the underlying deductive systems were unwieldy and admitted a proof structure which was far removed from the structure of actual mathematical reasoning. In 1935, Gentzen [41] introduced the proof systems of the *sequent calculus* and *natural deduction* for propositional and first-order classical and intuitionistic logic, which were designed so that their inference rules more closely reflected the structure of reasoning in proofs as developed in mathematical practice. As a consequence of this design, the new proof systems were able to express normal forms for proofs which were not observable in Hilbert-Frege systems, providing a better setting in which to systematically investigate the normalisation theory of proofs. The sequent calculus expresses logical judgements in the form of *sequents*

$$A_1, \dots, A_n \vdash B_1, \dots, B_m$$

where  $A_1, \dots, A_n, B_1, \dots, B_m$  are formulae, which express that if every  $A_i$  is true for  $i \in \{1, \dots, n\}$  then at least one  $B_j$  is true for  $j \in \{1, \dots, m\}$  (i.e.,  $\overline{A_1} \vee \dots \vee \overline{A_n} \vee B_1 \vee \dots \vee B_m$  is a valid formula). Proofs in the sequent calculus are comprised of inference rules which contain a non-negative integer number of sequent premises and a single sequent conclusion. Inference rules are composed to form proofs so that the resultant sequent-calculus proofs exhibit a tree-like structure, with inference rules as vertices and formulae as edges.

Gentzen proved a fundamental normalisation result for the sequent calculus in the form of the *Hauptsatz*, or cut-elimination theorem, which states that the cut rule, which is given in the sequent calculus by

$$\frac{\Gamma \vdash \Delta, A \quad A, \Sigma \vdash \Pi}{\Gamma, \Sigma \vdash \Delta, \Pi} \text{Cut}$$

where  $\Gamma, \Delta, \Sigma$  and  $\Pi$  are sequences of formulae and  $A$  is a formula, is eliminable from sequent-calculus proofs. Observe that the cut rule may be understood as being equivalent to modus ponens due to the classical logical equivalence  $A \rightarrow B \equiv \overline{A} \vee B$ . The cut rule is the only inference rule in the sequent calculus which is not *analytic*, meaning it is the only inference rule which may contain symbols in its premise which are not contained in its conclusion. Consequently, cut-free proofs in the sequent calculus admit the *subformula property*: the property that every formula which occurs in the proof is a subformula of the conclusion of the proof. Consistency of the sequent calculus is an immediate corollary of the

subformula property, since it follows that a cut-free proof of the false unit would necessarily assume the false unit as a premise. Gentzen would also later employ the cut-elimination theorem to provide the first proof of the consistency of Peano arithmetic relative to the axioms of ZFC set theory [40].

Natural deduction is a restriction of the sequent calculus so that judgements are of the form

$$A_1, \dots, A_n \vdash B$$

where  $A_1, \dots, A_n$  and  $B$  are formulae. Prawitz [76] would later extend the *Hauptsatz* and other normalisation results of the sequent calculus to natural deduction in 1965. The sequent calculus and natural deduction have provided the setting in which most contemporary research in proof theory is conducted. Several adaptations of these formalisms have been developed in order to demonstrate cut elimination for various logics which do not admit cut elimination in the standard sequent calculus, such as *labelled sequent calculi* [59], *hypersequent calculi* [10, 11] and *nested sequent calculi* [27].

This thesis explores the proof theory of *deep inference* [45], which is a relatively new design methodology for proof formalisms which breaks from the conventions of Gentzen proof theory by allowing inference rules to apply at arbitrary depth inside of formulae. The greater freedom in applying inference rules in deep-inference formalisms enables the expression of new normal forms for proofs which are not observable in Gentzen systems as well as for speedups over conventional proof systems. The central proof system of this thesis, the falsifier calculus, is defined in the deep-inference formalism of *open deduction* [47] and exploits the properties of deep-inference proof theory to provide novel insights into the complexity and normalisation theory of proofs in first-order predicate logic. The proof theory of deep inference is overviewed in more detail in Subsection 1.5 below.

This thesis contributes to proof theory in two main ways: by providing a new normalisation result for proofs in first-order predicate logic and a new perspective on the phenomenon of the non-elementary compression of cut-free proofs. The falsifier calculus admits a new decomposition theorem for first-order proofs, Theorem 4.13 the Falsifier Decomposition Theorem, which is an analogue of Herbrand’s Theorem in which  $\varepsilon$ -terms may appear in the uppermost segment of the proof alongside its propositional content. The Falsifier Decomposition Theorem induces the notion of falsifier disjunctions, which are analogues to Herbrand disjunctions in the language of the epsilon-calculus such that certain classes of first-order theorems admit non-elementarily smaller falsifier disjunctions than Herbrand disjunctions. This yields a novel insight into the structure and complexity of Herbrand disjunctions and first-order proofs more broadly. The fact that certain proof systems admit non-elementarily smaller cut-free proofs than Gentzen’s sequent calculus has long been of interest in the study of first-order proof complexity [2, 13, 17] and the Falsifier Decomposition Theorem provides a new perspective on the properties that make a proof system admit this compression and the role of  $\varepsilon$ -terms in this compression.

## 1.2 Predicate Logic

In his *Begriffsschrift* [39], in an effort to develop a formal language in which to express and reason about arithmetical truths, Frege introduced the first formalisation of first-order predicate logic. Remarkably, the syntax of his predicate logic has seen widespread application and utility in a range of disciplines through to the present day, seeing only stylistic and presentational adaptations since its publication in 1879. The syntax of first-order predicate logic extends the language of propositional classical logic by the universal quantifier  $\forall$  and the existential quantifier  $\exists$ , which are respectively read “for all” and “there exists”, so that predicate logic allows for the expression of reasoning over elements of a domain. The first completeness proof for first-order predicate logic appeared in Gödel’s 1929 thesis [49] and the first formal semantics were provided by Tarski in 1931 [84].

In this thesis, the epsilon-calculus is employed as an alternate syntax for first-order predicate logic with the motivation of expressing normalisation and complexity results for



first-order proofs which are not observable in the standard syntax of predicate logic. Other alternate syntaxes for predicate logic include those which allow for greater expressivity of the dependence relations between quantifiers than the standard syntax, such as Henkin quantification [53] and independence-friendly logic [58], and those which compare the cardinalities of the subsets of the domain which satisfy certain predicates, such as the Hartig quantifier [55] and the Frege quantifier [7]. There is growing interest in investigating the proof theory of such syntaxes [15], placing this work in the broader space of understanding the proof theory of alternative formulations of predicate logic.

### 1.3 Herbrand’s Theorem

*Herbrand’s Theorem* [54] is one of the central theorems of first-order proof theory. It essentially provides a means of reducing provability in first-order predicate logic to provability in propositional classical logic by means of eliminating the existential quantifiers from a first-order formula. The theorem was first proved in Herbrand’s thesis in 1930 for arbitrary first-order formulae, however many statements and restricted forms of the theorem can be found in the literature. Following an overview due to Buss [30, 31] (with a correction due to McKinley [68]), the general version of Herbrand’s Theorem for arbitrary formulae in first-order predicate logic may be stated as follows.

**Theorem 1.1** (Herbrand’s Theorem). *A formula  $A$  is valid if and only if it admits a Herbrand proof, which consists of:*

- *A strong  $\vee$ -expansion  $A_\vee$  of  $A$ , which is a formula obtained by repeatedly replacing subformulae of the form  $\exists xB$  in  $A$  with  $\exists xB \vee \exists xB$*
- *A prenexation  $A^*$  of  $A_\vee$ , which is a formula obtained by renaming quantifiers and bound variables in  $A_\vee$  so that no two quantifiers bind the same variable and then using prenex operations to transform the resultant formula into prenex normal form*
- *A witnessing substitution  $\sigma$  for  $A^*$ , which is a sequence of terms  $t_1, \dots, t_n$  where  $\exists x_1, \dots, \exists x_n$  are all of the existential quantifiers in  $A^*$ , such that  $A_{qf}^*[t_1/x_1] \dots [t_n/x_n]$  is a propositional tautology, where  $A_{qf}^*$  is the formula obtained by deleting every existential and universal quantifier in  $A^*$  and  $B[t/x]$  denotes the formula obtained by replacing every free occurrence of the variable  $x$  in the formula  $B$  with the term  $t$*

Such statements of Herbrand’s Theorem are usually proved in the sequent calculus as a corollary of cut elimination. In “Buss-style” proofs of Herbrand’s Theorem, cut rules are first eliminated from a sequent-calculus proof to make the proof amenable to rule permutations. The cut-free proof is then appropriately modified, permuting existential contraction rules and existential witnessing rules down the proof, to yield a propositional proof of the desired propositional tautology. Herbrand’s Theorem does not admit a direct statement in the sequent calculus, instead requiring statements in the form of metatheorems involving formulae which are not directly included in any formal proofs, such as the one given above. However, a direct statement and proof of Herbrand’s Theorem in a deep-inference formalism has been provided by Brunnler [21] in the form of a decomposition theorem, which is given as follows.

**Theorem 1.2** (Herbrand’s Theorem in Deep Inference). *Every valid formula  $A$  has a proof*

of the form

$$\begin{array}{c}
\forall x_1 \dots \forall x_n \boxed{\begin{array}{c} \prod \text{Propositional rules} \\ A' \end{array}} \\
\parallel \{\exists\} \\
A'' \\
\parallel \{r1\downarrow, r2\downarrow, r3\downarrow, r4\downarrow\} \\
A''' \\
\parallel \{qc\downarrow\} \\
A
\end{array}$$

which is said to be in Herbrand normal form, where the formula  $\forall x_1 \dots \forall x_n A'$  is called a Herbrand disjunction for  $A$ , the formula  $A''$  is in prenex normal form and the relevant inference rules are the existential witnessing rule  $\exists$ , given by

$$\exists \frac{B(t)}{\exists x B(x)}$$

where  $t$  is free for  $x$  in  $B(x)$ , the quantifier-shift rules  $r1\downarrow, r2\downarrow, r3\downarrow, r4\downarrow$ , given by

$$\begin{array}{cccc}
r1\downarrow \frac{\forall x(B(x) \vee C)}{\forall x B(x) \vee C} & r2\downarrow \frac{\forall x(B(x) \wedge C)}{\forall x B(x) \wedge C} & r3\downarrow \frac{\exists x(B(x) \vee C)}{\exists x B(x) \vee C} & r4\downarrow \frac{\exists x(B(x) \wedge C)}{\exists x B(x) \wedge C}
\end{array}$$

where  $x$  does not occur free in  $C$ , and the existential contraction rule  $qc\downarrow$ , given by

$$qc\downarrow \frac{\exists x B \vee \exists x B}{\exists x B}$$

Observe that this resembles the version of the theorem given by Theorem 1.1 above, in that the decomposed proof of  $A$  in Herbrand normal form is similar to a Herbrand proof of  $A$ . However, the proof in Herbrand normal form may be expressed directly in a deep-inference formalism, unlike the Herbrand proof of  $A$  which requires an ad hoc definition which is not expressed directly in any formalism. Furthermore, the procedure for transforming a first-order deep-inference proof into Herbrand normal form provided by Br nnler does not require all cut rules to be eliminated from the proof, establishing a kind of independence between Herbrand’s Theorem and cut elimination which is not observable in the sequent calculus.

Since first-order predicate logic is undecidable and propositional classical logic is decidable, Herbrand’s Theorem may be understood to isolate the undecidable content contained in a first-order proof in the existential rules in the lower segments of a proof in Herbrand normal form. In particular, the existential contraction rule  $qc\downarrow$ , given by

$$qc\downarrow \frac{\exists x B \vee \exists x B}{\exists x B}$$

may be understood as a case distinction on the witnesses to the two existential quantifiers in the premise of the rule, in that the existential quantifier in the conclusion of the rule is witnessed by one of the witnesses to the two existential quantifiers in the premise of the rule. As such, we may understand the undecidable content of a first-order proof as lying in the case distinctions contained in instances of the existential contraction rule  $qc\downarrow$ , which are isolated in the lowermost segment of a proof in Herbrand normal form.

A simple yet elucidating example of Herbrand’s Theorem is provided by the “drinker’s paradox”, which is the formula  $\exists y \forall x (P(x) \vee \bar{P}(y))$  (or, expressed in a logically equivalent form using an implication symbol,  $\exists y \forall x (P(y) \rightarrow P(x))$ ) and expresses that in any pub, there exists a person such that if they are drinking then everyone in the pub is drinking. The drinker’s paradox is true by the following logical argument:

There are two cases: either (1) everyone in the pub is drinking,  $\forall xP(x)$ , or (2) there exists a person  $y$  in the pub who is not drinking,  $\exists y\bar{P}(y)$ . In case (1), since everyone in the pub is drinking, we may assign an arbitrary witness to the existential quantifier  $\exists y$  in the drinker's paradox to obtain  $\exists y\forall x(P(x) \vee \bar{P}(y))$  as desired. In case (2), since there is a person  $y$  in the pub who is not drinking, it follows by the principle of explosion that if  $y$  is drinking then everyone in the pub is drinking so that we have  $\exists y\forall x(P(x) \vee \bar{P}(y))$  as desired.

A Herbrand disjunction for the formula is  $\forall x_1\forall x_2[(P(x_1) \vee \bar{P}(c)) \vee (P(x_2) \vee \bar{P}(x_1))]$ , which reflects the two cases in this argument in the two disjuncts. The universally quantified variable  $x_1$  is used to perform a case distinction on  $P(x_1)$ , so that the first disjunct  $P(x_1) \vee \bar{P}(c)$  represents case (1), that everyone in the pub is drinking ( $\forall x_1P(x_1)$ ) and the second disjunct  $P(x_2) \vee \bar{P}(x_1)$  represents case (2), that there exists a person in the pub who is not drinking ( $\exists y\bar{P}(y)$ , where the existential quantifier  $\exists y$  is witnessed by  $x_1$ ). A deep-inference proof in Herbrand normal form for the formula using this Herbrand disjunction is given by

$$\begin{array}{c}
\forall x_1\forall x_2 \quad = \quad \boxed{\boxed{\boxed{\boxed{\boxed{\frac{\text{t}}{P(x_1) \vee \bar{P}(x_1)}}{\text{i}\downarrow}}}}}} \\
\boxed{\boxed{\boxed{\boxed{P(x_1) \vee \boxed{\boxed{\frac{\text{f}}{\bar{P}(c)}}{\text{w}\downarrow}}}} \vee \boxed{\boxed{\frac{\text{f}}{P(x_2)}} \vee \bar{P}(x_1)}}}} \\
\parallel \{\exists\} \\
\exists y_1\forall x_1\exists y_2\forall x_2(P(x_1) \vee \bar{P}(y_1) \vee P(x_2) \vee \bar{P}(y_2)) \\
\parallel \{\text{r1}\downarrow, \text{r3}\downarrow\} \\
\text{qc}\downarrow \frac{\exists y_1(\forall x_1P(x_1) \vee \bar{P}(y_1)) \vee \exists y_2(\forall x_2P(x_2) \vee \bar{P}(y_2))}{\exists y(\forall xP(x) \vee \bar{P}(y))}
\end{array}$$

Observe that, as given by Theorem 1.2 above, the upper segment of the proof is a propositional proof of the Herbrand disjunction  $\forall x_1\forall x_2[(P(x_1) \vee \bar{P}(c)) \vee (P(x_2) \vee \bar{P}(x_1))]$  and is followed by the lower first-order segments of the proof, which contract the two cases contained in the Herbrand disjunction into a single existential quantifier.

In general, there is no recursive bound on the size of a Herbrand disjunction for a first-order formula in terms of the size of the formula [30]. However, as a consequence of the ‘‘Buss-style’’ procedure for extracting Herbrand disjunctions from sequent-calculus proofs, there is a non-elementary function bounding the size of a Herbrand disjunction for a first-order formula in terms of the size of a given sequent-calculus proof for the formula [30]. This bound is essentially sharp, in that Statman [81] has demonstrated that there is no elementary function bounding the size of a Herbrand disjunction for a first-order formula in terms of the size of a sequent-calculus proof of the formula. This is related to recent work studying the non-elementary compression of cut-free proofs in first-order predicate logic [2, 13, 17], which will also be one of the central focuses of this thesis.

Herbrand’s Theorem provides the foundation for Miller’s *expansion proofs* [71], which is a graphical proof system for first-order predicate logic in which each witness to an existential quantifier is represented as an edge connecting the existential quantifier to a disjunct in a Herbrand disjunction. Ralph [78, 79] has provided translations between expansion proofs and deep-inference proofs in Herbrand normal form, demonstrating the equivalence of the two systems. Various extensions of the expansion proofs formalism have been developed which introduce cut rules into the formalism and provide a notion of cut elimination for expansion proofs, such as Heijltjes’ *proof forestry* [51, 52] and an extension of expansion proofs due to Aschieri, Hetzl and Weller [9]. Other examples of proof systems and approaches to first-order proof theory which are based on Herbrand’s Theorem include McKinley’s

*Herbrand nets* [69] and a recent game-semantic approach due to Alcolei, Clairambault, Hyland and Winskel [3].

The proof system for first-order predicate introduced in this thesis, the falsifier calculus, provides a new decomposition theorem for first-order proofs which may be understood as a non-constructive analogue of Herbrand's Theorem. The new decomposition theorem gives rise to the notion of *falsifier disjunctions*, which are analogues to Herbrand disjunctions in the language of the epsilon-calculus such that certain classes of theorems admit non-elementarily smaller falsifier disjunctions than Herbrand disjunctions. The  $\varepsilon$ -terms in a falsifier disjunction represent elements which are drawn from the domain non-constructively, resulting in a compression over Herbrand disjunctions, which are comprised only of ground terms so that they represent the content of a first-order proof in a purely constructive manner.

## 1.4 The Epsilon-Calculus

The *epsilon-calculus* [57, 87, 88] was introduced as part of Hilbert's program in the early twentieth century, which sought to provide a formal foundation for mathematics and establish its consistency, and is one of the oldest proof-theoretic formalisms for predicate logic. The epsilon-calculus extends the language of predicate logic by  $\varepsilon$ -terms  $\varepsilon_x A$  for all variables  $x$  and formulae  $A$  such that, for a given semantics  $\llbracket - \rrbracket_{\mathbb{D}}$  with domain  $\mathbb{D}$ , each  $\varepsilon$ -term  $\varepsilon_x A(x)$  is assigned an element in  $\mathbb{D}$  by

$$\llbracket \varepsilon_x A(x) \rrbracket_{\mathbb{D}} = \begin{cases} d & \text{if there exists some } d \in \mathbb{D} \text{ such that } \llbracket A(d) \rrbracket_{\mathbb{D}} \\ a & \text{for some arbitrary } a \in \mathbb{D}, \text{ otherwise} \end{cases}$$

where  $d$  is chosen by a choice function on  $\mathcal{P}(\mathbb{D})$  and  $a \in \mathbb{D}$  is fixed. As such, the epsilon-calculus provides a means of incorporating a choice operator into the syntax of predicate logic. In the epsilon-calculus, quantifiers in a proof may be encoded by  $\varepsilon$ -terms using the logical equivalences  $\exists x A(x) \equiv A(\varepsilon_x A(x))$  and  $\forall x A(x) \equiv A(\varepsilon_x \bar{A}(x))$  so that all quantifier information in a first-order proof may be expressed within  $\varepsilon$ -terms.

In the traditional epsilon-calculus as introduced by Hilbert,  $\varepsilon$ -terms are introduced into a proof by *critical axioms*, which are inferences of the form

$$A(t) \implies A(\varepsilon_x A(x))$$

and, in the presence of the encodings of quantifiers by  $\varepsilon$ -terms, form a complete proof system for first-order predicate logic when incorporated into a proof system for propositional classical logic [89]. The central theorems of the proof theory of the epsilon-calculus are the *epsilon theorems*, which are conservativity results for the epsilon-calculus and provide a means of eliminating  $\varepsilon$ -terms from proofs. The *first epsilon theorem* and *second epsilon theorem* (see [75]) respectively establish the conservativity of extending propositional classical logic and first-order predicate logic by critical axioms. The *extended first epsilon theorem* [14, 75] further establishes that for any quantifier-free, epsilon-free formula  $A(x_1, \dots, x_n)$  and  $\varepsilon$ -terms  $\varepsilon_{x_1} B_1, \dots, \varepsilon_{x_n} B_n$  which are respectively free for  $x_1, \dots, x_n$  in  $A(x_1, \dots, x_n)$ , if  $A(\varepsilon_{x_1} B_1, \dots, \varepsilon_{x_n} B_n)$  is provable in the epsilon-calculus, then there exist epsilon-free terms  $t_j^i$  for  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, m\}$  such that

$$\bigvee_{i=1}^{i=n} A(t_1^i, \dots, t_m^i)$$

is a propositional tautology. Herbrand's Theorem for existential formulae is an immediate corollary of this since existential quantifiers may be encoded by  $\varepsilon$ -terms using the logical equivalence  $\exists x A(x) \equiv A(\varepsilon_x A(x))$ . These theorems are traditionally proved using a procedure known as *epsilon substitution*, which recursively eliminates critical axioms from a proof. The epsilon theorems and epsilon substitution are usually stated and proved for the

equality-free case, though a treatment of the case with equality has recently been provided by Miyamoto and Moser [73]. Hetzl [56] has also recently provided a novel proof of the epsilon theorems using expansion proofs.

Aside from its historical interest, the epsilon-calculus has seen a renewal of interest in recent years [8, 14, 17, 72, 73, 75], in part due to its complexity advantages over traditional proof systems: the sequent-calculus-based epsilon-calculus admits non-elementarily smaller cut-free proofs of certain classes of theorems than traditional sequent-calculus systems for first-order predicate logic [17]. This is related to the design of a recently introduced sequent-calculus proof system for first-order predicate logic [2] which admits this non-elementary compression for cut-free proofs over the standard sequent calculus by the use of inference rules which are locally unsound, but which admits a global correctness condition which is inspired by the epsilon-calculus. The methodology for designing global correctness conditions based on the epsilon-calculus for proof systems which include locally unsound rules has also motivated the definition an analytic sequent-calculus proof system for “reasonable” fragments of Henkin quantification [16], which do not admit locally sound sequent-calculus proof systems, as well as to demonstrate that an intermediate logic admits Skolemisation if and only if admits a certain class of rules known as *quantifier-shifts*, which will be discussed later in this thesis (see Subsection 3.3) [12].

Despite its complexity benefits and interesting proof-theoretic properties, the epsilon-calculus has received little attention from the wider proof theory community or from the designers of automated and interactive theorem provers, largely due to its unwieldy nature. In particular, the encodings of quantifiers by  $\varepsilon$ -terms can result in cumbersome encoded formulae in the epsilon-calculus. For instance, borrowing an example from [14], the translation of the simple first-order formula  $\exists x \exists y \exists z A(x, y, z)$  into the epsilon-calculus is given by

$$\begin{aligned} & A(\varepsilon_x A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \varepsilon_y A(\varepsilon_x A(x, \\ & \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), y, \varepsilon_z A(\varepsilon_x A(x, \varepsilon_y A(x, y, \\ & \varepsilon_z A(x, y, z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), y, z)), \varepsilon_z A(\varepsilon_x A(x, \varepsilon_y A(x, y, \\ & \varepsilon_z A(x, y, z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \varepsilon_y A(\varepsilon_x A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, \\ & z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), y, \varepsilon_z A(\varepsilon_x A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \\ & \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), z))y, z)), z)) \end{aligned}$$

Some solutions to this problem have been suggested, such as an alternative syntax for the epsilon-calculus using Skolem functions due to Davis and Fechter [36, 38] and the introduction of *function variable proofs* by Baaz, Leitsch and Lolić [14].

This thesis contributes to the proof theory of the epsilon-calculus by providing a new proof system for first-order predicate logic in the language of the epsilon-calculus and demonstrating several normalisation and complexity results for this system. One of the aims of this work is to provide a more accessible treatment of the epsilon-calculus, guided by modern motivations of complexity and normalisation in place of the foundational concerns for completeness and consistency when the epsilon-calculus was originally introduced. The system and results in this thesis also do not use the cumbersome encodings of quantifiers by  $\varepsilon$ -terms described above, circumventing one of the most notorious shortfalls of the traditional epsilon-calculus.

## 1.5 Deep Inference

*Deep inference* is a design methodology for proof systems which breaks from the conventions of traditional Gentzen proof theory by allowing inference rules to apply at arbitrary depth inside of formulae. This allows for a more flexible composition mechanism for composing proofs and derivations than in traditional proof systems such as the sequent calculus and natural deduction, giving rise to a new proof theory within the more expressive formalism. The freedom to apply inference rules inside formula contexts yields benefits in proof normalisation, such as more freedom in permuting inference rules around a proof and a finer

granularity of inference rules than is observable in traditional proof systems, in that inference rules on arbitrary formulae may be decomposed into derivations of smaller inference rules on atomic formulae and quantifiers. This thesis introduces a new deep-inference proof system for first-order predicate logic, the falsifier calculus, and proves several normalisation results for this system, exploiting these properties of deep-inference proof theory in doing so.

The proof theory of deep inference was first investigated when it was employed by Guglielmi to develop a formalism which could express BV [45], an extension of multiplicative linear logic by a self-adjoint non-commutative binary connective, which is not expressible in the sequent calculus [86]. A proliferation of deep-inference proof systems for other more standard logics followed, such as for propositional classical logic [29], first-order predicate logic [28], linear logic [83], intuitionistic logic [85] and modal logics [82]. Early research in deep inference primarily used the deep-inference proof formalism of *the calculus of structures* [45], which represents proofs and derivations as sequences of formulae connected vertically by inference rules which act at arbitrary depth inside formulae, so that a total vertical ordering of inference rules exists for every derivation. Contemporary research generally prefers to use the deep-inference formalism of *open deduction* [47], which is a generalisation of the calculus of structures and centers its definition around compositionality, so that derivations are constructed inductively by composing derivations by connectives, by quantifiers and by inference rules.

Early investigations into the proof theory of deep inference were concerned primarily with reestablishing the central theorems of Gentzen proof theory within the new methodology. In proof systems of the Gentzen tradition, inference rules act on the outermost connectives of formulae, a property which is lost in deep inference due to the greater freedom in applying inference rules. This meant that new techniques had to be developed within the deep-inference methodology to prove standard proof-theoretic results such as cut elimination, Herbrand’s Theorem and Craig interpolation. The technique of *splitting* [5, 21] was developed to demonstrate cut elimination for various logics in deep inference, which splits the proof above a cut rule in a deep-inference proof into two separate proofs composed by a conjunction, essentially locally replicating the tree-like structure of the sequent calculus within a deep-inference formalism to make proofs amenable to cut elimination.

The greater freedom in applying inference rules in deep-inference proof systems as compared to traditional Gentzen proof systems is also a source of complexity difference between the two methodologies. As such, much work has been done to demonstrate speedup results for deep-inference proof systems over conventional proof systems, such as the demonstration of exponential speedups for cut-free proofs of certain tautologies in propositional classical logic [24], the existence of a quasipolynomial-complexity cut-elimination procedure for propositional classical logic due to Jeřábek [25, 63] (in contrast to such procedures in the sequent calculus, which are inherently of exponential complexity) and the existence of quasipolynomial-size cut-free proofs of the pigeonhole principle [35] (which are of at least exponential size in the sequent calculus). The complexity of deep-inference proofs for propositional classical logic was the subject of Das’ thesis [34]. Like the epsilon-calculus, deep-inference proof systems for first-order predicate logic also admit non-elementarily smaller cut-free proofs of certain classes of theorems than the sequent calculus. Despite having been known for some time, this fact will be formally presented for the first time in this thesis (see Proposition 3.9). A discussion on the non-elementary compression of cut-free proofs in deep inference may also be found in Ralph’s thesis [78]. The deep-inference proof system for first-order predicate logic introduced in this thesis admits this non-elementary compression for cut-free proofs, providing a new perspective on this compression through a novel approach to the epsilon-calculus.

The complexity difference between deep-inference proof systems and traditional proof systems may be seen as resulting from the greater expressivity of deep-inference formalisms: since the formalism is able to express more proofs, smaller proofs become possible within the new methodology. As such, deep inference is also part of a broader research effort towards



abstracting away from the inessential syntactic bureaucracy of conventional proof systems like the sequent calculus, attempting to capture the “essence” of proofs that is independent of any particular proof formalism. Other notable research programs with this goal include Girard’s *proof nets* [43] and Hughes’ *combinatorial proofs* [60, 61, 62].

A fundamental normalisation mechanism in deep-inference proof theory is the existence of *decomposition theorems* for various logics. The liberal composition mechanism of deep-inference formalisms allows for great freedom in permuting inference rules around proofs and derivations, so that they may be factorised into a vertical composition of independent derivations which are comprised only of certain inference rules, of the form

$$\begin{array}{c}
A_1 \\
\parallel \\
\{\rho_1, \dots, \rho_{k_1}\} \\
A_2 \\
\parallel \\
\vdots \\
\parallel \\
A_{n-1} \\
\parallel \\
\{\rho_{k_{n-2}+1}, \dots, \rho_{k_{n-1}}\} \\
A_n
\end{array}$$

The corresponding normal forms are called decomposition theorems, and may be used to express many proof-theoretic phenomena and provide insight into the shape and structure of proofs in various logics. Decomposition theorems have been proved for many logics, including propositional classical logic [29], first-order predicate logic [20, 62] and linear logic [83]. In particular, several of the central theorems of classical proof theory may be expressed as decomposition theorems in deep inference, such as the general version of Herbrand’s Theorem [21] and Craig interpolation [26]. In traditional proof theory, such theorems are usually proved as corollaries of cut elimination, requiring a procedure which substantially alters the structure of a proof in doing so. Their statements and proofs as decomposition theorems in deep inference, however, do not require cuts to be fully eliminated from a proof and instead reveal a kind of independence between these theorems and cut elimination which is not observable in traditional formalisms, providing new perspectives on the relationships between these theorems. The central theorem of this thesis is a new decomposition theorem for first-order predicate logic using the newly introduced falsifier calculus, providing a new perspective on the structure of first-order proofs.

*Subatomic logic* [4, 5, 19] is a recent development in deep-inference proof theory which provides several novel means of normalising proofs using an alternative syntax for various logics. In subatomic logic, atoms are represented as non-commutative binary connectives which represent a superposition of truth values (e.g., in propositional classical logic, the formula  $0 \mathbf{a} 1$  is interpreted as 0 when  $\mathbf{a}$  is “false” and as 1 when  $\mathbf{a}$  is “true” so that the formula is semantically equivalent to a propositional atom in the standard syntax). By adopting this syntax, it is possible to form complete deep-inference proof systems for several logics, including propositional classical logic, first-order predicate logic and linear logics, whose inference rules are all generated by a single rule shape called the *subatomic rule shape* [4, 5]. This allows for a uniform treatment of proof normalisation for several logics in a logic-agnostic setting, as well as a semantically natural cut-elimination procedure for propositional classical logic based on the projection of the atomic connectives in a proof [19]. The work in this thesis may be seen as sharing some of the spirit of subatomic logic, in that the epsilon-calculus is employed as an alternative syntax for first-order predicate logic with the motivation of proof normalisation, rather than the usual focus on expressivity when designing a language.

Deep inference has also inspired the design of several other proof systems, including nested sequent calculi [27] and combinatorial proofs [60, 61]. Nested sequent calculi incorporate deep inference into the sequent calculus, enabling the definition of complete proof

systems for modal logics which do not admit complete sequent-calculus systems. Combinatorial proofs provide a graphical representation of proofs in various logics which resemble deep-inference decomposition theorems, in that the contraction-weakening segment of a combinatorial proof is contained in a “skew fibration” which occurs below the rest of the proof. Another notable proof formalism which resembles deep inference is Belnap’s *display calculus* [67].

This thesis contributes to deep-inference proof theory by introducing a new deep-inference proof system for first-order predicate logic in the language of the epsilon-calculus and providing a new decomposition theorem for first-order proofs using this system. This system also admits non-elementarily smaller cut-free proofs of certain classes of theorems than the sequent calculus, providing a new perspective on this compression. Although most research in deep inference has focused on propositional logics, a strong foundation has been laid for first-order logic by Brünnler, Ralph et al. [20, 21, 22, 26, 77, 78, 79]. The work in this thesis may in a certain sense be seen as a continuation of these efforts. There are, however, several other directions in which to study first-order proof theory using deep inference, particularly from the perspective of more recent developments in deep-inference proof theory such as subatomic logic discussed above.

## 1.6 Overview of Thesis

At the technical level, the main contributions of this thesis are the introduction a new deep-inference proof system for first-order predicate logic, the falsifier calculus, in the language of the epsilon-calculus and a decomposition theorem for first-order predicate logic using this system. More broadly, the aims of this research are to better understand the phenomenon of the non-elementary compression of cut-free proofs and to investigate the shape and structure of first-order proofs using the methodology of deep inference. The falsifier calculus also aims to provide a simpler and more accessible proof system for the epsilon-calculus than the traditional Hilbert-Frege and sequent-calculus systems. The central technical concepts of this thesis have already appeared in [6], of which the present work may be considered an elaboration.

The sections of the thesis are structured as follows:

- Section 2, *Preliminaries*, provides definitions and preliminary results which serve as a technical foundation for the rest of the thesis. In particular, the language of first-order predicate logic and the epsilon-calculus are introduced along with a semantics for this language. Open deduction, the primary proof formalism used to define proof systems in this thesis, is also introduced.
- Section 3, *The Structure of First-Order Proofs*, introduces several open-deduction proof systems for first-order predicate logic and uses them to explore some of the properties of first-order proofs. In particular, two of the central topics of this thesis are introduced and explored: the non-elementary compression of cut-free proofs and the extraction of the case analyses contained in quantifier contraction rules from proofs.
- Section 4, *The Falsifier Calculus*, introduces the central proof system of this thesis, the falsifier calculus, and explores some of its properties. The falsifier calculus is used to state the main result of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem, a decomposition theorem for proofs in first-order predicate logic.
- Section 5, *Proof of The Falsifier Decomposition Theorem*, provides a proof of Theorem 4.13 the Falsifier Decomposition Theorem by formalising a procedure for extracting the case analyses contained in quantifier contraction rules from proofs with elementary complexity.
- Section 6, *The Restricted Falsifier Calculus*, introduces the restricted falsifier calculus, which is a restriction of the falsifier calculus that replaces the usual universal



instantiation rule with a trivial variant that may only be instantiated by variables. The restricted falsifier calculus is used to state and prove an analogue of the Falsifier Decomposition Theorem, demonstrating that only the trivial universal instantiation rule is required to attain the desirable complexity and normalisation properties of the falsifier calculus.

- Section 7, *Conclusion*, briefly overviews the main results of the thesis and suggests some directions for future research.

## 2 Preliminaries

All language is but a poor translation.

*F. Kafka*

In this section I will provide definitions and preliminary results which will serve as the technical foundation for the rest of the thesis. In Subsection 2.1 I define the language of first-order predicate logic and the epsilon-calculus. In Subsection 2.2 I provide a semantics for this language. In Subsection 2.3 I define the deep-inference proof formalism of *open deduction* [47], which will serve as the primary formalism for defining proof systems in this thesis.

### 2.1 Language

I begin by recalling the standard definitions of the language of equality-free first-order predicate logic and the epsilon-calculus.

The syntax defined here for the epsilon-calculus is essentially the same as that originally used by Hilbert [57] when the epsilon-calculus was first introduced under the name of “the logical  $\epsilon$ -function”. For a given semantics  $\llbracket - \rrbracket_{\mathbb{D}}$  with domain  $\mathbb{D}$ , each  $\epsilon$ -term  $\epsilon_x A(x)$  in the language is assigned an element in  $\mathbb{D}$  by

$$\llbracket \epsilon_x A(x) \rrbracket_{\mathbb{D}} = \begin{cases} d & \text{if there exists some } d \in \mathbb{D} \text{ such that } \llbracket A(d) \rrbracket_{\mathbb{D}} \\ a & \text{for some arbitrary } a \in \mathbb{D}, \text{ otherwise} \end{cases}$$

where  $d$  is chosen by a choice function on  $\mathcal{P}(\mathbb{D})$  and  $a \in \mathbb{D}$  is fixed. A formal semantics for this language and  $\epsilon$ -terms will be provided in the following subsection, Subsection 2.2.

The decision to not include equality in the language follows the conventions of existing research into the deep-inference proof theory of first-order predicate logic [20, 21, 78], for which the case with equality has yet to be thoroughly investigated.

**Definition 2.1.** Fix three disjoint countably infinite sets of symbols  $\mathbf{Var} = \{x, y, z, \dots\}$ ,  $\mathbf{Fun} = \{f, g, h, \dots\}$ ,  $\mathbf{Pred} = \{P, Q, R, \dots\}$ , whose respective elements are called *variable symbols*, *function symbols* and *predicate symbols*, where every function symbol and predicate symbol has an associated non-negative integer arity and every predicate symbol  $P$  has a corresponding *dual* predicate symbol  $\bar{P}$  of the same arity such that  $\bar{\bar{P}} = P$  and  $\bar{P} \neq P$ .

I define *terms*  $t$  and *formulae*  $A$  by the following grammars by mutual recursion:

$$\begin{aligned} t &::= x \mid f(t, \dots, t) \mid \epsilon_x A \\ A &::= \mathbf{t} \mid \mathbf{f} \mid P(t, \dots, t) \mid A \vee A \mid A \wedge A \mid \exists x A \mid \forall x A \end{aligned}$$

where  $x$  is a variable symbol, called a *variable*,  $f$  is a function symbol of arity  $n$ , each  $f(t_1, \dots, t_n)$  is called a *function term*, function terms of arity 0 are called *constant terms*,  $\epsilon_x A$  is called an  $\epsilon$ -term,  $\mathbf{t}$  (true) and  $\mathbf{f}$  (false) are called *units*,  $P$  is a predicate symbol of arity  $n$  and each  $P(t_1, \dots, t_n)$  is called an *atomic formula*.

**Remark 2.2.** In the traditional epsilon-calculus [17, 57, 74, 75], quantifiers are encoded by  $\epsilon$ -terms using the logical equivalences  $\exists x A(x) \equiv A(\epsilon_x A(x))$  and  $\forall x A(x) \equiv A(\epsilon_x \bar{A}(x))$  so that quantifiers need not be included as explicit symbols in the language. This can result in unwieldy translated formulae in the epsilon-calculus. For instance, borrowing an example from [14], the translation of the simple first-order formula  $\exists x \exists y \exists z A(x, y, z)$  under

this encoding is given by

$$\begin{aligned} & A(\varepsilon_x A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \varepsilon_y A(\varepsilon_x A(x, \\ & \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), y, \varepsilon_z A(\varepsilon_x A(x, \varepsilon_y A(x, y, \\ & \varepsilon_z A(x, y, z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), y, z), \varepsilon_z A(\varepsilon_x A(x, \varepsilon_y A(x, y, \\ & \varepsilon_z A(x, y, z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \varepsilon_y A(\varepsilon_x A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, \\ & z))), \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), y, \varepsilon_z A(\varepsilon_x A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z))), \\ & \varepsilon_z A(x, \varepsilon_y A(x, y, \varepsilon_z A(x, y, z)))y, z))), z)) \end{aligned}$$

Such encodings are not used in the proof systems and results of this thesis, so that these cumbersome encodings are avoided and existential quantifiers  $\exists x$  and universal quantifiers  $\forall x$  are included as explicit symbols in the language, alongside epsilon symbols  $\varepsilon_x$ .

The duals of formulae are defined using standard De Morgan duals.

**Definition 2.3.** The *dual* formula  $\bar{A}$  to a formula  $A$  is defined recursively as follows. For all formulae  $A$  and  $B$ , all atomic formulae  $P(t_1, \dots, t_n)$  and all variables  $x$ :  $\bar{\bar{t}} = t$ ,  $\bar{\bar{f}} = f$ ,  $\bar{P(t_1, \dots, t_n)} = \bar{P}(t_1, \dots, t_n)$ ,  $\overline{A \vee B} = \bar{A} \wedge \bar{B}$ ,  $\overline{A \wedge B} = \bar{A} \vee \bar{B}$ ,  $\overline{\exists x A} = \forall x \bar{A}$  and  $\overline{\forall x A} = \exists x \bar{A}$ .

**Remark 2.4.** As per convention when defining deep-inference systems, implication  $\rightarrow$  and negation  $\neg$  have not been included as explicit symbols in the language. This is because standard inference rules are unsound when applied inside the context of an odd number of negation symbols  $\neg$  or the left-hand side of implication symbols  $\rightarrow$  (the premise and conclusion of an inference rule  $\rho \frac{A}{B}$  must be reversed in such a context to yield a sound inference rule  $\neg \rho \frac{B}{A}$ ). Instead, implication may be encoded by a disjunction and formula dual by the logical equivalence  $A \rightarrow B \equiv \bar{A} \vee B$  and negation may be encoded simply as the dual of a formula by  $\neg A \equiv \bar{A}$ .

I introduce the following definitions to distinguish formulae in which quantifiers may occur inside, but not outside, the scope of  $\varepsilon$ -terms.

**Definition 2.5.** A formula  $A_{wqf}$  is said to be *weakly quantifier-free* if it is generated by the grammar

$$A_{wqf} ::= t \mid f \mid P(t_1, \dots, t_n) \mid A_{wqf} \vee A_{wqf} \mid A_{wqf} \wedge A_{wqf}$$

and a formula  $A_{wef}$  is said to be *weakly existential-free* if it is generated by the grammar

$$A_{wef} ::= t \mid f \mid P(t_1, \dots, t_n) \mid A_{wef} \vee A_{wef} \mid A_{wef} \wedge A_{wef} \mid \forall x A_{wef}$$

where  $P(t_1, \dots, t_n)$  is an atomic formula and  $x$  is a variable.

Note that the terms  $t_1, \dots, t_n$  in the definition above may contain  $\varepsilon$ -terms which contain quantifiers.

I now introduce some definitions which will be used when reasoning about terms.

A term is defined to *occur in* another term or formula in the expected way, as follows.

**Definition 2.6.** A term  $t$  is said to *occur in* a term or formula as follows:

- $t$  occurs in itself.
- If  $t$  occurs in a term  $s$ , then  $t$  occurs in all function terms of the form  $f(t_1, \dots, s, \dots, t_n)$  and all atomic formulae of the form  $P(t_1, \dots, s, \dots, t_n)$ .
- If  $t$  occurs in a formula  $A$ , then for all formulae  $B$  and all variables  $x$ ,  $t$  occurs in the formulae  $A \vee B$ ,  $B \vee A$ ,  $A \wedge B$ ,  $B \wedge A$ ,  $\forall x A$  and  $\exists x A$  and the term  $\varepsilon_x A$ .

I define the notion of a *free occurrence* of a variable in a term or formula in the standard way and recall a standard notation for formulae and terms.

**Definition 2.7.** An occurrence of a variable  $x$  in a term or formula is said to be a *free occurrence* if it does not occur inside the scope of any  $\exists x$ ,  $\forall x$  or  $\varepsilon_x$  symbols.

If there is a free occurrence of a variable  $x$  in a term or formula,  $x$  is said to *occur free* in that term or formula.

If a formula is denoted  $A(x_1, \dots, x_n)$  for variables  $x_1, \dots, x_n$ , I denote by  $A(t_1, \dots, t_n)$  the formula obtained by replacing every free occurrence of  $x_i$  in  $A(x_1, \dots, x_n)$  by the term  $t_i$ , for  $i \in \{1, \dots, n\}$ . Likewise, if a term is denoted  $t(x_1, \dots, x_n)$ , I denote by  $t(s_1, \dots, s_n)$  the term obtained by replacing every free occurrence of  $x_i$  in  $t(x_1, \dots, x_n)$  by the term  $s_i$ , for  $i \in \{1, \dots, n\}$ .

I further define the notion of a *bound occurrence* of a variable in a term or formula in the standard way.

**Definition 2.8.** An occurrence of a variable  $x$  in a term or formula is said to be a *bound occurrence* if it occurs inside the scope of a  $\exists x$ ,  $\forall x$  or  $\varepsilon_x$  symbol. Furthermore, in such a case, the innermost  $\exists x$ ,  $\forall x$  or  $\varepsilon_x$  symbol which contains the bound occurrence of  $x$  in its scope is said to *bind* the occurrence of  $x$ .

The following example illustrates the usage of the above two definitions.

**Example 2.9.** Let  $P(x, y)$  be an atomic formula, where  $x$  and  $y$  are variables, and consider the formula  $A$  given by  $P(x, y) \wedge P(\varepsilon_x \forall y P(x, y), \varepsilon_y P(x, y))$ . Observe that the occurrences of  $x$  and  $y$  in  $P(x, y)$  are free occurrences in  $A$  since they do not occur within the scope of any quantifiers or epsilon symbols. In the term  $\varepsilon_x \forall y P(x, y)$ , the occurrence of  $x$  is contained within the scope of the epsilon symbol  $\varepsilon_x$  and the occurrence of  $y$  is contained within the scope of the universal quantifier  $\forall y$  so that both of the occurrences of  $x$  and  $y$  are bound occurrences in  $A$ . In the term  $\varepsilon_y P(x, y)$ , the occurrence of  $x$  is not bound by any quantifiers or epsilon symbols and hence is a free occurrence in  $A$ , but the occurrence of  $y$  is contained within the scope of the epsilon symbol  $\varepsilon_y$  and hence is a bound occurrence.  $A$  therefore contains two free occurrences and one bound occurrence of  $x$ , and one free occurrence and two bound occurrences of  $y$ .

The following notation will be used when substituting one term into another.

**Definition 2.10.** For all terms  $t$  and  $s$  and variables  $x$ , I denote by  $t[s/x]$  the term obtained by replacing every free occurrence of  $x$  in  $t$  with  $s$ .

I define the notion of a term being *free for* a particular variable in a formula in the standard way. This expresses that the term may be substituted for the variable in the formula without creating any variable binding conflicts.

**Definition 2.11.** A term  $t$  is said to be *free for* a variable  $x$  in a formula  $A$  if for all variables  $y$  which occur free in  $t$ , no free occurrence of  $x$  in  $A$  occurs inside the scope of a  $\exists y$ ,  $\forall y$  or  $\varepsilon_y$  symbol.

**Example 2.12.** The term  $f(x, y)$ , where  $x$  and  $y$  are variables, is free for  $z$  in the atomic formula  $P(z)$  but is not free for  $z$  in the formula  $\exists y P(z)$  since  $z$  occurs within the scope of an existential quantifier  $\exists y$  and  $y$  occurs free in  $f(x, y)$ . Observe that substituting  $f(x, y)$  for  $z$  in  $\exists y P(z)$  results in the formula  $\exists y P(f(x, y))$ , in which  $y$  occurs within the scope of the existential quantifier  $\exists y$ , causing a variable binding conflict.

I define the *size* of terms and formulae in the usual way and introduce the notion of  $\varepsilon$ -size, which measures the size of terms and formulae accounting for formulae inside the scope of epsilon symbols.

**Definition 2.13.** The *size*  $|t|$ ,  $\varepsilon$ -*size*  $|t|_\varepsilon$  of terms  $t$  and *size*  $|A|$  and  $\varepsilon$ -*size*  $|A|_\varepsilon$  of formulae  $A$  are defined recursively as follows:

- For constant terms  $c$ ,  $|c| = |c|_\varepsilon = 1$ .
- For variables  $x$ ,  $|x| = |x|_\varepsilon = 1$ .
- For function terms  $f(t_1, \dots, t_n)$ ,  $|f(t_1, \dots, t_n)| = 1 + \sum_{i=1}^n |t_i|$  and  $|f(t_1, \dots, t_n)|_\varepsilon = 1 + \sum_{i=1}^n |t_i|_\varepsilon$ .
- For  $\varepsilon$ -terms  $\varepsilon_x A$ ,  $|\varepsilon_x A| = 1$  and  $|\varepsilon_x A|_\varepsilon = |A|_\varepsilon$ .
- For formulae  $A$  and  $B$ ,  $|A \vee B| = |A \wedge B| = |A| + |B| + 1$  and  $|A \vee B|_\varepsilon = |A \wedge B|_\varepsilon = |A|_\varepsilon + |B|_\varepsilon + 1$ .
- For formulae  $A$  and variables  $x$ ,  $|\exists x A| = |\forall x A| = |A| + 1$  and  $|\exists x A|_\varepsilon = |\forall x A|_\varepsilon = |A|_\varepsilon + 1$ .
- For atomic formulae  $P(t_1, \dots, t_n)$ ,  $|P(t_1, \dots, t_n)| = 1 + \sum_{i=1}^n |t_i|$  and  $|P(t_1, \dots, t_n)|_\varepsilon = 1 + \sum_{i=1}^n |t_i|_\varepsilon$ .
- $|t| = |f| = |t|_\varepsilon = |f|_\varepsilon = 1$ .

**Remark 2.14.** In this thesis, complexity is of interest primarily for the sake of proving the existence of elementary bounds for the sizes of various proofs and formulae. I have thus chosen to measure the complexity of  $\varepsilon$ -terms in the maximal reasonable way, by the size of the formula inside the scope of the epsilon symbol. Note that this differs from traditional complexity measures of  $\varepsilon$ -terms, such as *rank* and *degree* (see [75]).

## 2.2 Semantics

In this thesis I provide a primarily syntactic treatment of proofs in first-order predicate logic and the epsilon-calculus. However, for the sake of completeness, I will define a semantics for the language defined above in Definition 2.1.

The intended interpretation of an  $\varepsilon$ -term  $\varepsilon_x A(x)$  in a semantics with domain  $\mathbb{D}$  is “some element  $d \in \mathbb{D}$  such that  $A(d)$  is true if such a  $d$  exists, or an arbitrary value otherwise” or, more formally,

$$\llbracket \varepsilon_x A(x) \rrbracket_{\mathbb{D}} = \begin{cases} d & \text{if there exists some } d \in \mathbb{D} \text{ such that } \llbracket A(d) \rrbracket_{\mathbb{D}} \\ a & \text{for some arbitrary } a \in \mathbb{D}, \text{ otherwise} \end{cases} \quad (1)$$

where  $d$  is chosen by a choice function on  $\mathcal{P}(\mathbb{D})$  and  $a \in \mathbb{D}$  is fixed. This leaves open the question of how the choice function on  $\mathcal{P}(\mathbb{D})$  should be defined. In general, this is dependent on the context one is working in and the intended application of the epsilon-calculus. For instance, it is sometimes useful in an arithmetic setting to assume that  $\varepsilon_x A(x)$  is interpreted as the *least* value  $x$  such that  $A(x)$  is true, if such an  $x$  exists, or an arbitrary value otherwise.

One choice which must be made when defining a semantics for the epsilon-calculus is whether the semantics should be *extensional* or *intensional*. In an extensional semantics for the epsilon-calculus, for all formulae  $A(x)$  and  $B(x)$  such that  $A(x) \equiv B(x)$  and all variables  $x$  and  $y$  such that  $y$  is free for  $x$  in  $B(x)$ , the  $\varepsilon$ -terms  $\varepsilon_x A(x)$  and  $\varepsilon_y B(y)$  are interpreted such that  $\varepsilon_x A(x) = \varepsilon_y B(y)$  under the interpretation. An intensional semantics for the epsilon-calculus, on the other hand, imposes no such condition on the interpretation of  $\varepsilon$ -terms so that we may have  $\varepsilon_x A(x) \neq \varepsilon_y B(y)$  under the interpretation. See [89] for an overview and definition of both an extensional and intensional semantics for the epsilon-calculus.

This thesis treats the epsilon-calculus in the general setting of proofs in first-order logic and, as such, requires no particular conditions on the semantics so long as (1) holds for all  $\varepsilon$ -terms  $\varepsilon_x A(x)$ . I have chosen to define an extensional semantics for the epsilon-calculus

following [89], but this choice is essentially arbitrary and all of the results in this thesis would also hold using an intensional semantics or any other valid semantics for the epsilon-calculus.

The semantics of the epsilon-free syntax of first-order predicate logic are standard and well-established, with the first formal semantics having been developed by Tarski [84]. Other notable approaches for defining the semantics of predicate logic include game semantics [33, 58], in which the truth of a formula is defined as the existence of a winning strategy for one of the players in a two-player game determined by the formula.

Following [89], I define a Tarskian semantics for first-order predicate logic and an extensional semantics for the epsilon-calculus as follows.

I begin by defining the notion of a *structure*, which is intended as an interpretation of all function symbols and all predicate symbols in the language in a given domain.

**Definition 2.15.** A *structure* is defined to be a tuple  $\mathcal{A} = \langle |\mathcal{A}|, \llbracket - \rrbracket_{\mathcal{A}}^{\text{Fun}}, \llbracket - \rrbracket_{\mathcal{A}}^{\text{Pred}} \rangle$  consisting of a non-empty set  $|\mathcal{A}|$  called the *domain* of  $\mathcal{A}$ , a function

$$\llbracket - \rrbracket_{\mathcal{A}}^{\text{Fun}} : \text{Fun} \rightarrow \bigcup_{n \geq 0} |\mathcal{A}|^{|\mathcal{A}|^n}$$

such that if  $f \in \text{Fun}$  is a function symbol of arity  $n$  then  $\llbracket f \rrbracket_{\mathcal{A}}^{\text{Fun}} \in |\mathcal{A}|^{|\mathcal{A}|^n}$  and a function

$$\llbracket - \rrbracket_{\mathcal{A}}^{\text{Pred}} : \text{Pred} \rightarrow \bigcup_{n \geq 0} \mathcal{P}(|\mathcal{A}|^n)$$

where  $\mathcal{P}(S) = \{X \mid X \subseteq S\}$  denotes the powerset of  $S$ , such that if  $P \in \text{Pred}$  is a predicate symbol of arity  $n$  then  $\llbracket P \rrbracket_{\mathcal{A}}^{\text{Pred}} \subseteq |\mathcal{A}|^n$  and  $\llbracket \bar{P} \rrbracket_{\mathcal{A}}^{\text{Pred}} = |\mathcal{A}|^n \setminus \llbracket P \rrbracket_{\mathcal{A}}^{\text{Pred}}$ .

For ease of expression, for all structures  $\mathcal{A}$ ,  $f \in \text{Fun}$  and  $P \in \text{Pred}$ , I will write  $\llbracket f \rrbracket_{\mathcal{A}}$  to denote  $\llbracket f \rrbracket_{\mathcal{A}}^{\text{Fun}}$  and  $\llbracket P \rrbracket_{\mathcal{A}}$  to denote  $\llbracket P \rrbracket_{\mathcal{A}}^{\text{Pred}}$ .

I define the notion of an *extensional choice function* on a structure, which will be the choice function used to draw elements from the domain of the structure when interpreting  $\varepsilon$ -terms.

**Definition 2.16.** For all structures  $\mathcal{A}$ , a function  $\Phi : \mathcal{P}(|\mathcal{A}|) \rightarrow |\mathcal{A}|$  is called an *extensional choice function on  $\mathcal{A}$*  if for all  $X \in \mathcal{P}(|\mathcal{A}|) \setminus \{\emptyset\}$ ,  $\Phi(X) \in X$ .

Observe that for all extensional choice functions  $\Phi$  on a structure  $\mathcal{A}$ , since  $\Phi$  is a total function on  $\mathcal{P}(|\mathcal{A}|)$ , this definition implies that  $\Phi(\emptyset) \in |\mathcal{A}|$ , where  $\Phi(\emptyset)$  represents the arbitrary value  $a$  in (1) above.

Note also that in assuming that for all structures  $\mathcal{A}$  an extensional choice function on  $\mathcal{A}$  exists, we are implicitly assuming the axiom of choice in our metatheory.

I define the notion of an *assignment* on a structure, which is an assignment of an element of the domain of the structure to each variable.

**Definition 2.17.** For all structures  $\mathcal{A}$ , an *assignment on  $\mathcal{A}$*  is a function  $s : \text{Var} \rightarrow |\mathcal{A}|$ .

For all structures  $\mathcal{A}$ , all assignments  $s$  on  $\mathcal{A}$ ,  $x \in \text{Var}$  and  $a \in |\mathcal{A}|$ , the assignment  $s[a/x]$  on  $\mathcal{A}$  is defined by

$$s[a/x](y) = \begin{cases} a & \text{if } y = x \\ s(y) & \text{otherwise} \end{cases}$$

for all  $y \in \text{Var}$ .

I now define the *value*  $\text{val}_{\mathcal{A}, \Phi, s}(t)$  of all terms  $t$  and the *satisfaction relation*  $\mathcal{A}, \Phi, s \models A$  of all formulae  $A$  under a structure  $\mathcal{A}$  with extensional choice function  $\Phi$  on  $\mathcal{A}$  and assignment  $s$  on  $\mathcal{A}$ . The value  $\text{val}_{\mathcal{A}, \Phi, s}(t)$  of a term  $t$  is the element of  $|\mathcal{A}|$  assigned to  $t$  under the interpretation and  $\mathcal{A}, \Phi, s \models A$  if and only if the formula  $A$  is true in the interpretation.

**Definition 2.18.** For all structures  $\mathcal{A}$ , all extensional choice functions  $\Phi$  on  $\mathcal{A}$  and all assignments  $s$  on  $\mathcal{A}$ , the *value*  $\text{val}_{\mathcal{A}, \Phi, s}(t)$  of all terms  $t$  and the *satisfaction relation*  $\mathcal{A}, \Phi, s \models A$  on all formulae  $A$  are defined by mutual recursion as follows:

- For all  $x \in \text{Var}$ ,  $\text{val}_{\mathcal{A}, \Phi, s}(x) = s(x)$ .
- $\mathcal{A}, \Phi, s \models \text{t}$  and  $\mathcal{A}, \Phi, s \not\models \text{f}$ .
- For all function terms  $f(t_1, \dots, t_n)$ ,

$$\text{val}_{\mathcal{A}, \Phi, s}(f(t_1, \dots, t_n)) = \llbracket f \rrbracket_{\mathcal{A}}(\text{val}_{\mathcal{A}, \Phi, s}(t_1), \dots, \text{val}_{\mathcal{A}, \Phi, s}(t_n))$$

- For all atomic formulae  $P(t_1, \dots, t_n)$ ,

$$\mathcal{A}, \Phi, s \models P(t_1, \dots, t_n) \text{ if and only if } \langle \text{val}_{\mathcal{A}, \Phi, s}(t_1), \dots, \text{val}_{\mathcal{A}, \Phi, s}(t_n) \rangle \in \llbracket P \rrbracket_{\mathcal{A}}$$

- For all  $\varepsilon$ -terms  $\varepsilon_x A(x)$ ,

$$\text{val}_{\mathcal{A}, \Phi, s}(\varepsilon_x A(x)) = \Phi(\{a \in |\mathcal{A}| \mid \mathcal{A}, \Phi, s[a/x] \models A(x)\})$$

- For all formulae  $A$  and  $B$ ,

$$\mathcal{A}, \Phi, s \models A \vee B \text{ if and only if } \mathcal{A}, \Phi, s \models A \text{ or } \mathcal{A}, \Phi, s \models B$$

and

$$\mathcal{A}, \Phi, s \models A \wedge B \text{ if and only if } \mathcal{A}, \Phi, s \models A \text{ and } \mathcal{A}, \Phi, s \models B$$

- For all formulae  $A(x)$ ,

$$\mathcal{A}, \Phi, s \models \exists x A(x) \text{ if and only if for some } a \in |\mathcal{A}|, \mathcal{A}, \Phi, s[a/x] \models A(x)$$

and

$$\mathcal{A}, \Phi, s \models \forall x A(x) \text{ if and only if for all } a \in |\mathcal{A}|, \mathcal{A}, \Phi, s[a/x] \models A(x)$$

The following lemmas are standard properties of the semantics of predicate logic and will be useful when reasoning about these semantics.

**Lemma 2.19.** *For all structures  $\mathcal{A}$ , all extensional choice functions  $\Phi$  on  $\mathcal{A}$ , all assignments  $s$  on  $\mathcal{A}$  and all formulae  $A$ , either  $\mathcal{A}, \Phi, s \models A$  or  $\mathcal{A}, \Phi, s \models \bar{A}$ .*

**Lemma 2.20.** *For all structures  $\mathcal{A}$ , all extensional choice functions  $\Phi$  on  $\mathcal{A}$ , all assignments  $s$  on  $\mathcal{A}$ , all terms  $t$ , all variables  $x$  and all formulae  $A(x)$ ,  $\mathcal{A}, \Phi, s \models A(t)$  if and only if  $A(x)$ ,  $\mathcal{A}, \Phi, s[t/x] \models A(x)$ .*

Finally, I can now define the *valid* formulae of the language as those formulae which are true in all structures  $\mathcal{A}$  for all choices of extensional choice functions on  $\mathcal{A}$  and all assignments on  $\mathcal{A}$ .

**Definition 2.21.** A formula  $A$  is defined to be *valid*, denoted  $\models A$ , if for all structures  $\mathcal{A}$ , all extensional choice functions  $\Phi$  on  $\mathcal{A}$  and all assignments  $s$  on  $\mathcal{A}$ , we have  $\mathcal{A}, \Phi, s \models A$ .

## 2.3 Open Deduction

The fundamental proof objects of deep-inference formalisms are called *derivations*, which are syntactic constructions representing logical arguments that a formula premise  $A$  logically entails a formula conclusion  $B$ , i.e., that  $\models A$  implies  $\models B$ . The notation

$$\frac{A}{\phi \parallel s} B$$

denotes a derivation  $\phi$  with premise  $A$  and conclusion  $B$  which is constructed using only inference rules from the set  $\mathcal{S}$ . A *proof* is a derivation with premise of the true unit  $t$

$$\begin{array}{c} t \\ \parallel \\ A \end{array}$$

which demonstrates the logical validity of the formula  $A$ , i.e.,  $\models A$ , and is denoted by

$$\begin{array}{c} \parallel \\ A \end{array}$$

Several deep-inference proof formalisms exist in the literature, the most notable being the *calculus of structures* [45] and *open deduction* [47] (sometimes called the *functorial calculus* [48]).

The open deduction formalism places compositionality at the forefront of its design, in that open-deduction proof systems are determined entirely by a set of connectives and a set of inference rules which are used to compose derivations. In the case of the language for first-order predicate logic and the epsilon-calculus given in Definition 2.1, open-deduction derivations are constructed using vertical composition of derivations by inference rules:

$$\phi;_{\rho}\psi \begin{array}{c} A \\ \parallel \\ D \end{array} = \rho \frac{\boxed{\begin{array}{c} A \\ \phi \\ \parallel \\ B \end{array}}}{\boxed{\begin{array}{c} C \\ \psi \\ \parallel \\ D \end{array}}}$$

horizontal composition of derivations by binary connectives:

$$\begin{array}{c} A \vee B \\ \phi \vee \psi \\ \parallel \\ C \vee D \end{array} = \boxed{\begin{array}{c} A \\ \phi \\ \parallel \\ C \end{array}} \vee \boxed{\begin{array}{c} B \\ \psi \\ \parallel \\ D \end{array}} \quad \begin{array}{c} A \wedge B \\ \phi \wedge \psi \\ \parallel \\ C \wedge D \end{array} = \boxed{\begin{array}{c} A \\ \phi \\ \parallel \\ C \end{array}} \wedge \boxed{\begin{array}{c} B \\ \psi \\ \parallel \\ D \end{array}}$$

and application of quantifiers to derivations:

$$\begin{array}{c} \exists x A \\ \exists x \phi \\ \parallel \\ \exists x B \end{array} = \exists x \boxed{\begin{array}{c} A \\ \phi \\ \parallel \\ B \end{array}} \quad \begin{array}{c} \forall x A \\ \forall x \phi \\ \parallel \\ \forall x B \end{array} = \forall x \boxed{\begin{array}{c} A \\ \phi \\ \parallel \\ B \end{array}}$$

Note that when displaying derivations, boxes denote parentheses. Note also that there is no mechanism present for the application of epsilon symbols to derivations, i.e.,

$$\begin{array}{c} \varepsilon_x A \\ \varepsilon_x \phi \\ \parallel \\ \varepsilon_x B \end{array} = \varepsilon_x \boxed{\begin{array}{c} A \\ \phi \\ \parallel \\ B \end{array}}$$

as such constructions are in general unsound using the standard semantics of the epsilon-calculus. This issue is discussed in more detail at the beginning of Section 4.

The open deduction formalism provides a finer representation of proofs than traditional proof formalisms due to its more flexible composition mechanism and the resultant property that inference rules may be applied inside formula contexts:

$$K \left\{ \boxed{\begin{array}{c} A \\ \rho \\ B \end{array}} \right\}$$



As such, I have chosen to use open deduction as the primary proof formalism of this thesis for the purposes of investigating the structure of proofs. Observe that the composition mechanism of open deduction described above results in a greater freedom in applying inference rules over conventional proof systems. For instance, the fact that quantifiers may be applied to derivations allows for the application of inference rules inside the context of quantifiers, a property which is not present in proof systems of the Gentzen tradition, such as the sequent calculus or natural deduction. The disjunctive composition mechanism present in open deduction is also not present in proof systems of the Gentzen tradition, enabling expression at the level of the formalism of proof-theoretic phenomena which involve case distinctions. This has led to novel normalisation procedures which are unique to the deep-inference methodology, such as the *experiments method* [78] (see also [19]), a semantically natural cut-elimination procedure for propositional classical logic which eliminates cuts from a proof by disjunctively composing “projections” of the proof in which atoms are either true or false.

The other most notable deep-inference formalism, the calculus of structures, represents proofs and derivations as sequences of formulae connected vertically by instances of inference rules which act at arbitrary depth inside formulae, so that a total vertical ordering on instances of inference rules exists for every derivation. Derivations in the open deduction formalism may be translated into the calculus of structures by a process known as *sequentialisation* [78, 79] so that the calculus of structures may be considered as a restriction of open deduction. Most modern deep-inference research uses the more flexible formalism of open deduction for defining proof systems and conducting investigations into the deep-inference proof theory of logics and, as such, I have chosen to use open deduction for the purposes of this thesis.

I structure the definition of proofs and derivations in the open deduction formalism as follows: I first define the notion of *inference rules*, which are relations on formulae and will serve as the fundamental building blocks of derivations; I then define *derivations* in the open deduction formalism, which are recursively constructed by composing derivations by inference rules, connectives and quantifiers.

## Inference Rules

To begin, I define the notion of an *inference rule* and an *instance* of an inference rule.

**Definition 2.22.** An *inference rule* is a relation on formulae and an *instance* of an inference rule  $\rho$  is an ordered pair of formulae  $(A, B)$  such that  $(A, B) \in \rho$ , denoted by  $\rho \frac{A}{B}$ .

For all instances  $\rho \frac{A}{B}$  of an inference rule  $\rho$ ,  $A$  is called the *premise* of the instance and  $B$  is called the *conclusion* of the instance.

Overloading the above notation, if  $\rho$  is an inference rule such that for all formulae  $A'$  and  $B'$ ,  $(A', B') \in \rho$  if and only if  $A'$  is of the form  $A$  and  $B'$  is of the form  $B$ , it will be denoted by  $\rho \frac{A}{B}$ .

An inference rule  $\rho$  is intended to be a sound relation on formulae, in that  $(A, B) \in \rho$  should imply that if  $\models A$  then  $\models B$ .

I now define the notion of *dual* inference rules, which reflect the principle of contraposition, that  $A \rightarrow B$  implies  $\overline{B} \rightarrow \overline{A}$ , and will later be used to define the notion of dual derivations (see Definition 2.29).

**Definition 2.23.** For every inference rule  $\rho \frac{A}{B}$ , its *dual* inference rule  $\bar{\rho}$  given by  $\bar{\rho} \frac{\overline{B}}{\overline{A}}$ , i.e.,  $(A, B) \in \rho$  if and only if  $(\overline{B}, \overline{A}) \in \bar{\rho}$ .

Observe that duality of inference rules is an involution, i.e., for all inference rules  $\rho$ ,  $\overline{\bar{\rho}} = \rho$ .

**Example 2.24.** The inference rule *cocontraction*  $\mathsf{c}\uparrow \frac{A}{A \wedge A}$  is given by  $\mathsf{c}\uparrow = \{(B, B \wedge B) \mid B \text{ is a formula}\}$  and for each instance  $\mathsf{c}\uparrow \frac{B}{B \wedge B}$  of  $\mathsf{c}\uparrow$ , the premise of the instance is  $B$  and the conclusion of the instance is  $B \wedge B$ . The dual of the cocontraction rule  $\mathsf{c}\uparrow$  is the *contraction* rule  $\mathsf{c}\downarrow$ , given by  $\mathsf{c}\downarrow \frac{A \vee A}{A}$ .

## Derivations

I define derivations in the *open deduction* formalism [47] as follows.

**Definition 2.25.** I define *derivations*  $\phi$  with formula premises  $A$  and conclusions  $B$ , denoted  $\phi \parallel \frac{A}{B}$ , inductively as follows:

- Every formula  $A$  is a derivation, with premise  $A$  and conclusion  $A$ .

For all derivations  $\psi \parallel \frac{A}{A'}$  and  $\chi \parallel \frac{B}{B'}$ , we have the following:

- Composition by inference: if  $\rho \frac{A'}{B}$  is an instance of an inference rule  $\rho$ ,

$$\psi;_{\rho}\chi \parallel \frac{A}{B'} = \rho \frac{\boxed{\psi \parallel \frac{A}{A'}}}{\boxed{\chi \parallel \frac{B}{B'}}}$$

is a derivation.

- Composition by connective: for  $\star \in \{\vee, \wedge\}$ ,

$$\psi \star \chi \parallel \frac{A \star B}{A' \star B'} = \boxed{\psi \parallel \frac{A}{A'}} \star \boxed{\chi \parallel \frac{B}{B'}}$$

is a derivation.

- Composition by quantifier: for  $Qx \in \{\forall x, \exists x\}$ , where  $x$  is any variable,

$$\begin{array}{c} QxA \\ Qx\psi \parallel \\ QxA' \end{array} = Qx \boxed{\psi \parallel \frac{A}{A'}}$$

is a derivation.

Composition by inference and composition by connective are defined to be associative: for  $\star \in \{\vee, \wedge\}$ , all inference rules  $\rho_1$  and  $\rho_2$  and all derivations  $\psi$ ,  $\phi$  and  $\chi$ ,

$$(\psi \star \phi) \star \chi = \psi \star (\phi \star \chi) = \psi \star \phi \star \chi$$

and

$$\psi;_{\rho_1} (\phi;_{\rho_2} \chi) = (\psi;_{\rho_1} \phi);_{\rho_2} \chi = \psi;_{\rho_1} \phi;_{\rho_2} \chi$$

Furthermore, if every instance of an inference rule in a derivation  $\phi$  is an instance of an inference rule contained in the set  $\mathcal{S}$ ,  $\phi$  is said to be a *derivation in  $\mathcal{S}$* , which is denoted by

$$\phi \parallel_{\mathcal{S}} \frac{A}{B}$$

If a derivation  $\phi$  with conclusion  $A$  has premise  $t$ , it is called a *proof*, denoted by

$$\phi \parallel \frac{}{A}$$

An example of a derivation is given in Example 3.5, after several sets of inference rules have been defined.

The *subderivations* of a derivation are defined in the expected way, as follows.

**Definition 2.26.** For every derivation  $\phi$ , its *subderivations* are defined recursively as follows:

- $\phi$  is a subderivation of itself.
- If  $\phi = \psi;_{\rho} \chi$  for some derivations  $\psi$  and  $\chi$  and inference rule  $\rho$ , then every subderivation of  $\psi$  and every subderivation of  $\chi$  is a subderivation of  $\phi$ .
- If  $\phi = \psi \star \chi$  for some derivations  $\psi$  and  $\chi$  and  $\star \in \{\vee, \wedge\}$ , then every subderivation of  $\psi$  and every subderivation of  $\chi$  is a subderivation of  $\phi$ .
- If  $\phi = Qx\psi$  for some derivation  $\psi$  and  $Qx \in \{\exists x, \forall x\}$ , then every subderivation of  $\psi$  is a subderivation of  $\phi$ .

I define the notion of an *occurrence* of an inference rule in a derivation, which refers to the subderivation corresponding to a particular composition by inference in a derivation.

**Definition 2.27.** Let  $\phi$  be a derivation and let  $\chi$  be a subderivation of  $\phi$  of the form  $\rho \frac{A}{B}$ . Then  $\chi$  is said to be an *occurrence of  $\rho$  in  $\phi$* .

The standard definition of a *subformula* of a formula may be defined using the notion of subderivations, as follows.

**Definition 2.28.** A formula  $A$  which is a subderivation of a formula  $B$  is said to be a *subformula* of  $B$ .

A remarkable property of deep-inference proof systems is that the symmetry observed for their inference rules may be extended to derivations so that the principle of contraposition may be observed at the level of the formalism. I define the *duals* of derivations as follows.

**Definition 2.29.** For every derivation  $\phi \parallel \frac{A}{B}$ , its *dual*  $\bar{\phi} \parallel \frac{\bar{B}}{\bar{A}}$  is defined recursively as follows:

- If  $\phi$  is a formula  $A$ , then  $\bar{\phi} = \bar{A}$ .
- If  $\phi = \psi;_{\rho} \chi$  for some derivations  $\psi$  and  $\chi$  and inference rule  $\rho$ , then  $\bar{\phi} = \bar{\chi};_{\bar{\rho}} \bar{\psi}$ .
- If  $\phi = \psi \star \chi$  for some derivations  $\psi$  and  $\chi$  and  $\star \in \{\vee, \wedge\}$ , then  $\bar{\phi} = \bar{\psi} \bar{\star} \bar{\chi}$ , where  $\bar{\vee} = \wedge$  and  $\bar{\wedge} = \vee$ .
- If  $\phi = Qx\psi$  for some derivation  $\psi$  and  $Qx \in \{\exists x, \forall x\}$ , then  $\bar{\phi} = \bar{Qx}\bar{\psi}$ , where  $\bar{\exists x} = \forall x$  and  $\bar{\forall x} = \exists x$ .

Observe that duality of derivations is an involution, i.e., for all derivations  $\phi$ ,  $\overline{\overline{\phi}} = \phi$ .

I extend the notion of a term occurring in a formula that was given in Definition 2.6 to derivations, as follows.

**Definition 2.30.** A term is defined to *occur* in a derivation  $\phi$  recursively as follows:

- If  $\phi = \psi;_{\rho} \chi$  for some derivations  $\psi$  and  $\chi$  and inference rule  $\rho$ , then every term which occurs in  $\psi$  or  $\chi$  also occurs in  $\phi$ .
- If  $\phi = \psi \star \chi$  for some derivations  $\psi$  and  $\chi$  and  $\star \in \{\vee, \wedge\}$ , then every term which occurs in  $\psi$  or  $\chi$  also occurs in  $\phi$ .
- If  $\phi = Qx\psi$  for some derivation  $\psi$  and  $Qx \in \{\exists x, \forall x\}$ , then every term which occurs in  $\psi$  also occurs in  $\phi$ .

I also extend the notion of a free occurrence of a variable in a formula that was given in Definition 2.7 to derivations, as follows.

**Definition 2.31.** An occurrence of a variable  $x$  in a derivation  $\phi$  is said to be a *free occurrence* if it does not occur inside the scope of any  $\exists x$ ,  $\forall x$  or  $\varepsilon_x$  symbols in  $\phi$ .

I extend the notions of size and  $\varepsilon$ -size that were defined for formulae in Definition 2.13 to derivations, as follows.

**Definition 2.32.** For every derivation  $\phi$ , its *size*  $|\phi|$  and  $\varepsilon$ -*size*  $|\phi|_{\varepsilon}$  are defined recursively as follows:

- If  $\phi$  is a formula  $A$ , then

$$|\phi| = |A| \text{ and } |\phi|_{\varepsilon} = |A|_{\varepsilon}$$

- If  $\phi = \psi;_{\rho} \chi$  for some derivations  $\psi$  and  $\chi$  and inference rule  $\rho$ , then

$$|\phi| = |\chi| + |\psi| \text{ and } |\phi|_{\varepsilon} = |\chi|_{\varepsilon} + |\psi|_{\varepsilon}$$

- If  $\phi = \psi \star \chi$  for some derivations  $\psi$  and  $\chi$  and  $\star \in \{\vee, \wedge\}$ , then

$$|\phi| = |\chi| + |\psi| + 1 \text{ and } |\phi|_{\varepsilon} = |\chi|_{\varepsilon} + |\psi|_{\varepsilon} + 1$$

- If  $\phi = Qx\psi$  for some derivation  $\psi$  and  $Qx \in \{\exists x, \forall x\}$ , then

$$|\phi| = |\psi| + 1 \text{ and } |\phi|_{\varepsilon} = |\psi|_{\varepsilon} + 1$$

A useful consequence of the definitions of size and  $\varepsilon$ -size is that the  $\varepsilon$ -size of a term, formula or derivation is bounded by the product of its size and the  $\varepsilon$ -size of the largest  $\varepsilon$ -term which occurs in it. This lemma will be used when providing bounds for the  $\varepsilon$ -size of various formulae and derivations throughout this thesis.

**Lemma 2.33.** For all terms  $t$  which are not *epsilon-free*,  $|t|_{\varepsilon} \leq |\varepsilon_x B|_{\varepsilon} |t|$ , where  $\varepsilon_x B$  is the largest  $\varepsilon$ -term which occurs in  $t$ .

Likewise, for all derivations  $\phi$  which are not *epsilon-free*,  $|\phi|_{\varepsilon} \leq |\varepsilon_x B|_{\varepsilon} |\phi|$ , where  $\varepsilon_x B$  is the largest  $\varepsilon$ -term which occurs in  $\phi$ .

*Proof.* By structural induction on the term or derivation. □

A formula or derivation is defined to be *epsilon-free* if it contains no  $\varepsilon$ -terms, as follows.

**Definition 2.34.** A formula  $A$  is said to be *epsilon-free* if no  $\varepsilon$ -terms occur in  $A$ .

Likewise, a derivation  $\phi$  is said to be *epsilon-free* if no  $\varepsilon$ -terms occur in  $\phi$ .

I introduce the following notation to denote substituting a term for a variable in a derivation, as follows.

**Definition 2.35.** For every derivation  $\phi$ , term  $t$  and variable  $x$ , I denote by  $\phi[t/x]$  the derivation obtained by replacing every free occurrence of  $x$  in  $\phi$  with  $t$ .

I introduce the notion of a *formula context*, which will be useful when reasoning about the structure of derivations.

**Definition 2.36.** A *formula context*  $K\{\}$  is a function from derivations to derivations which is a formula with exactly one occurrence of the *hole*  $\{-\}$  in the position of an atomic formula. For all derivations  $\phi$ ,  $K\{\phi\}$  is given by replacing the hole in  $K\{\}$  with  $\phi$ .

For convenience in normalisation and describing the structure of derivations, I introduce the notion of the *sequential composition* of any two derivations  $\phi$  and  $\psi$  for which the conclusion of  $\phi$  is the premise  $\psi$ , which is the derivation obtained by connecting the two derivations by this shared formula.

**Definition 2.37.** Let  $\phi \parallel \begin{smallmatrix} A \\ A' \end{smallmatrix}$  and  $\psi \parallel \begin{smallmatrix} A' \\ B \end{smallmatrix}$  be derivations. The *sequential composition*  $\phi; \psi$  of  $\phi$  and  $\psi$  is defined recursively as follows:

- If  $\phi$  is a formula, then  $\phi; \psi = \psi$ . Likewise, if  $\psi$  is a formula, then  $\phi; \psi = \phi$ .
- If  $\phi = \chi;_{\rho} \omega$ , where  $\chi$  and  $\omega$  are derivations and  $\rho$  is an inference rule, then  $\phi; \psi = \chi;_{\rho} (\omega; \psi)$ . Likewise, if  $\psi = \chi;_{\rho} \omega$ , then  $\phi; \psi = (\phi; \chi);_{\rho} \omega$ .
- If  $\phi = \chi \star \omega$  and  $\psi = \chi' \star \omega'$ , where  $\star \in \{\vee, \wedge\}$  and  $\chi, \chi', \omega$  and  $\omega'$  are derivations such that the conclusion of  $\chi$  is the premise of  $\chi'$  and the conclusion of  $\omega$  is the premise of  $\omega'$ , then  $\phi; \psi = (\chi; \chi') \star (\omega; \omega')$ .
- If  $\phi = Qx\chi$  and  $\psi = Qx\chi'$ , where  $Qx \in \{\forall x, \exists x\}$  for some variable  $x$  and  $\chi$  and  $\chi'$  are derivations, then  $\phi; \psi = Qx(\chi; \chi')$ .

I will write

$$\phi \parallel \begin{smallmatrix} A \\ A' \\ \psi \parallel B \end{smallmatrix}$$

to denote  $\phi; \psi$ .

As one would expect, sequential composition of derivations is associative so that we may unambiguously use the notation

$$\phi \parallel \begin{smallmatrix} A \\ B \\ \psi \parallel C \\ \chi \parallel D \end{smallmatrix}$$

to denote the sequential composition of three derivations  $\phi$ ,  $\psi$  and  $\chi$ , as follows.

**Lemma 2.38.** *Sequential composition is associative, i.e., for all derivations  $\phi \parallel \begin{smallmatrix} A \\ B \end{smallmatrix}$ ,  $\psi \parallel \begin{smallmatrix} B \\ C \end{smallmatrix}$  and  $\chi \parallel \begin{smallmatrix} C \\ D \end{smallmatrix}$ , we have*

$$(\psi; \phi); \chi = \psi; (\phi; \chi)$$

*Proof.* By induction on  $|\phi| + |\chi|$  (see [78], Lemma 1.10).  $\square$

I define the notion of an occurrence of an inference rule *occurring above* or *occurring below* another occurrence of an inference rule in a derivation, which will be useful when reasoning about permuting occurrences of inference rules up or down a derivation. Such permutations are central to the proofs of many normalisation results and decomposition theorems in deep inference, such as the main result of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem.

**Definition 2.39.** Let  $\chi_1$  and  $\chi_2$  be occurrences of inference rules in a derivation  $\phi$ . If  $\phi$  may be expressed in the form

$$\begin{array}{c} C \\ \psi_1 \parallel \\ K\{\chi_1\} \\ \psi_2 \parallel \\ J\{\chi_2\} \\ \psi_3 \parallel \\ C' \end{array}$$

for some derivations  $\psi_1, \psi_2, \psi_3$  and formula contexts  $K\{\}$  and  $J\{\}$  but not in the form

$$\begin{array}{c} D \\ \psi'_1 \parallel \\ K'\{\chi_2\} \\ \psi'_2 \parallel \\ J'\{\chi_1\} \\ \psi'_3 \parallel \\ D' \end{array}$$

for any derivations  $\psi'_1, \psi'_2, \psi'_3$  and formula contexts  $K'\{\}$  and  $J'\{\}$ , then  $\chi_1$  is said to *occur above*  $\chi_2$  in  $\phi$  and  $\chi_2$  is said to *occur below*  $\chi_1$  in  $\phi$ .

An occurrence of an inference rule  $\rho$  in a derivation  $\phi$  is said to be a *lowermost rule occurrence (of  $\rho$ )* in  $\phi$  if it does not occur above any other rule occurrences (of  $\rho$ ) in  $\phi$ . Likewise, it is said to be an *uppermost rule occurrence (of  $\rho$ )* in  $\phi$  if it does not occur below any other rule occurrences (of  $\rho$ ) in  $\phi$ .

When permuting occurrences of inference rules up or down a derivation, they are successively permuted up or down through occurrences of inference rules which neighbour them vertically. This notion is encapsulated by the following definition of an occurrence of an inference rule in a derivation *occurring immediately above* or *occurring immediately below* another occurrence of an inference rule in the derivation.

**Definition 2.40.** An occurrence  $\chi$  of an inference rule in a derivation  $\phi$  is said to *occur immediately above* an occurrence  $\chi'$  of an inference rule in  $\phi$  if  $\chi$  occurs above  $\chi'$  in  $\phi$  and there is no occurrence  $\chi''$  of an inference rule in  $\phi$  which occurs above  $\chi'$  in  $\phi$  and occurs below  $\chi$  in  $\phi$ . Likewise,  $\chi'$  is said to *occur immediately below*  $\chi$  in  $\phi$ .

Note that several distinct occurrences of inference rules may occur immediately above or immediately below an occurrence of an inference rule in a derivation, so that these notions do not describe unique occurrences of inference rules in a derivation.

I now define the notion of a set of occurrences of inference rules in a derivation *occurring in parallel* in the derivation. This will be useful when performing complexity assessments on procedures which permute occurrences of inference rules up or down a derivation, such as in the proof of Theorem 4.13 the Falsifier Decomposition Theorem, as a single occurrence of an inference rule cannot be permuted through two occurrences of inference rules which occur in parallel in a derivation.

**Definition 2.41.** Let  $\phi$  be a derivation and let  $S$  be a set of occurrences of inference rules in  $\phi$  such that for all  $\chi_1 \in S$  and  $\chi_2 \in S$ ,  $\chi_1$  neither occurs above nor occurs below  $\chi_2$  in  $\phi$ . Then the rule occurrences in  $S$  are said to *occur in parallel* in  $\phi$ .

### 3 The Structure of First-Order Proofs

What we observe is not nature itself, but nature exposed to our method of questioning.

---

W. Heisenberg

In this section I will introduce several open-deduction proof systems for first-order predicate logic and use them to explore the structure and complexity of first-order proofs. In particular, I discuss two phenomena in first-order proof theory which are central to this thesis: the non-elementary compression of cut-free proofs and the extraction of case analyses contained in quantifier contraction rules from proofs. These topics are the primary motivations for the introduction of the falsifier calculus and the main result of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem, which will be introduced in the following section, Section 4.

In Subsection 3.1 I define several sets of inference rules which will be used to form open-deduction proof systems. In particular, I define a set of inference rules  $\text{SKSg1}$  which form a sound and complete open-deduction proof system for epsilon-free first-order predicate logic, which will be used as the main reference proof system for first-order logic throughout this thesis, and demonstrate that  $\text{SKSg1}$  may be used to simulate several other proof systems in the proof theory literature. In Subsection 3.2 I demonstrate that, as is standard in deep inference, instances of inference rules in  $\text{SKSg1}$  may be decomposed into instances of smaller inference rules in the set  $\text{SKS1}$ , which is comprised of atomic variants of certain inference rules in  $\text{SKSg1}$  as well as certain other restricted inference rules. In Subsection 3.3 I explore the proof theory of the inference rules in  $\text{SKSg1}$  known as quantifier-shifts and their connection to the phenomenon of the non-elementary compression of cut-free proofs admitted by some proof systems over the sequent calculus. In Subsection 3.4 I discuss the extraction of case analyses contained in the existential contraction rule and the universal cocontraction rule from first-order proofs, how this extraction is related to Herbrand's Theorem, and why a certain procedure for extracting such case analyses is non-terminating in the standard syntax of predicate logic.

#### 3.1 Proof Systems

I define the following sets of deep-inference inference rules, which will be used to construct derivations and proofs.

$\text{SKSg1}$  is a set of inference rules which will be used to form a sound and complete proof system for first-order predicate logic (see Theorem 3.11), which will serve as the main reference proof system for first-order predicate logic in this thesis.  $\text{SKSg}_P$  is a subset of the inference rules in  $\text{SKSg1}$  which are restricted so that certain inference rules act only on weakly quantifier-free formulae so that the inference rules are “propositional”, which will be used later when defining the falsifier calculus  $\text{SKSg}_\varepsilon$  (see Definition 4.3). The inference rules  $r1\downarrow, r1\uparrow, r2\downarrow, r2\uparrow, r3\downarrow, r3\uparrow, r4\downarrow$  and  $r4\uparrow$  in  $\text{SKSg1}$  are known as *quantifier-shifts* and will be discussed in more detail in Section 3.3.

The set of inference rules  $\text{SKS1}$  is also a restriction of  $\text{SKSg1}$ , containing atomic variants of certain inference rules in  $\text{SKSg1}$ , which act only on atomic formulae and units, the *existential contraction* rule  $qc\downarrow$  and its dual the *universal cocontraction* rule  $qc\uparrow$ , which are restrictions of the contraction rule  $c\downarrow$  and the cocontraction rule  $c\uparrow$  to respectively act only on existential formulae and universal formulae, as well as restrictions of the equality rules  $=\exists$  and  $=\forall$ . Following conventional deep-inference taxonomy, the “g” in “ $\text{SKSg1}$ ” stands for “general” rules and is hence dropped in the name of the restricted system “ $\text{SKS1}$ ”. The compositionality admitted by deep-inference proof systems results in a finer granularity of the inference rules, so that instances of inference rules in  $\text{SKSg1}$  may be decomposed into derivations in  $\text{SKS1}$  (see Lemma 3.17), meaning that the inference rules of  $\text{SKS1}$  may also be used to form a sound and complete proof system for first-order predicate logic. Such



decompositions will be the subject of the following subsection, Subsection 3.2. Instances of the medial rule  $\mathbf{m}$  may also be decomposed into derivations in  $\mathbf{SKS1} \setminus \{\mathbf{m}\}$ , but I have chosen to include  $\mathbf{m}$  in  $\mathbf{SKS1}$  in this thesis.

**Definition 3.1.** I define the following sets of inference rules:

$\mathbf{SKSg}_P$  the set of inference rules given at the top of Figure 1, with the restriction that all instances of the inference rules  $\mathbf{i}\downarrow, \mathbf{i}\uparrow, \mathbf{w}\downarrow, \mathbf{w}\uparrow, \mathbf{c}\downarrow$  and  $\mathbf{c}\uparrow$  may include only weakly quantifier-free formulae in the premise and conclusion

$\mathbf{SKSg1}$  the set of all inference rules given in Figure 1, without any such restriction

$\mathbf{SKS1}$  the set of all inference rules given in Figure 2

where:

- For all of the symmetric relations  $= \in \{=\mathbf{P}, =\exists, =\forall\}$  given by  $A = B$  in Figures 1 and 2, the corresponding inference rules are given by  $= \frac{A}{B}$
- The inference rules corresponding to  $A = QxA$  for  $Qx \in \{\exists x, \forall x\}$  are called *vacuous*  $=$  rules

For ease of expression, I will often denote the inference rules  $=\mathbf{P}$ ,  $=\exists$  and  $=\forall$  simply by  $=$  and sometimes omit instances of  $=$  rules when displaying derivations. The equality rule  $=$  has been divided into the three inference rules of propositional equality  $=\mathbf{P}$ , existential equality  $=\exists$  and universal equality  $=\forall$  since certain results in this thesis require this distinction between the three rules. For instance, the falsifier calculus  $\mathbf{SKSg}_\varepsilon$  is defined in Definition 4.3 to contain  $=\mathbf{P}$  and  $=\forall$ , but not  $=\exists$ .

I introduce the following definition to distinguish the terms which occur in instances of the inference rules  $\exists$  and  $\forall$ .

**Definition 3.2.** For any instance  $\exists \frac{A(t)}{\exists xA(x)}$  of the  $\exists$  rule, the term  $t$  is said to *witness* the instance of the  $\exists$  rule.

Likewise, for any instance  $\forall \frac{\forall xA(x)}{A(t)}$  of the  $\forall$  rule, the term  $t$  is said to *instantiate* the instance of the  $\forall$  rule.

The completeness of the open-deduction proof system with rules  $\mathbf{SKSg1}$  with respect to the semantics of epsilon-free first-order predicate logic will be given later by Theorem 3.11.

Observe that for each inference rule  $\rho \frac{A}{B}$  in each of the sets defined above, its dual inference rule  $\bar{\rho} \frac{\bar{B}}{\bar{A}}$  also belongs to the set, as follows.

**Proposition 3.3.** *The sets  $\mathbf{SKSg}_P$ ,  $\mathbf{SKSg1}$  and  $\mathbf{SKS1}$  are closed under dual inference rules, i.e., for all  $\mathcal{S} \in \{\mathbf{SKSg}_P, \mathbf{SKSg1}, \mathbf{SKS1}\}$ , if  $\rho \in \mathcal{S}$ , then  $\bar{\rho} \in \mathcal{S}$ .*

**Remark 3.4.** Closure under dual inference rules is a standard property of the inference rules of deep-inference proof systems, and most of the inference rules defined are named to reflect this: each inference rule with a name of the form “ $\rho\downarrow$ ” is known as a *down rule* and its corresponding dual inference rule with the name “ $\rho\uparrow$ ” is known as an *up rule*. The inference rule  $\exists$  may be considered as a down rule and its dual  $\forall$  may be considered as an up rule. The inference rules switch  $\mathbf{s}$  and medial  $\mathbf{m}$  are self-dual.

I provide the following example of a derivation in  $\mathbf{SKSg1}$ .

The inference rules of  $\text{SKSg}_P$  and  $\text{SKSg1}$ :

$$\boxed{\begin{array}{cccc} \text{i}\downarrow \frac{t}{A \vee \bar{A}} & \text{w}\downarrow \frac{f}{A} & \text{c}\downarrow \frac{A \vee A}{A} & \text{s} \frac{A \wedge (B \vee C)}{(A \wedge B) \vee C} \\ \text{i}\uparrow \frac{A \wedge \bar{A}}{f} & \text{w}\uparrow \frac{A}{t} & \text{c}\uparrow \frac{A}{A \wedge A} & \text{m} \frac{(A \wedge B) \vee (C \wedge D)}{(A \vee C) \wedge (B \vee D)} \end{array}}$$

where in  $\text{SKSg}_P$ , the formula  $A$  is weakly quantifier-free in the inference rules  $\text{i}\downarrow$ ,  $\text{i}\uparrow$ ,  $\text{w}\downarrow$ ,  $\text{w}\uparrow$ ,  $\text{c}\downarrow$  and  $\text{c}\uparrow$ .

The equality rules of  $\text{SKSg}_P$  and  $\text{SKSg1}$ :

$$\boxed{\begin{array}{llll} A \vee f =_P A & f \wedge f =_P f & t \vee t =_P t & A \wedge t =_P A \\ A \vee B =_P B \vee A & (A \vee B) \vee C =_P A \vee (B \vee C) & & \\ A \wedge B =_P B \wedge A & (A \wedge B) \wedge C =_P A \wedge (B \wedge C) & & \end{array}}$$

The remaining inference rules of  $\text{SKSg1}$ :

$$\boxed{\begin{array}{ccccc} \exists \frac{A(t)}{\exists x A(x)} & \text{r1}\downarrow \frac{\forall x(A(x) \vee B)}{\forall x A(x) \vee B} & \text{r2}\downarrow \frac{\forall x(A(x) \wedge B)}{\forall x A(x) \wedge B} & \text{r3}\downarrow \frac{\exists x(A(x) \vee B)}{\exists x A(x) \vee B} & \text{r4}\downarrow \frac{\exists x(A(x) \wedge B)}{\exists x A(x) \wedge B} \\ \forall \frac{\forall x A(x)}{A(t)} & \text{r1}\uparrow \frac{\exists x A(x) \wedge B}{\exists x(A(x) \wedge B)} & \text{r2}\uparrow \frac{\exists x A(x) \vee B}{\exists x(A(x) \vee B)} & \text{r3}\uparrow \frac{\forall x A(x) \wedge B}{\forall x(A(x) \wedge B)} & \text{r4}\uparrow \frac{\forall x A(x) \vee B}{\forall x(A(x) \vee B)} \end{array}}$$

where  $t$  is free for  $x$  in  $A(x)$  in the inference rules  $\exists$  and  $\forall$  and  $x$  does not occur free in  $B$  in the remaining inference rules.

The remaining equality rules of  $\text{SKSg1}$ :

$$\boxed{\begin{array}{lll} A =_{\exists} \exists x A & \exists x \exists y A =_{\exists} \exists y \exists x A & \exists x A(x) =_{\exists} \exists y A(y) \\ A =_{\forall} \forall x A & \forall x \forall y A =_{\forall} \forall y \forall x A & \forall x A(x) =_{\forall} \forall y A(y) \end{array}}$$

where  $x$  does not occur free in  $A$  in  $A =_{\exists} \exists x A$  and  $A =_{\forall} \forall x A$  and  $y$  is free for  $x$  in  $A(x)$  in  $\exists x A(x) =_{\exists} \exists y A(y)$  and  $\forall x A(x) =_{\forall} \forall y A(y)$ .

**Figure 1:** The inference rules of  $\text{SKSg}_P$  and  $\text{SKSg1}$

The inference rules of SKS1:

$$\begin{array}{cccc}
\text{ai}\downarrow \frac{\mathbf{t}}{a \vee \bar{a}} & \text{aw}\downarrow \frac{\mathbf{f}}{a} & \text{ac}\downarrow \frac{a \vee a}{a} & \text{s} \frac{A \wedge (B \vee C)}{(A \wedge B) \vee C} \\
\text{ai}\uparrow \frac{a \wedge \bar{a}}{\mathbf{f}} & \text{aw}\uparrow \frac{a}{\mathbf{t}} & \text{ac}\uparrow \frac{a}{a \wedge a} & \text{m} \frac{(A \wedge B) \vee (C \wedge D)}{(A \vee C) \wedge (B \vee D)}
\end{array}$$

where  $a$  is either an atomic formula or a unit  $\mathbf{t}$  or  $\mathbf{f}$ .

$$\begin{array}{ccc}
\exists \frac{A(t)}{\exists x A(x)} & \text{r1}\downarrow \frac{\forall x(A(x) \vee B)}{\forall x A(x) \vee B} & \text{qc}\downarrow \frac{\exists x A \vee \exists x A}{\exists x A} \\
\forall \frac{\forall x A(x)}{A(t)} & \text{r1}\uparrow \frac{\exists x A(x) \wedge B}{\exists x(A(x) \wedge B)} & \text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}
\end{array}$$

where  $t$  is free for  $x$  in  $A(x)$  in the inference rules  $\exists$  and  $\forall$  and  $x$  does not occur free in  $B$  in the inference rules  $\text{r1}\downarrow$  and  $\text{r1}\uparrow$ .

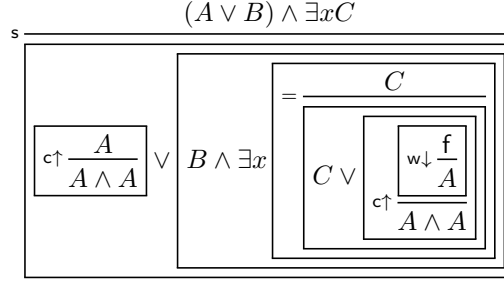
The equality rules of SKS1:

$$\begin{array}{cccc}
A \vee \mathbf{f} =_{\mathbf{P}} A & \mathbf{f} \wedge \mathbf{f} =_{\mathbf{P}} \mathbf{f} & \mathbf{t} \vee \mathbf{t} =_{\mathbf{P}} \mathbf{t} & A \wedge \mathbf{t} =_{\mathbf{P}} A \\
A \vee B =_{\mathbf{P}} B \vee A & & (A \vee B) \vee C =_{\mathbf{P}} A \vee (B \vee C) & \\
A \wedge B =_{\mathbf{P}} B \wedge A & & (A \wedge B) \wedge C =_{\mathbf{P}} A \wedge (B \wedge C) & \\
=_{\exists} \frac{\exists x A}{A} & =_{\forall} \frac{A}{\forall x A} & \exists x A(x) =_{\exists} \exists y A(y) & \forall x A(x) =_{\forall} \forall y A(y)
\end{array}$$

where  $x$  does not occur free in  $A$  in the inference rules  $=_{\exists} \frac{\exists x A}{A}$  and  $=_{\forall} \frac{A}{\forall x A}$  and  $y$  is free for  $x$  in  $A(x)$  in  $\exists x A(x) =_{\exists} \exists y A(y)$  and  $\forall x A(x) =_{\forall} \forall y A(y)$ .

**Figure 2:** The inference rules of SKS1

**Example 3.5.** The following is a derivation in **SKSg1** with premise  $(A \vee B) \wedge \exists x C$  and conclusion  $(A \wedge A) \vee (B \wedge \exists x(C \vee (A \wedge A)))$



where  $A$ ,  $B$  and  $C$  are formulae and  $x$  is a variable.

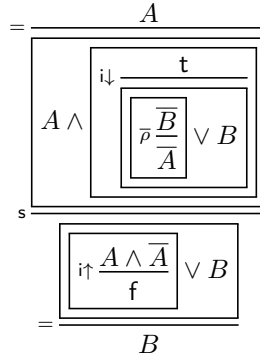
In the derivation above, the instance  $\text{c}\uparrow \frac{A}{A \wedge A}$  of the  $\text{c}\uparrow$  rule is used in two separate composition by inferences. As such, these correspond to two distinct occurrences of  $\text{c}\uparrow$  in the derivation, despite the instances being technically identical when considered as ordered pairs of formulae.

The following proposition demonstrates that every inference rule  $\rho$  is derivable using its dual  $\bar{\rho}$  and the cut rule  $\text{i}\uparrow$  in deep inference. Although this fact will not be utilised in this thesis, it is commonly used as a lemma when proving admissibility of the cut rule  $\text{i}\uparrow$  from a deep-inference proof system, to demonstrate that all up rules are also admissible from the system (see, for instance, [21]). It follows that in any deep-inference proof system in which the cut rule  $\text{i}\uparrow$  is admissible, only the down rules of the system are required to attain completeness.

**Proposition 3.6.** *For all inferences rules  $\rho$  and all instances  $\rho \frac{A}{B}$  of  $\rho$ , there exists a derivation*

$$\begin{array}{c}
 A \\
 \parallel_{\{\bar{\rho}, \text{i}\downarrow, \text{i}\uparrow, \text{s}, =_P\}} \\
 B
 \end{array}$$

*Proof.* The following derivation is of the desired form.



□

I now define the notion of a *cut-free* derivation to be a derivation which contains no occurrences of the cut rule  $\text{i}\uparrow$  or its atomic variant  $\text{ai}\uparrow$ .

As demonstrated by Proposition 3.6 above, every up rule in a deep-inference proof system is derivable using its corresponding down rule and the cut rule  $\text{i}\uparrow$ . As such, a cut-free derivation in deep inference is sometimes defined to be a derivation which contains no occurrences of up rules (see, e.g., [21, 26]). Furthermore, unlike in the sequent calculus, cut-free deep-inference proofs do not necessarily admit the subformula property, the property

that every formula in the proof is a subformula of the conclusion of the proof, so that the notion of cut freeness for deep inference may not necessarily be considered to be a direct analogue of that of the sequent calculus. See [23, 24] for a discussion of the notion of analyticity which cut-free deep-inference proofs admit in place of the subformula property.

Furthermore, an instance of the cut rule  $i\uparrow$  on a quantified formula  $\forall xA(x)$  may be simulated by an instance of the universal instantiation rule  $\forall$ , an instance of the quantifier-shift  $r1\uparrow$  and an instance of the cut rule  $i\uparrow$  on the formula  $A(x)$  in deep inference, as follows.

$$\begin{array}{c}
\frac{i\uparrow \quad \frac{\forall xA(x) \wedge \exists x\bar{A}(x)}{f}}{f} \quad \rightarrow \quad \frac{r1\uparrow \quad \frac{\frac{\frac{\forall xA(x)}{\forall yA(y)} \wedge \exists x\bar{A}(x)}{\frac{\forall yA(y)}{A(x)} \wedge \bar{A}(x)}}{\exists x \quad \frac{i\uparrow \quad \frac{\frac{\forall yA(y)}{A(x)} \wedge \bar{A}(x)}{f}}{f}}}{f}
\end{array}$$

This can be problematic when considering cut-free derivations since a derivation may contain no occurrences of the cut rule  $i\uparrow$ , but contain other rule occurrences which simulate certain parts of a cut. As such, it is sometimes relevant when reasoning about cut-free derivations to also require that the derivations contain no occurrences of the universal instantiation rule  $\forall$ , as will be the case several times in this thesis.

The following definition of a cut-free derivation used for this thesis follows the conventions of modern deep-inference literature [25, 46].

**Definition 3.7.** A derivation is said to be *cut-free* if it contains no occurrences of the rules  $i\uparrow$  or  $ai\uparrow$ .

### Simulation of Other Proof Systems

Several results in this thesis will be proved by means of translating proofs between different proof systems in the literature and those presented in this thesis. I present here some lemmas providing the necessary simulation results and some of their consequences.

In [2], a sequent-calculus proof system  $\mathbf{LK}_{shift}$  is presented for first-order predicate logic which admits non-elementarily smaller cut-free proofs of an adaptation of a class of theorems due to Statman [81] than the standard sequent calculus  $\mathbf{LK}$ .  $\mathbf{LK}_{shift}$  is an extension of  $\mathbf{LK}$  by the quantifier-shift rules  $r1\downarrow, r1\uparrow, r2\downarrow, r2\uparrow, r3\downarrow, r3\uparrow, r4\downarrow, r4\uparrow$  and their corresponding variants for implication  $\rightarrow$  obtained by the logical equivalence  $A \rightarrow B \equiv \bar{A} \vee B$  (see Subsection 3.3 for further discussion). The following lemma demonstrates that the cut-free open-deduction proof system with rules  $\mathbf{SKSg1} \setminus \{\forall, i\uparrow\}$  simulates cut-free  $\mathbf{LK}_{shift}$  with polynomial complexity. This fact will be used when demonstrating that the falsifier calculus admits non-elementarily smaller cut-free proofs of certain theorems than the sequent calculus  $\mathbf{LK}$  (see Corollaries 4.15 and 4.16) and also demonstrates that deep-inference proof systems for first-order predicate logic admit non-elementarily smaller cut-free proofs of certain theorems than  $\mathbf{LK}$  (see Proposition 3.9). Note that the lemma provides a simulation result for the cut-free system  $\mathbf{SKSg1} \setminus \{\forall, i\uparrow\}$  rather than  $\mathbf{SKSg1} \setminus \{i\uparrow\}$ . When considering cut-free derivations in  $\mathbf{SKSg1}$ , it is sometimes convenient to consider the set  $\mathbf{SKSg1} \setminus \{\forall, i\uparrow\}$  since, as discussed before Definition 3.7 above, instances of  $\forall$  may be used to simulate instances of  $i\uparrow$ .

**Lemma 3.8.** *For every cut-free proof  $\phi$  of a valid epsilon-free formula  $A$  in  $\mathbf{LK}_{shift}$ , there exists an epsilon-free proof of  $A$  in  $\mathbf{SKSg1} \setminus \{\forall, i\uparrow\}$  of polynomial size with respect to the size of  $\phi$ .*

*Proof.* By local simulation of inference rules. See the proof of Theorem 3.3.1 of [20] for an analogous simulation argument.  $\square$

Since  $\mathbf{LK}_{shift}$  admits non-elementarily smaller cut-free proofs of certain classes of theorems than the standard sequent calculus  $\mathbf{LK}$ , a consequence of Lemma 3.8 is that the open-deduction proof system with rules  $\mathbf{SKSg1} \setminus \{\forall, i\uparrow\}$  also admits this compression. The non-elementary compression for cut-free proofs in deep-inference proof systems which include quantifier-shifts over those which do not has already been observed in Ralph's thesis [78].

**Proposition 3.9.** *There is no elementary function bounding the size of the smallest cut-free  $\mathbf{LK}$  proof of an epsilon-free formula in terms of the size of its smallest epsilon-free proof in  $\mathbf{SKSg1} \setminus \{\forall, i\uparrow\}$ .*

*Proof.* By Theorem 3.3 of [2], there is no elementary function bounding the size of the smallest cut-free  $\mathbf{LK}$  proof of an epsilon-free formula in terms of the size of its smallest cut-free  $\mathbf{LK}_{shift}$  proof, where  $\mathbf{LK}_{shift}$  is the system presented in [2]. The result follows by Lemma 3.8.  $\square$

In [21], Brünnler presents a set of deep-inference inference rules  $\mathbf{SKSgq}$  and demonstrates that the calculus-of-structures proof system with rules  $\mathbf{SKSgq}$  is sound and complete with respect to the semantics of epsilon-free first-order predicate logic and also admits cut elimination. Since the calculus of structures formalism is a restriction of the open deduction formalism, it follows that the open-deduction proof system with rules  $\mathbf{SKSgq}$  also admits these properties. The following lemma demonstrates that the open-deduction proof system with rules  $\mathbf{SKSg1}$  polynomially simulates the open-deduction proof system with rules  $\mathbf{SKSgq}$ .

**Lemma 3.10.** *For every epsilon-free proof  $\phi$  of a formula  $A$  in  $\mathbf{SKSgq}$ , there exists an epsilon-free proof of  $A$  in  $\mathbf{SKSg1}$  of polynomial size with respect to the size of  $\phi$ .*

*Proof.* By local simulation of inference rules: for every instance  $\rho \frac{A}{B}$  of an inference rule  $\rho$  in  $\mathbf{SKSgq}$  such that  $A$  and  $B$  are epsilon-free, there exists an epsilon-free derivation

$$\frac{A}{\frac{}{B} \mathbf{SKSg1}}$$

of polynomial size with respect to  $|A| + |B|$ .  $\square$

It is also now possible to establish the completeness of the open-deduction proof system with rules  $\mathbf{SKSg1}$  with respect to the semantics of epsilon-free first-order predicate logic, as an immediate corollary of Lemma 3.10, by simulating  $\mathbf{SKSgq}$ .

**Theorem 3.11.** *Every valid epsilon-free formula has an epsilon-free proof in  $\mathbf{SKSg1}$ .*

*Proof.* This follows from Lemma 3.10, given the completeness of  $\mathbf{SKSgq}$ .  $\square$

### 3.2 Rule Decomposition

As discussed in Subsection 2.3, I have chosen to use the deep-inference formalism of open deduction in this thesis to investigate the structure of proofs due to its ability to more finely express the structure of proofs than traditional proof formalisms. One remarkable property of deep-inference formalisms such as open deduction is that the compositionality of derivations results in a finer granularity of the inference rules, in that instances of inference rules on formulae may be decomposed into derivations consisting of instances of inference rules on smaller formulae. In particular, inference rules may be decomposed to atomic form so that they act only on atomic formulae. The following lemmas provide results for decomposing the inference rules in  $\mathbf{SKSg1}$  into derivations in  $\mathbf{SKS1}$ . Similar decomposition results for

deep-inference systems for first-order predicate logic have been provided by Brünnler and Ralph [21, 78]. Note that the “decomposition” of inference rules refers to a different kind of decomposition from decomposition theorems, which are factorisation results for proofs in deep inference.

Such decomposition results will be fundamental to the proof of the central theorem of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem. Before permuting inference rules around a first-order proof to decompose it into stratified segments consisting only of certain inference rules, occurrences of inference rules in the proof are first decomposed to make the proof amenable to these permutations. Each of the following lemmas contains further information about the size and other complexity metrics of the derivations resulting from decomposing inference rules, which will be relevant when performing a complexity assessment on the Falsifier Decomposition Theorem.

The following lemma provides a decomposition of instances of the identity rule  $i\downarrow$  and its dual the cut rule  $i\uparrow$  to atomic form. Note that instances of the quantifier-shifts  $r1\downarrow$  and  $r1\uparrow$  are included in the decomposed derivation if the instance of the inference rule contains quantifiers.

**Lemma 3.12.** *For every instance of the rule  $i\downarrow$  of the form*

$$i\downarrow \frac{t}{A \vee \overline{A}}$$

*there exists a derivation*

$$\phi \parallel \frac{t}{\{ai\downarrow, s, \exists, r1\downarrow, =_P, =_\exists\}} \\ A \vee \overline{A}$$

*such that  $\phi$  does not contain any occurrences of  $=$  rules of the form  $= \frac{\exists x \exists y B}{\exists y \exists x B}$  or  $= \frac{B}{\exists x B}$  and  $|\phi| = O(|A|^2)$ . Furthermore, if  $A$  is epsilon-free then  $\phi$  may be chosen to be epsilon-free.*

*Dually, for every instance of the rule  $i\uparrow$  of the form*

$$i\uparrow \frac{A \wedge \overline{A}}{f}$$

*there exists a derivation*

$$\phi \parallel \frac{A \wedge \overline{A}}{\{ai\uparrow, s, \forall, r1\uparrow, =_P, =_\forall\}} \\ f$$

*such that  $\phi$  does not contain any occurrences of  $=$  rules of the form  $= \frac{\forall x \forall y B}{\forall y \forall x B}$  or  $= \frac{\forall x B}{B}$  and  $|\phi| = O(|A|^2)$ . Furthermore, if  $A$  is epsilon-free then  $\phi$  may be chosen to be epsilon-free.*

*Proof.* I present derivations for decomposing  $i\downarrow$  rule instances by structural induction.  $i\uparrow$  rule instances may be decomposed dually.

For all formulae  $A$ , let  $D(A)$  denote the derivation in SKS1 obtained by decomposing the rule instance  $i\downarrow \frac{t}{A \vee \overline{A}}$ . For all formulae  $A$  and  $B$ , all atomic formulae  $a$  and all variables  $x$ , we have

$$D(A \vee B) = \frac{\frac{\frac{t}{D(A) \parallel \frac{t}{A \vee \overline{A}}} \wedge \frac{t}{D(B) \parallel \frac{t}{B \vee \overline{B}}}}{\parallel \{s\}}}{(A \vee B) \vee (\overline{A} \wedge \overline{B})}$$

so that  $|D(A \vee B)| \leq k_1|A \vee B| + |D(A)| + |D(B)|$  for some constant  $k_1$ .

$$D(A \wedge B) = \frac{\frac{\frac{\mathbf{t}}{D(A)} \parallel A \vee \bar{A} \quad \frac{\mathbf{t}}{D(B)} \parallel B \vee \bar{B}}{\wedge}}{\parallel_{\{s\}} (A \wedge B) \vee (\bar{A} \vee \bar{B})}$$

so that  $|D(A \wedge B)| \leq k_2|A \wedge B| + |D(A)| + |D(B)|$  for some constant  $k_2$ .

$$D(\forall x A(x)) = \frac{\forall x \frac{\frac{\mathbf{t}}{D(A(x))} \parallel A(x) \vee \frac{\bar{A}(x)}{\exists y \bar{A}(y)}}{\forall x \downarrow}}{\frac{\forall x A(x) \vee \frac{\exists y \bar{A}(y)}{\exists x \bar{A}(x)}}{r1 \downarrow}}$$

where  $y$  is some variable term that is free for  $x$  in  $A(x)$ , so that  $|D(\forall x A(x))| \leq k_3|\forall x A(x)| + |D(A(x))|$  for some constant  $k_3$ .

$$D(\exists x A(x)) = \frac{\forall x \frac{\frac{\mathbf{t}}{D(A(x))} \parallel \frac{A(x)}{\exists y A(y)} \vee \bar{A}(x)}{\forall x \downarrow}}{\frac{\frac{\exists y A(y)}{\exists x A(x)} \vee \forall x \bar{A}(x)}{r1 \downarrow}}$$

where  $y$  is some variable term that is free for  $x$  in  $A(x)$ , so that  $|D(\exists x A(x))| \leq k_4|\exists x A(x)| + |D(A(x))|$  for some constant  $k_4$ .

$$D(\mathbf{t}) = \frac{\mathbf{t}}{\mathbf{t} \vee \mathbf{f}}$$

so that  $|D(\mathbf{t})| = 4$ .

$$D(\mathbf{f}) = \frac{\mathbf{t}}{\mathbf{f} \vee \mathbf{t}}$$

so that  $|D(\mathbf{f})| = 4$ .

$$D(a) = \text{ai} \downarrow \frac{\mathbf{t}}{a \vee \bar{a}}$$

so that  $|D(a)| = 2|a| + 2$ .

It follows by appropriate choice of constant  $K$  that for all formulae  $A$ ,  $|D(A)| \leq K|A|^2$ . The result follows.  $\square$

The following lemma provides a decomposition of instances of the contraction rule  $\mathbf{c} \downarrow$  and its dual the cocontraction rule  $\mathbf{c} \uparrow$  into derivations in **SKS1**. I keep track of the number of occurrences of the existential contraction rule  $\mathbf{qc} \downarrow$  and universal cocontraction rule  $\mathbf{qc} \uparrow$  in the



decomposed derivations, as this information will be relevant when assessing the complexity of the procedure for proving Theorem 4.13 the Falsifier Decomposition Theorem.

**Lemma 3.13.** *For every instance of the rule  $\text{c}\downarrow$  of the form*

$$\text{c}\downarrow \frac{A \vee A}{A}$$

*there exists a derivation*

$$\frac{A \vee A}{\phi \parallel \{\text{ac}\downarrow, \text{m}, \text{qc}\downarrow, \forall, =_P, =_\forall\}} A$$

*such that  $\phi$  does not contain any occurrences of  $=$  rules of the form  $\frac{\forall x \forall y B}{\forall y \forall x B}$  or  $\frac{\forall x B}{B}$ ,  $|\phi| = O(|A|^2)$ ,  $\phi$  contains at most  $n$  occurrences of  $\text{qc}\downarrow$ , where  $n$  is the number of existential quantifiers in  $A$  which occur outside the scope of epsilon symbols, and the occurrences of  $\text{qc}\downarrow$  in  $\phi$  occur in parallel in  $\phi$ . Furthermore, if  $A$  is epsilon-free then  $\phi$  may be chosen to be epsilon-free.*

*Dually, for every instance of the rule  $\text{c}\uparrow$  of the form*

$$\text{c}\uparrow \frac{A}{A \wedge A}$$

*there exists a derivation*

$$\frac{A}{\phi \parallel \{\text{ac}\uparrow, \text{m}, \text{qc}\uparrow, \exists, =_P, =_\exists\}} A \wedge A$$

*such that  $\phi$  does not contain any occurrences of  $=$  rules of the form  $\frac{\exists x \exists y B}{\exists y \exists x B}$  or  $\frac{B}{\exists x B}$ , does not contain any occurrences of  $\exists$  or  $=_\exists$  if  $A$  is weakly existential-free,  $|\phi| = O(|A|^2)$ ,  $\phi$  contains at most  $n$  occurrences of  $\text{qc}\uparrow$ , where  $n$  is the number of universal quantifiers in  $A$  which occur outside the scope of epsilon symbols, and the occurrences of  $\text{qc}\uparrow$  in  $\phi$  occur in parallel in  $\phi$ . Furthermore, if  $A$  is epsilon-free then  $\phi$  may be chosen to be epsilon-free.*

*Proof.* I present derivations for decomposing  $\text{c}\downarrow$  rule instances by structural induction.  $\text{c}\uparrow$  rule instances may be decomposed dually.

For all formulae  $A$ , let  $D(A)$  denote the derivation in SKS1 obtained by decomposing the rule instance  $\text{c}\downarrow \frac{A \vee A}{A}$ . For all formulae  $A$  and  $B$ , all atomic formulae  $a$  and all variables  $x$ , we have

$$D(A \vee B) = \frac{(A \vee B) \vee (A \vee B)}{\boxed{\boxed{\frac{A \vee A}{D(A) \parallel A} \vee \frac{B \vee B}{D(B) \parallel B}}}}$$

so that  $|D(A \vee B)| \leq k_1|A \vee B| + |D(A)| + |D(B)|$  for some constant  $k_1$  and  $D(A \vee B)$  contains  $n+m$  occurrences of  $\text{qc}\downarrow$ , where  $n$  is the number of occurrences of  $\text{qc}\downarrow$  in  $D(A)$  and  $m$  is the number of occurrences of  $\text{qc}\downarrow$  in  $D(B)$ . It also follows from the derivation above that if all occurrences of  $\text{qc}\downarrow$  in  $D(A)$  occur in parallel in  $D(A)$  and all occurrences of  $\text{qc}\downarrow$  in  $D(B)$  occur in parallel in  $D(B)$ , then all occurrences of  $\text{qc}\downarrow$  in  $D(A \vee B)$  occur in parallel in  $D(A \vee B)$ .

$$D(A \wedge B) = \frac{(A \wedge B) \vee (A \wedge B)}{\boxed{\boxed{\frac{A \vee A}{D(A) \parallel A} \wedge \frac{B \vee B}{D(B) \parallel B}}}}$$

so that  $|D(A \wedge B)| = 2|A \wedge B| + 2 + |D(A)| + |D(B)|$  and  $D(A \wedge B)$  contains  $n + m$  occurrences of  $\mathbf{qc}\downarrow$ , where  $n$  is the number of occurrences of  $\mathbf{qc}\downarrow$  in  $D(A)$  and  $m$  is the number of occurrences of  $\mathbf{qc}\downarrow$  in  $D(B)$ . It also follows from the derivation above that if all occurrences of  $\mathbf{qc}\downarrow$  in  $D(A)$  occur in parallel in  $D(A)$  and all occurrences of  $\mathbf{qc}\downarrow$  in  $D(B)$  occur in parallel in  $D(B)$ , then all occurrences of  $\mathbf{qc}\downarrow$  in  $D(A \wedge B)$  occur in parallel in  $D(A \wedge B)$ .

$$D(\forall xA) = \frac{\forall xA \vee \forall xA}{\forall x \left( \frac{\frac{\forall y(A[y/x])}{\forall} A \vee \frac{\forall y(A[y/x])}{\forall} A}{D(A) \parallel A} \right)}$$

where  $y$  is some variable term that is free for  $x$  in  $A$ , so that  $|D(\forall xA)| \leq k_2|\forall xA| + |D(A)|$  for some constant  $k_2$  and  $D(\forall xA)$  contains the same number of occurrences of  $\mathbf{qc}\downarrow$  as  $D(A)$ . It also follows from the derivation above that if all occurrences of  $\mathbf{qc}\downarrow$  in  $D(A)$  occur in parallel in  $D(A)$ , then all occurrences of  $\mathbf{qc}\downarrow$  in  $D(\forall xA)$  occur in parallel in  $D(\forall xA)$ .

$$D(\exists xA) = \mathbf{qc}\downarrow \frac{\exists xA \vee \exists xA}{\exists xA}$$

so that  $|D(\exists xA)| = 3|\exists xA|$  and  $D(\exists xA)$  contains precisely one occurrence of  $\mathbf{qc}\downarrow$ .

$$D(\mathbf{t}) = \frac{\mathbf{t} \vee \mathbf{t}}{\mathbf{t}}$$

so that  $|D(\mathbf{t})| = 4$  and  $D(\mathbf{t})$  contains no occurrences of  $\mathbf{qc}\downarrow$ .

$$D(\mathbf{f}) = \frac{\mathbf{f} \vee \mathbf{f}}{\mathbf{f}}$$

so that  $|D(\mathbf{f})| = 4$  and  $D(\mathbf{f})$  contains no occurrences of  $\mathbf{qc}\downarrow$ .

$$D(a) = \mathbf{ac}\downarrow \frac{a \vee a}{a}$$

so that  $|D(a)| = 3|a| + 1$  and  $D(a)$  contains no occurrences of  $\mathbf{qc}\downarrow$ .

It follows by appropriate choice of constant  $K$  that for all formulae  $A$ ,  $|D(A)| \leq K|A|^2$  and that  $D(A)$  contains at most  $n$  occurrences of  $\mathbf{qc}\downarrow$ , where  $n$  is the number of existential quantifiers in  $A$  which occur outside the scope of epsilon symbols. Furthermore, all occurrences of  $\mathbf{qc}\downarrow$  in  $D(A)$  occur in parallel in  $D(A)$ . The result follows.  $\square$

The following lemma provides a decomposition of instances of the weakening rule  $\mathbf{w}\downarrow$  and its dual the coweakening rule  $\mathbf{w}\uparrow$  into derivations in **SKS1**.

**Lemma 3.14.** *For every instance of the rule  $\mathbf{w}\downarrow$  of the form*

$$\mathbf{w}\downarrow \frac{\mathbf{f}}{A}$$

*there exists a derivation*

$$\frac{\mathbf{f}}{\phi \parallel \{ \mathbf{aw}\downarrow, \exists, =_P, =_V \}} \frac{}{A}$$

*such that  $\phi$  does not contain any occurrences of  $=$  rules of the form  $\frac{\forall x \forall y B}{\forall y \forall x B}, \frac{\forall x B}{B}$  and  $|\phi| = O(|A|)$ . Furthermore, if  $A$  is epsilon-free then  $\phi$  may be chosen to be epsilon-free.*

*Dually, for every instance of the rule  $\mathbf{w}\uparrow$  of the form*

$$\mathbf{w}\uparrow \frac{A}{\mathbf{t}}$$

there exists a derivation

$$\frac{A}{\frac{\frac{}{t}}{\{aw\uparrow, \forall, =_P, =_\exists\}}}$$

such that  $\phi$  does not contain any occurrences of  $=$  rules of the form  $= \frac{\exists x \exists y B}{\exists y \exists x B}$  or  $= \frac{B}{\exists x B}$ , does not contain any occurrences of  $=_\exists$  if  $A$  is weakly existential-free and  $|\phi| = O(|A|)$ . Furthermore, if  $A$  is epsilon-free then  $\phi$  may be chosen to be epsilon-free.

*Proof.* I present derivations for decomposing  $w\downarrow$  rule instances by structural induction.  $w\uparrow$  rule instances may be decomposed dually.

For all formulae  $A$ , let  $D(A)$  denote the derivation in SKS1 obtained by decomposing the rule instance  $w\downarrow \frac{f}{A}$ . For all formulae  $A$  and  $B$ , all atomic formulae  $a$  and all variables  $x$ , we have

$$D(A \vee B) = \frac{f}{\frac{\frac{f}{D(A) \parallel A} \vee \frac{f}{D(B) \parallel B}}$$

so that  $|D(A \vee B)| = 2 + |D(A)| + |D(B)|$ .

$$D(A \wedge B) = \frac{f}{\frac{\frac{f}{D(A) \parallel A} \wedge \frac{f}{D(B) \parallel B}}$$

so that  $|D(A \wedge B)| = 2 + |D(A)| + |D(B)|$ .

$$D(\forall x A) = \frac{f}{\forall x \frac{f}{D(A) \parallel A}}$$

so that  $|D(\forall x A)| = 2 + |D(A)|$ .

$$D(\exists x A) = \frac{f}{\exists x \frac{f}{D(A) \parallel A}}$$

so that  $|D(\exists x A)| = 2 + |D(A)|$ .

$$D(t) = aw\downarrow \frac{f}{t}$$

so that  $|D(t)| = 2$ .

$$D(f) = f$$

so that  $|D(f)| = 1$ .

$$D(a) = aw\downarrow \frac{f}{a}$$

so that  $|D(a)| = 1 + |a|$ .

It follows by appropriate choice of constant  $K$  that for all formulae  $A$ ,  $|D(A)| \leq K|A|$ . The result follows.  $\square$

The following lemma provides decompositions of instances of the equality rules which are contained in **SKSg1** but not **SKS1** into derivations in **SKS1**.

**Lemma 3.15.** *For every instance*

$$\begin{array}{c} \frac{A}{B} \\ \text{of an inference rule of the form } = \frac{\exists x \exists y C}{\exists y \exists x C}, = \frac{\forall x \forall y C}{\forall y \forall x C}, = \frac{C}{\exists x C} \text{ or } = \frac{\forall x C}{C}, \text{ there exists a deriva-} \\ \text{tion} \\ \frac{A}{\phi \parallel \{\exists, \forall, =\exists, =\forall\}} \\ B \end{array}$$

such that  $\phi$  does not contain any occurrences of  $=$  rules of the form  $= \frac{\exists x \exists y C}{\exists y \exists x C}, = \frac{\forall x \forall y C}{\forall y \forall x C}, = \frac{C}{\exists x C}$  or  $= \frac{\forall x C}{C}$  and  $|\phi| = O(|A|)$ . Furthermore, if  $A$  is epsilon-free then  $\phi$  may be chosen to be epsilon-free.

*Proof.* I present derivations for each inference rule.

$$\begin{array}{ccc} = \frac{\exists x \exists y C}{\exists y \exists x C} & \rightarrow & = \frac{\boxed{\begin{array}{c} \exists x \exists y \quad \boxed{\frac{C}{\exists x C}} \end{array}}}{\exists y \exists x C} \\ \\ = \frac{\forall x \forall y C}{\forall y \forall x C} & \rightarrow & = \frac{\boxed{\begin{array}{c} \forall y \forall x \quad \boxed{\frac{\forall y C}{C}} \end{array}}}{\forall x \forall y C} \\ \\ = \frac{C}{\exists x C} & \rightarrow & \exists \frac{C}{\exists x C} \\ \\ = \frac{\forall x C}{C} & \rightarrow & \forall \frac{\forall x C}{C} \end{array}$$

□

Remarkably, in a deep-inference setting, most of the quantifier-shift rules are trivial, in that they may be decomposed into derivations in **SKS1** with linear complexity. The following proposition demonstrates that, with the exceptions of  $r1\downarrow$  and  $r1\uparrow$ , instances of the quantifier-shifts in **SKSg1** may be decomposed into derivations in  $\mathbf{SKS1} \setminus \{r1\downarrow, r1\uparrow\}$  with linear complexity. This resembles the equivalence of quantifier-shifts and traditional first-order deep-inference rules that is observed in Ralph's thesis [78]. The consequences of this fact and the properties of the non-trivial quantifier-shifts  $r1\downarrow$  and  $r1\uparrow$  will be explored in more detail in the following two subsections, Subsections 3.3 and 3.4.

**Proposition 3.16** (Decomposition of quantifier-shifts). *For every instance  $\rho \frac{A}{B}$  of an inference rule  $\rho \in \{r2\downarrow, r2\uparrow, r3\downarrow, r3\uparrow, r4\downarrow, r4\uparrow\}$ , there exists a derivation*

$$\frac{A}{\phi \parallel \{\mathbf{aw}\downarrow, \mathbf{aw}\uparrow, \mathbf{qc}\downarrow, \mathbf{qc}\uparrow, \exists, \forall, =\mathbf{p}, =\exists, =\forall\}} \\ B$$

such that  $\phi$  does not contain any occurrences of  $=$  rules of the form  $= \frac{\exists x \exists y C}{\exists y \exists x C}, = \frac{\forall x \forall y C}{\forall y \forall x C}, = \frac{C}{\exists x C}$  or  $= \frac{\forall x C}{C}$  and  $|\phi| = O(|A|)$ . Furthermore, if  $A$  is epsilon-free then  $\phi$  may be chosen to be epsilon-free.

*Proof.* I present derivations for  $r2\downarrow$ ,  $r3\downarrow$  and  $r4\downarrow$ . The remaining rules may be derived dually.

$$r2\downarrow \frac{\forall x(A(x) \wedge B)}{\forall xA(x) \wedge B} \rightarrow \begin{array}{c} \text{qc}\uparrow \\ \boxed{\begin{array}{c} \forall x(A(x) \wedge B) \\ \hline \forall x \left( A(x) \wedge \boxed{\text{w}\uparrow \frac{B}{t}} \right) \wedge \forall \left( \boxed{\text{w}\uparrow \frac{A(x)}{t}} \wedge B \right) \\ \hline \forall xA(x) \wedge B \end{array}} \end{array}$$

where the occurrences of  $\text{w}\uparrow$  in the derivation above are replaced with derivations in  $\{\text{aw}\uparrow, \forall, =_{\mathbf{P}}, =_{\exists}\}$  using Lemma 3.14 and sequential composition.

$$r3\downarrow \frac{\exists x(A(x) \vee B)}{\exists xA(x) \vee B} \rightarrow \begin{array}{c} \exists x \left( \boxed{\exists \frac{A(x)}{\exists yA(y)}} \vee B \right) \\ \hline \exists xA(x) \vee B \end{array}$$

where  $y$  is some variable that is free for  $x$  in  $A(x)$ .

$$r4\downarrow \frac{\exists x(A(x) \wedge B)}{\exists xA(x) \wedge B} \rightarrow \begin{array}{c} \exists x \left( \boxed{\exists \frac{A(x)}{\exists yA(y)}} \wedge B \right) \\ \hline \exists xA(x) \wedge B \end{array}$$

where  $y$  is some variable that is free for  $x$  in  $A(x)$ .

Observe that the size of each construction is linear with respect to the size of the premise of each inference rule. The result follows.  $\square$

The following lemma combines the above results to provide a full decomposition from derivations in  $\text{SKSg1}$  to derivations in  $\text{SKS1}$ .

**Lemma 3.17.** *For every derivation  $\phi$  with premise  $A$  and conclusion  $B$  in  $\text{SKSg1}$ , there exists a proof  $\phi'$  with premise  $A$  and conclusion  $B$  in  $\text{SKS1}$  such that  $|\phi'| = O(|\phi|^2)$ . Furthermore, if  $\phi$  is epsilon-free then  $\phi'$  may be chosen to be epsilon-free and if  $\phi$  is cut-free then  $\phi'$  may be chosen to be cut-free.*

*Proof.* By replacing all occurrences of inference rules in  $\text{SKSg1} \setminus \text{SKS1}$  in  $\phi$  with derivations in  $\text{SKS1}$  using sequential composition, by Lemmas 3.12, 3.13, 3.14 and 3.15 and Proposition 3.16.  $\square$

### 3.3 Quantifier-Shifts and Non-Elementary Compression

Of particular interest in the study of first-order proof complexity and normalisation are the inference rules in  $\text{SKSg1}$  known as *quantifier-shifts*, which are given by

$$\begin{array}{cccc} r1\downarrow \frac{\forall x(A(x) \vee B)}{\forall xA(x) \vee B} & r2\downarrow \frac{\forall x(A(x) \wedge B)}{\forall xA(x) \wedge B} & r3\downarrow \frac{\exists x(A(x) \vee B)}{\exists xA(x) \vee B} & r4\downarrow \frac{\exists x(A(x) \wedge B)}{\exists xA(x) \wedge B} \\ r1\uparrow \frac{\exists xA(x) \wedge B}{\exists x(A(x) \wedge B)} & r2\uparrow \frac{\exists xA(x) \vee B}{\exists x(A(x) \vee B)} & r3\uparrow \frac{\forall xA(x) \wedge B}{\forall x(A(x) \wedge B)} & r4\uparrow \frac{\forall xA(x) \vee B}{\forall x(A(x) \vee B)} \end{array}$$

where  $x$  does not occur free in  $B$ .

The quantifier-shifts were first introduced in Herbrand's thesis in 1929 [54] to be used as part of his proof of Herbrand's Theorem under the name of "rules of passage". Herbrand's thesis also included quantifier-shifts for shifting negation symbols  $\neg$  inside and outside of quantifiers, given by

$$\boxed{\begin{array}{cccc} \frac{\neg\forall xA}{\exists x\neg A} & \frac{\neg\exists xA}{\forall x\neg A} & \frac{\exists x\neg A}{\neg\forall xA} & \frac{\forall x\neg A}{\neg\exists xA} \end{array}}$$

and it is sometimes also convenient to consider variants of the quantifier-shifts for implication  $\rightarrow$  which are generated by the classical logical equivalence  $A \rightarrow B \equiv \overline{A} \vee B$  and are given by

$$\boxed{\begin{array}{cccc} \frac{\forall x(A(x) \rightarrow B)}{\exists xA(x) \rightarrow B} & \frac{\exists x(A(x) \rightarrow B)}{\forall xA(x) \rightarrow B} & \frac{\forall x(B \rightarrow A(x))}{B \rightarrow \forall xA(x)} & \frac{\exists x(B \rightarrow A(x))}{B \rightarrow \exists xA(x)} \\ \frac{\exists xA(x) \rightarrow B}{\forall x(A(x) \rightarrow B)} & \frac{\forall xA(x) \rightarrow B}{\exists x(A(x) \rightarrow B)} & \frac{B \rightarrow \forall xA(x)}{\forall x(B \rightarrow A(x))} & \frac{B \rightarrow \exists xA(x)}{\exists x(B \rightarrow A(x))} \end{array}}$$

where  $x$  does not occur free in  $B$ . The quantifier-shifts were in fact the source of an error in Herbrand's thesis, which was discovered and corrected by Aanderaa, Andrews and Dreben in 1963 [37].

A remarkable phenomenon in first-order proof theory is that some proof systems admit non-elementarily smaller cut-free proofs of certain classes of theorems than other proof systems. In [81], Statman provides a class  $S$  of first-order theorems for which there is no elementary function bounding the size of the smallest Herbrand disjunction for a formula in  $S$  in terms of the size of its smallest proof in Gentzen's sequent calculus **LK**. In [2], Aguilera and Baaz define a sequent-calculus proof system **LK<sub>shift</sub>** which is obtained by extending **LK** by the quantifier-shift rules in **SKSg1** and those for implication and demonstrate that for a class  $S'$  of first-order theorems obtained by adapting  $S$ , there is no elementary function bounding the size of the smallest cut-free **LK** proof of a formula in  $S'$  in terms of the size of its smallest cut-free **LK<sub>shift</sub>** proof. Notably, the sequent-calculus-based epsilon-calculus also admits this compression over the sequent calculus for cut-free proofs [17] since critical axioms can simulate quantifier-shifts in the presence of the encodings of quantifiers by  $\varepsilon$ -terms. Since the quantifier-shifts involve rewriting inside a formula, they are natural deep-inference inference rules and it follows that deep-inference proof systems for first-order predicate logic admit non-elementarily smaller cut-free proofs than **LK** for certain classes of theorems (see Proposition 3.9). The non-elementary compression of cut-free proofs in deep inference arising from the quantifier-shifts is also discussed in Ralph's thesis [78]. Given the finer granularity of inference rules and more flexible composition mechanism exhibited by open deduction over the sequent calculus, it is natural to consider what properties of the quantifier-shifts may be revealed by analysing them in the open deduction formalism and what insights this might provide into the properties which make a proof system admit the non-elementary compression over the sequent calculus for cut-free proofs.

As was demonstrated by Proposition 3.16, most of the quantifier-shifts are trivial in open deduction, in that they may be decomposed into derivations in **SKS1**  $\setminus \{\mathbf{r1}\downarrow, \mathbf{r1}\uparrow\}$  with linear complexity. The quantifier-shifts  $\mathbf{r1}\downarrow$  and  $\mathbf{r1}\uparrow$ , however, are more delicate and do not admit such decompositions, warranting a closer analysis of these rules and the role they play in the phenomenon of the non-elementary compression of cut-free proofs. As was noted in Lemma 3.12, instances of the identity rule  $\mathbf{i}\downarrow$  may be decomposed into derivations in  $\{\mathbf{ai}\downarrow, \mathbf{s}, \exists, \mathbf{r1}\downarrow, =\mathbf{p}, =\exists, =\forall\}$  and, dually, instances of the cut rule  $\mathbf{i}\uparrow$  may be decomposed into derivations in  $\{\mathbf{ai}\uparrow, \mathbf{s}, \forall, \mathbf{r1}\uparrow, =\mathbf{p}, =\exists, =\forall\}$ , establishing connections between the identity rule  $\mathbf{i}\downarrow$  and the quantifier-shift  $\mathbf{r1}\downarrow$  and between the cut rule  $\mathbf{i}\uparrow$  and the quantifier-shift  $\mathbf{r1}\uparrow$ . More specifically, an instance of the identity rule  $\mathbf{i}\downarrow$  on a quantified formula  $\forall xA(x)$  may

be decomposed into an instance of the  $r1\downarrow$  quantifier-shift, an instance of the existential witnessing rule  $\exists$  and an instance of the identity rule  $i\downarrow$  on the formula  $A(x)$ :

$$i\downarrow \frac{t}{\forall x A(x) \vee \exists x \bar{A}(x)} \rightarrow \begin{array}{c} = \frac{t}{\forall x \left( \begin{array}{c} i\downarrow \frac{t}{A(x) \vee \exists y \bar{A}(y)} \end{array} \right)} \\ r1\downarrow \frac{\forall x A(x) \vee \frac{\exists y \bar{A}(y)}{\exists x \bar{A}(x)}}{\forall x A(x) \vee \exists x \bar{A}(x)} \end{array}$$

and, dually, an instance of the cut rule  $i\uparrow$  on a quantified formula  $\forall x A(x)$  may be decomposed into an instance of the  $r1\uparrow$  quantifier-shift, an instance of the universal instantiation rule  $\forall$  and an instance of the cut rule  $i\uparrow$  on the formula  $A(x)$ :

$$i\uparrow \frac{\forall x A(x) \wedge \exists x \bar{A}(x)}{f} \rightarrow \begin{array}{c} \frac{\frac{\forall x A(x)}{\forall y A(y)} \wedge \exists x \bar{A}(x)}{r1\uparrow} \\ \frac{\forall y A(y) \wedge \bar{A}(x)}{\exists x \left( \frac{\forall y A(y)}{A(x)} \wedge \bar{A}(x) \right)} \\ i\uparrow \frac{f}{f} \\ = f \end{array}$$

Given the fundamental roles of the identity rule  $i\downarrow$  and the cut rule  $i\uparrow$  in the proof theory of first-order predicate logic, these decomposition results suggest that the quantifier-shifts  $r1\downarrow$  and  $r1\uparrow$  share similarly important roles, hence why they do not admit trivial decompositions like the other quantifier-shifts. Indeed, this connects the elimination of occurrences of the quantifier-shifts  $r1\downarrow$  and  $r1\uparrow$  from a first-order proof to the non-trivial phenomenon of cut elimination and, consequently, Herbrand's Theorem. Both cut elimination [31] and Herbrand's Theorem [81] are associated with non-elementary blowups in proof complexity, suggesting a fundamental connection between these theorems, the non-trivial quantifier-shifts  $r1\downarrow$  and  $r1\uparrow$  and the phenomenon of non-elementary proof compression. As will be discussed in the following subsection, Subsection 3.4, the quantifier-shifts  $r1\downarrow$  and  $r1\uparrow$  also present certain challenges for the normalisation of first-order proofs when permuting occurrences of the existential contraction rule  $qc\downarrow$  down a proof and occurrences of the universal cocontraction rule  $qc\uparrow$  up a proof.

Interestingly, the quantifier-shift  $r1\downarrow$ , which is given by

$$r1\downarrow \frac{\forall x (A(x) \vee B)}{\forall x A(x) \vee B}$$

where  $x$  does not occur free in  $B$ , is also the only quantifier-shift in  $SKSg1$  which is not intuitionistically valid, since  $B$  may contain existential quantifiers which are witnessed by terms which depend on  $x$ . The role of by quantifier-shifts in intuitionistic and intermediate logics has recently been studied by Baaz, Gamsakhurdia, Iemhoff and Jalali, who have demonstrated that an intermediate logic admits Skolemisation if and only if it admits all of the quantifier-shifts (including those for implication) [12].

### 3.4 Case Analysis Extraction and Herbrand's Theorem

In the proof theory of first-order predicate logic, contractions on existential formulae may be understood as case analyses on the witnesses to the existential quantifiers in the premise of the contraction:

$$\text{qc}\downarrow \frac{\exists x A \vee \exists x A}{\exists x A}$$

the witness to the existential quantifier  $\exists x$  in the conclusion may be understood as a case analysis on the witness to the existential quantifier  $\exists x$  on the left side of the disjunction in the premise and the witness to the existential quantifier  $\exists x$  on the right side of the disjunction in the premise. Dually, cocontractions on universal formulae

$$\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}$$

may also be understood as case analyses on the terms that instantiate the universal quantifiers in the conclusion later in a proof. A semantically natural operation to perform on a first-order proof is thus to extract the case analyses contained within such quantifier contraction rules from the proof, deriving a disjunction of terms which witness the existential quantifiers in the conclusion. This is one of the essential notions of *Herbrand's Theorem* [54], a fundamental theorem of classical proof theory, and the propositional disjunction collecting the term witnesses is called a *Herbrand disjunction* for the conclusion of the proof.

In the sequent calculus, Herbrand's Theorem is traditionally proved as a corollary to cut elimination, such as in an exposition of a proof of the general version of the theorem due to Buss [30, 31] (with a correction due to McKinley [68]). In “Buss-style” proofs of Herbrand's Theorem, cut rules are first eliminated from a sequent-calculus proof to make the proof amenable to rule permutations. Existential witnessing and existential contraction rules are then permuted down the cut-free proof to yield a propositional proof of a Herbrand disjunction for the conclusion of the proof.

In a deep-inference setting, Brünnler [21] has presented a statement and proof of the general version of Herbrand's Theorem in the form of a decomposition theorem which does not require cuts to be eliminated from the proof, establishing a kind of independence between cut elimination and Herbrand's Theorem in deep inference which is not observable in the sequent calculus. Brünnler presents a procedure which transforms a first-order deep-inference proof into a factorised proof of the form

$$\begin{array}{c} \forall x_1 \dots \forall x_n \boxed{\begin{array}{c} \prod \text{Propositional rules (SKSg}_P\text{)} \\ A' \end{array}} \\ \parallel \{\exists\} \\ A'' \\ \parallel \{r1\downarrow, r2\downarrow, r3\downarrow, r4\downarrow\} \\ A''' \\ \parallel \{\text{qc}\downarrow\} \\ A \end{array}$$

where  $A''$  is in prenex normal form, the formula  $\forall x_1 \dots \forall x_n A'$  is a Herbrand disjunction for  $A$  and the above proof is said to be in *Herbrand normal form*.

As was discussed in the previous subsection, Subsection 3.3, Statman [81] has shown that, in general, there is no elementary bound on the size of the smallest Herbrand disjunction for a first-order theorem in terms of the size of its smallest proof, which is related to the non-elementary compression of cut-free proofs yielded by proof systems which include quantifier-shifts. A natural question then is whether it is possible to extract the case analyses contained in quantifier contraction rules from a first-order proof without producing the non-elementary blowups incurred by Herbrand's Theorem. This is possible using the central proof system of this thesis, the falsifier calculus, and is one of the essential notions of its corresponding



decomposition theorem, Theorem 4.13 the Falsifier Decomposition Theorem, both of which will be introduced and explored in the following section, Section 4. I will first, however, consider this problem in the setting of the standard syntax of predicate logic and illustrate why such case analysis extraction with elementary complexity is not possible in standard first-order deep-inference proof systems such as SKSg1.

In a deep-inference setting, to extract the case analyses contained within occurrences of the existential contraction rule  $\text{qc}\downarrow$  from a first-order proof, the rule occurrences may be permuted down the proof by recursively permuting them down through the rule occurrences immediately below them until they reach the bottom of the proof. However, when an occurrence of  $\text{qc}\downarrow$  is permuted down through certain occurrences of the  $\text{r1}\uparrow$  quantifier-shift, an occurrence of the universal cocontraction rule  $\text{qc}\uparrow$ , the dual of  $\text{qc}\downarrow$ , is introduced into the proof, as follows:

$$\begin{array}{c}
\boxed{\boxed{\text{qc}\downarrow \frac{\exists x A \vee \exists x A}{\exists x A}} \wedge \forall y B} \\
\text{r1}\uparrow \overline{\overline{\exists x(A \wedge \forall y B)}}
\end{array}
\rightarrow
\begin{array}{c}
\boxed{(\exists x A \vee \exists x A) \wedge \boxed{\text{qc}\uparrow \frac{\forall y B}{\forall y B \wedge \forall y B}}} \\
= \\
\boxed{\forall y B \wedge \boxed{s \frac{\forall y B \wedge (\exists x A \vee \exists x A)}{(\exists x A \wedge \forall y B) \vee \exists x A}}} \\
s \overline{\overline{\boxed{\boxed{\text{r1}\uparrow \frac{\exists x A \wedge \forall y B}{\exists x(A \wedge \forall y B)}} \vee \boxed{\text{r1}\uparrow \frac{\exists x A \wedge \forall y B}{\exists x(A \wedge \forall y B)}}}} \\
\text{qc}\downarrow \overline{\overline{\exists x(A \wedge \forall y B)}}
\end{array}$$

Dually, when an occurrence of  $\text{qc}\uparrow$  is permuted up through certain occurrences of the  $\text{r1}\downarrow$  quantifier-shift, an occurrence of  $\text{qc}\downarrow$  is introduced into the proof, as follows:

$$\begin{array}{c}
\forall x(A \vee \exists y B) \\
\text{r1}\downarrow \overline{\overline{\boxed{\boxed{\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}} \vee \exists y B}}}
\end{array}
\rightarrow
\begin{array}{c}
\text{qc}\uparrow \overline{\overline{\forall x(A \vee \exists y B)}} \\
\boxed{\boxed{\text{r1}\downarrow \frac{\forall x(A \vee \exists y B)}{\forall x A \vee \exists y B}} \wedge \boxed{\text{r1}\downarrow \frac{\forall x(A \vee \exists y B)}{\forall x A \vee \exists y B}}} \\
s \overline{\overline{\boxed{s \frac{(\forall x A \vee \exists y B) \wedge \forall x A}{(\forall x A \wedge \forall x A) \vee \exists y B}} \vee \exists y B}} \\
= \\
\boxed{(\forall x A \wedge \forall x A) \vee \boxed{\text{qc}\downarrow \frac{\exists y B \vee \exists y B}{\exists y B}}}
\end{array}
\tag{2}$$

Consequently, a procedure which successively permutes occurrences of  $\text{qc}\downarrow$  down and occurrences of  $\text{qc}\uparrow$  up a first-order proof is non-terminating in the standard syntax of predicate

logic. Consider, for instance, a proof which contains the following subderivation:

$$\frac{\begin{array}{c} \text{r1}\downarrow \\ \boxed{\begin{array}{c} \forall x(A \vee \exists yB) \\ \boxed{\text{qc}\uparrow \frac{\forall xA}{\forall xA \wedge \forall xA}} \vee \exists yB \end{array}} \wedge C \end{array}}{\text{s} \overline{\frac{(\forall xA \wedge \forall xA) \vee (\exists yB \wedge C)}{\text{m} \overline{\begin{array}{c} \text{r2}\uparrow \frac{\forall xA \vee \exists yB}{\exists y(\forall xA \vee B)} \wedge (\forall xA \vee C) \end{array}}} \text{r1}\uparrow \overline{\exists y((\forall xA \vee B) \wedge (\forall xA \vee C))}}}$$

When the occurrence of  $\mathbf{qc}\uparrow$  is permuted up through the occurrence of  $\mathbf{r1}\downarrow$  immediately above it, an occurrence of the existential contraction rule  $\mathbf{qc}\downarrow$  on the formula  $\exists yB$  is introduced into the proof. If this occurrence of  $\mathbf{qc}\downarrow$  is subsequently permuted down the proof, when it is permuted down through the occurrence of  $\mathbf{r1}\downarrow$  at the bottom of the subderivation, a further occurrence of  $\mathbf{qc}\uparrow$  is introduced into the proof. In this manner, each phase of permuting occurrences of  $\mathbf{qc}\uparrow$  up the proof will introduce further occurrences of  $\mathbf{qc}\downarrow$  into the proof and each phase of permuting occurrences of  $\mathbf{qc}\downarrow$  down the proof will introduce further occurrences of  $\mathbf{qc}\uparrow$  into the proof, resulting in non-termination.

The following section, Section 4, introduces a proof system which is able to overcome these difficulties and attain termination of the procedure described above through the use of an inference rule which generalises the quantifier-shift  $\mathbf{r1}\downarrow$  and introduces  $\varepsilon$ -terms into proofs. Through a certain normalisation, this system is able to avoid introducing unnecessary occurrences of  $\mathbf{qc}\downarrow$  into the proof during this procedure and instead introduces contractions on  $\varepsilon$ -terms, yielding termination. The resultant decomposition theorem is the central result of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem, which is able to extract the case analyses contained in quantifier contraction rules from a first-order proof with elementary complexity due to the presence of  $\varepsilon$ -terms in the corresponding decomposed proofs.

## 4 The Falsifier Calculus

Logic takes care of itself; all we have to do is to look and see how it does it.

---

*L. Wittgenstein*

The traditional epsilon-calculus as studied in Hilbert-Frege systems and the sequent calculus admits several notable and useful properties, such as admitting non-elementarily smaller cut-free proofs of certain classes of theorems than traditional proof systems [17] and a rich proof theory, as encapsulated in the epsilon theorems [75]. However, the traditional epsilon-calculus is notoriously unwieldy and these properties rely on the use of cumbersome encodings of quantifiers in proofs by  $\varepsilon$ -terms. In this section I will introduce the *falsifier calculus*, an open-deduction proof system for first-order predicate logic in the language of the epsilon-calculus which exhibits the desirable properties of the traditional epsilon-calculus while simplifying some of its more unwieldy aspects and providing new insights into the structure and complexity of first-order proofs. The falsifier calculus uses a new inference rule to introduce  $\varepsilon$ -terms into proofs, called the *falsifier rule  $\varepsilon$* , which is given by

$$\varepsilon \frac{\forall x(A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))}$$

where  $y$  is a variable such that  $y$  is free for  $x$  in  $A(x)$  and  $\varepsilon_y \bar{A}(y)$  is free for  $x$  in  $B(x)$ , and is distinct from the critical axioms used to introduce  $\varepsilon$ -terms into proofs in the traditional epsilon-calculus. Through the use of the falsifier rule  $\varepsilon$ , the falsifier calculus admits the non-elementary compression of cut-free proofs over the sequent calculus that is exhibited by the traditional epsilon-calculus, but without the need to encode quantifiers in proofs by  $\varepsilon$ -terms. The falsifier rule  $\varepsilon$  is a generalisation of the quantifier-shift  $r1\downarrow$ , providing a new perspective on the non-elementary compression of cut-free proofs yielded by quantifier-shifts. Due to the presence of the falsifier rule  $\varepsilon$  and the more expressive syntax of the epsilon-calculus, the falsifier calculus is also able to attain termination of the case analysis extraction procedure described in Subsection 3.4, yielding a new decomposition theorem for proofs in first-order predicate logic and the central result of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem. The Falsifier Decomposition Theorem decomposes a first-order proof into an upper segment of weakly existential-free formulae in the falsifier calculus and a lower segment of existential witnesses and existential contractions in  $\{\exists, qc\downarrow\}$ . The Falsifier Decomposition Theorem can be understood as an analogue to Herbrand's Theorem in which the upper segments of the corresponding proofs may be non-elementarily smaller than the corresponding smallest proofs in Herbrand normal form, providing insight into the role of  $\varepsilon$ -terms in the non-elementary compression of cut-free proofs.

In Subsection 4.1 I illustrate how the falsifier rule  $\varepsilon$  arises from an analysis of the quantifier-shift rules and use it to define the falsifier calculus  $SKSg\varepsilon$ . In Subsection 4.2 I state the main result of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem, which is a decomposition theorem for proofs in first-order predicate logic using the falsifier calculus  $SKSg\varepsilon$  and arises from attaining termination of the case analysis extraction procedure described in Subsection 3.4 using the falsifier rule  $\varepsilon$ . In Subsection 4.3 I provide examples of proofs which are in the normal form corresponding to this decomposition theorem and compare them with proofs in other systems. In Subsection 4.4 I compare the falsifier calculus  $SKSg\varepsilon$  and Theorem 4.13 the Falsifier Decomposition Theorem to other proof systems and proof interpretations for first-order predicate logic in the literature, namely Herbrand's Theorem, the traditional epsilon-calculus and Kreisel's no-counterexample interpretation.

I will first consider the problem of whether the compositionality of derivations exhibited by the open deduction formalism might be extended to  $\varepsilon$ -terms, which is a question that naturally arises when considering how to implement the epsilon-calculus in a deep-inference setting.

## Compositionality of Derivations Inside of Epsilon-Terms

Given the compositionality of derivations and context independence of inference rules exhibited by deep-inference proof systems, it is natural to ask whether this compositionality may be extended to  $\varepsilon$ -terms within the deep-inference methodology, so that epsilon symbols might be applied to derivations. I.e., given a sound derivation

$$\begin{array}{c} A \\ \phi \parallel \\ B \end{array}$$

is there some sense in which

$$\varepsilon_x \left( \begin{array}{c} A \\ \phi \parallel \\ B \end{array} \right)$$

is also sound, meaning that for any formula  $C(y)$  such that  $\varepsilon_x A$  and  $\varepsilon_x B$  are free for  $y$  in  $C(y)$ , if  $C(\varepsilon_x A)$  is valid then  $C(\varepsilon_x B)$  is also valid? In general, this is not the case for the standard semantics of the epsilon-calculus. Consider the following construction, which contains a sound derivation inside the context of an epsilon symbol  $\varepsilon_x$ :

$$A \left( \varepsilon_x \left( \begin{array}{c} \boxed{A(x)} \\ = \\ \boxed{A(x) \vee \boxed{\frac{f}{B(x)}}} \end{array} \right) \right)$$

For a given structure  $\mathcal{A}$ , assignment  $s$  on  $\mathcal{A}$  and extensional choice function  $\Phi$  on  $\mathcal{A}$ , we have that the premise  $A(\varepsilon_x A(x))$  of the construction satisfies  $\mathcal{A}, s, \Phi \models A(\varepsilon_x A(x))$  if and only if there exists some  $a \in |\mathcal{A}|$  such that  $\mathcal{A}, s[a/x], \Phi \models A(x)$ . The conclusion  $A(\varepsilon_x(A(x) \vee B(x)))$  of the construction, however, satisfies  $\mathcal{A}, s, \Phi \models A(\varepsilon_x(A(x) \vee B(x)))$  if and only if the element  $e = \text{val}_{\mathcal{A}, \Phi, s}(\varepsilon_x(A(x) \vee B(x)))$  assigned to  $\varepsilon_x(A(x) \vee B(x))$  by  $\Phi$  satisfies  $\mathcal{A}, s[e/x], \Phi \models A(x)$ . Since it may be the case that  $\mathcal{A}, s[e/x], \Phi \models \neg A(x) \wedge B(x)$  so that  $\mathcal{A}, s, \Phi \not\models A(\varepsilon_x(A(x) \vee B(x)))$ , the validity of the conclusion does not necessarily follow from the validity of the premise and hence the construction is unsound. This argument can be generalised to any valid semantics for the epsilon-calculus for which the condition

$$\llbracket \varepsilon_x A(x) \rrbracket_{\mathbb{D}} = \begin{cases} d & \text{if there exists some } d \in \mathbb{D} \text{ such that } \llbracket A(d) \rrbracket_{\mathbb{D}} \\ a & \text{for some arbitrary } a \in \mathbb{D}, \text{ otherwise} \end{cases}$$

holds for all  $\varepsilon$ -terms  $\varepsilon_x A(x)$ , demonstrating that the above construction is unsound for any standard semantics for the epsilon-calculus.

Likewise, reversing the premises and conclusions of inference rules and derivations inside the context of epsilon symbols also does not result in sound compositionality inside of  $\varepsilon$ -terms. Consider the following construction, which contains an occurrence of the reversed weakening rule  $\neg w\downarrow$  inside the context of an epsilon symbol  $\varepsilon_x$ :

$$A \left( \varepsilon_x \left( \begin{array}{c} \boxed{\neg w\downarrow \frac{B(x)}{f}} \end{array} \right) \right)$$

Once again, for a given structure  $\mathcal{A}$ , assignment  $s$  on  $\mathcal{A}$  and extensional choice function  $\Phi$  on  $\mathcal{A}$ , the premise  $A(\varepsilon_x B(x))$  of the construction satisfies  $\mathcal{A}, s, \Phi \models A(\varepsilon_x B(x))$  if and only if  $\mathcal{A}, s[e/x], \Phi \models A(x)$ , where  $e = \text{val}_{\mathcal{A}, s, \Phi}(\varepsilon_x B(x))$  is the element of  $|\mathcal{A}|$  assigned to  $\varepsilon_x B(x)$  by  $\Phi$ . The conclusion  $A(\varepsilon_x f)$  of the construction satisfies  $\mathcal{A}, s, \Phi \models A(\varepsilon_x f)$  if and only if  $\mathcal{A}, s[e'/x], \Phi \models A(x)$ , where  $e' = \text{val}_{\mathcal{A}, s, \Phi}(\varepsilon_x f)$  is the element of  $|\mathcal{A}|$  assigned to  $\varepsilon_x f$  by  $\Phi$ . Since  $e$  and  $e'$  are independent, we may have that  $\mathcal{A}, s[e/x], \Phi \models A(x)$  and

$\mathcal{A}, s[e'/x], \Phi \not\models A(x)$ . Hence the validity of the conclusion of the construction does not follow from the validity of its premise so that the construction is unsound.

The above examples demonstrate that the standard semantics of the epsilon-calculus are insufficient to obtain compositionality of derivations inside of  $\varepsilon$ -terms. However, it remains an open, and interesting, question as to whether imposing additional constraints on the semantics may be able to yield such compositionality. It is possible that deep inference may be able to inform the design of an alternate syntax for the epsilon-calculus with improved normalisation properties, similar to the recent development of subatomic logic [4, 5] as an alternate syntax for propositional logics with benefits in proof normalisation. Some alternate syntaxes for the epsilon-calculus have already been proposed in the literature [14, 36, 38], but usually with the motivation of simplifying the syntax of the standard epsilon-calculus rather than improving proof normalisation.

## 4.1 The Falsifier Rule

### Epsilon-Terms Arising From Quantifier-Shifts

As discussed in Subsection 3.3 and demonstrated by Proposition 3.16, most of the quantifier-shifts in  $\text{SKSg1}$  are derivable in  $\text{SKS1} \setminus \{\text{r1}\downarrow, \text{r1}\uparrow\}$ , with the exceptions of the rules  $\text{r1}\downarrow$  and  $\text{r1}\uparrow$  given by

$$\text{r1}\downarrow \frac{\forall x(A(x) \vee B)}{\forall xA(x) \vee B} \qquad \text{r1}\uparrow \frac{\exists xA(x) \wedge B}{\exists x(A(x) \wedge B)}$$

where  $x$  does not occur free in  $B$ . The more delicate nature of these rules suggests that a finer analysis is needed to understand their roles in the proof theory of first-order predicate logic and the non-elementary compression for cut-free proofs yielded by quantifier-shifts. As such, consider the following derivation, which contains an occurrence of the  $\text{r1}\downarrow$  quantifier-shift:

$$\text{r1}\downarrow \frac{\forall x \left( A(x) \vee \boxed{\begin{array}{c} \exists \frac{B(x)}{\exists y B(y)} \end{array}} \right)}{\forall xA(x) \vee \exists y B(y)}$$

Observe that the occurrence of  $\text{r1}\downarrow$  alters the witness to the existential quantifier. When the existential quantifier  $\exists y$  is instantiated, it is witnessed by the variable  $x$ , which is bound by a universal quantifier  $\forall x$ . After the application of  $\text{r1}\downarrow$  however, the existential quantifier  $\exists y$  no longer occurs in the scope of the universal quantifier  $\forall x$  and hence is not witnessed by  $x$  – there is no constructive witness to the existential quantifier  $\exists y$  given by the derivation.

However, by providing a semantic argument for the soundness of the derivation, we can provide a non-constructive witness for the existential quantifier  $\exists y$  in the conclusion, as follows:

We have that either (1)  $\forall xA(x)$  is true or (2)  $\exists x\bar{A}(x)$  is true. In case (1), we have that the conclusion of the derivation  $\forall xA(x) \vee \exists yB(y)$  is true. Otherwise, in case (2), there exists some element  $e$  of the domain such that  $\bar{A}(e)$  is true. From the premise of the derivation, we have that  $\forall x(A(x) \vee B(x))$  is true so that  $A(e) \vee B(e)$  is true and hence we must have that  $B(e)$  is true. It follows that  $\exists yB(y)$  is true and hence the conclusion of the derivation  $\forall xA(x) \vee \exists yB(y)$  is true.

This argument demonstrates that the conclusion of the derivation logically follows from its premise by assigning an arbitrary witness to the existential quantifier  $\exists y$  in case (1), that  $\forall xA(x)$  is true, and assigning it some element  $e$  of the domain such that  $\bar{A}(e)$  is true in case (2), that  $\exists x\bar{A}(x)$  is true. Observe that this definition for the witness assigned to the existential quantifier  $\exists y$  in the conclusion of the derivation coincides precisely with the element of the domain assigned to the  $\varepsilon$ -term  $\varepsilon_y\bar{A}(y)$ .

Indeed, to assign an explicit non-constructive witness to the existential quantifier  $\exists y$  in the conclusion of the derivation, I introduce the *falsifier rule*  $\varepsilon$ , given by

$$\varepsilon \frac{\forall x(A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))}$$

where  $y$  is a variable such that  $y$  is free for  $x$  in  $A(x)$  and  $\varepsilon_y \bar{A}(y)$  is free for  $x$  in  $B(x)$ . Observe that the falsifier rule  $\varepsilon$  is a generalisation of the quantifier-shift  $r1\downarrow$ . By replacing the occurrence of  $r1\downarrow$  in the derivation with an occurrence of the falsifier rule  $\varepsilon$ , we can permute the occurrence of the existential witnessing rule  $\exists$  down to obtain an explicit witness  $\varepsilon_y \bar{A}(y)$  for the existential quantifier  $\exists y$  in the conclusion, like so:

$$\begin{array}{c} \boxed{\forall x \left( A(x) \vee \boxed{\exists \frac{B(x)}{\exists y B(y)}} \right)} \\ r1\downarrow \\ \forall x A(x) \vee \exists y B(y) \end{array} \rightarrow \begin{array}{c} \boxed{\forall x A(x) \vee \boxed{\exists \frac{B(\varepsilon_y \bar{A}(y))}{\exists y B(y)}}} \\ \varepsilon \frac{\forall x(A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))} \end{array}$$

It is perhaps worth noting that the derivation above resembles the inference rule  $u\downarrow$  which sometimes appears in the deep-inference literature [20, 78], given by

$$u\downarrow \frac{\forall x(A \vee B)}{\forall x A \vee \exists x B}$$

so that the falsifier rule  $\varepsilon$  may be seen as arising from a decomposition of the inference rule  $u\downarrow$ .

The quantifier-shift  $r1\uparrow$ , on the other hand, may have all of its occurrences eliminated from a first-order proof by a particular normalisation. In the presence of the falsifier rule  $\varepsilon$ , by permuting occurrences of the existential contraction rule  $qc\downarrow$  and the existential witnessing rule  $\exists$  down a proof, an explicit disjunction of witnesses may be obtained for every existential quantifier in the proof, with  $\varepsilon$ -terms generated as witnesses when permuting occurrences of  $\exists$  down through occurrences of  $\varepsilon$ . Occurrences of the  $r1\uparrow$  quantifier-shift may then be eliminated from a proof by the following transformation:

$$\begin{array}{c} \boxed{\boxed{A(t_1) \vee \dots \vee A(t_n)} \\ \parallel \{\exists, qc\downarrow\} \\ \exists x A(x)} \wedge B \\ r1\uparrow \\ \exists x(A(x) \wedge B) \end{array} \rightarrow \begin{array}{c} \boxed{(A(t_1) \vee \dots \vee A(t_n)) \wedge \boxed{\begin{array}{c} B \\ \parallel \{\varepsilon\} \\ B \wedge \dots \wedge B \end{array}}} \\ \parallel \{\varepsilon\} \\ (A(t_1) \wedge B) \vee \dots \vee (A(t_n) \wedge B) \\ \parallel \{\exists, qc\downarrow\} \\ \exists x(A(x) \wedge B) \end{array}$$

Such permutations and transformations will be discussed in more detail in the following subsection, Subsection 4.2.

Furthermore, the falsifier rule  $\varepsilon$  may be used to overcome the issues associated with the quantifier-shift  $r1\downarrow$  which result in non-termination of the case analysis extraction procedure discussed in Subsection 3.4, attaining termination of this procedure. This will be the subject of the following subsection, Subsection 4.2, and yields the central result of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem.

## Falsifiers

I now formally introduce the falsifier rule  $\varepsilon$  as discussed above, as follows.

**Definition 4.1.** The falsifier rule  $\varepsilon$  is given by

$$\varepsilon \frac{\forall x(A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))}$$

for all formulae  $A(x)$ ,  $B(x)$  and all variables  $y$  such that  $y$  is free for  $x$  in  $A(x)$  and  $\varepsilon_y \bar{A}(y)$  is free for  $x$  in  $B(x)$ .

For any instance of the  $\varepsilon$  rule as above, the variable  $y$  is called the *variable name* of the instance of the  $\varepsilon$  rule.

The falsifier rule  $\varepsilon$  is sound by the following logical argument, which resembles the argument above for assigning an explicit witness to the existential quantifier in the conclusion of the derivation containing an occurrence of  $\text{r1}\downarrow$ :

Suppose that for some variable  $x$  and formulae  $A(x)$  and  $B(x)$ ,  $\forall x(A(x) \vee B(x))$  is true. We have that either (1)  $\forall x A(x)$  is true, or (2)  $\exists x \bar{A}(x)$  is true. In case (1), we have that  $\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))$  is true, as desired. Otherwise, in case (2), there must exist some element  $e$  in the domain such that  $\bar{A}(e)$  is true and hence  $\varepsilon_y \bar{A}(y)$  satisfies  $\bar{A}(\varepsilon_y \bar{A}(y))$ . Since  $\forall x(A(x) \vee B(x))$  is true, it follows that  $A(\varepsilon_y \bar{A}(y)) \vee B(\varepsilon_y \bar{A}(y))$  is true and, since  $\bar{A}(\varepsilon_y \bar{A}(y))$ , we therefore must have that  $B(\varepsilon_y \bar{A}(y))$  is true. It follows that  $\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))$  is true, as desired.

I formalise this argument using the semantics defined in Subsection 2.2, as follows.

**Proposition 4.2.** *The falsifier rule  $\varepsilon$  is sound, i.e., for all variables  $x$  and all formulae  $A(x)$  and  $B(x)$ , if  $\models \forall x(A(x) \vee B(x))$ , then for all variables  $y$  such that  $y$  is free for  $x$  in  $A(x)$  and  $\varepsilon_y \bar{A}(y)$  is free for  $x$  in  $B(x)$ ,  $\models \forall x A(x) \vee B(\varepsilon_y \bar{A}(y))$ .*

*Proof.* Let  $x$  be a variable,  $A(x)$  and  $B(x)$  be formulae,  $\mathcal{A}$  be a structure,  $\Phi$  be an extensional choice function on  $\mathcal{A}$  and  $s$  be an assignment on  $\mathcal{A}$  and suppose that  $\models \forall x(A(x) \vee B(x))$ . We have that  $\mathcal{A}, \Phi, s \models \forall x A(x) \vee B(\varepsilon_y \bar{A}(y))$  if and only if  $\mathcal{A}, \Phi, s \models \forall x A(x)$  or  $\mathcal{A}, \Phi, s \models B(\varepsilon_y \bar{A}(y))$  so that if  $\mathcal{A}, \Phi, s \models \forall x A(x)$ , then  $\mathcal{A}, \Phi, s \models \forall x A(x) \vee B(\varepsilon_y \bar{A}(y))$ . Otherwise, if  $\mathcal{A}, \Phi, s \not\models \forall x A(x)$ , then there must exist some  $a \in |\mathcal{A}|$  such that  $\mathcal{A}, \Phi, s[a/x] \not\models A(x)$  so that, by Lemma 2.19,  $\mathcal{A}, \Phi, s[a/x] \models \bar{A}(x)$  and hence  $\{b \in |\mathcal{A}| \mid \mathcal{A}, \Phi, s[b/y] \models \bar{A}(y)\}$  is non-empty. Thus  $\text{val}_{\mathcal{A}, \Phi, s}(\varepsilon_y \bar{A}(y)) = \Phi(\{b \in |\mathcal{A}| \mid \mathcal{A}, \Phi, s[b/y] \models \bar{A}(y)\})$  satisfies  $\mathcal{A}, \Phi, s[\text{val}_{\mathcal{A}, \Phi, s}(\varepsilon_y \bar{A}(y))/x] \models \bar{A}(x)$ . By assumption,  $\mathcal{A}, \Phi, s \models \forall x(A(x) \vee B(x))$  so that  $\mathcal{A}, \Phi, s[\text{val}_{\mathcal{A}, \Phi, s}(\varepsilon_y \bar{A}(y))/x] \models A(x) \vee B(x)$  and hence  $\mathcal{A}, \Phi, s[\text{val}_{\mathcal{A}, \Phi, s}(\varepsilon_y \bar{A}(y))/x] \models A(x)$  or  $\mathcal{A}, \Phi, s[\text{val}_{\mathcal{A}, \Phi, s}(\varepsilon_y \bar{A}(y))/x] \models B(x)$ . Since  $\mathcal{A}, \Phi, s[\text{val}_{\mathcal{A}, \Phi, s}(\varepsilon_y \bar{A}(y))/x] \models \bar{A}(x)$ , by Lemma 2.19, we must have  $\mathcal{A}, \Phi, s[\text{val}_{\mathcal{A}, \Phi, s}(\varepsilon_y \bar{A}(y))/x] \models B(x)$ . By Lemma 2.20, it follows that  $\mathcal{A}, \Phi, s \models B(\varepsilon_y \bar{A}(y))$  so that  $\mathcal{A}, \Phi, s \models \forall x A(x) \vee B(\varepsilon_y \bar{A}(y))$ . Therefore  $\models \forall x A(x) \vee B(\varepsilon_y \bar{A}(y))$ , as desired.  $\square$

### The Falsifier Calculus

I now introduce the *falsifier calculus*  $\text{SKSg}\varepsilon$  as the set of inference rules comprised of propositional rules  $\text{SKSg}_P$ , the universal instantiation rule  $\forall$ , the universal equality rule  $=_\forall$  and the falsifier rule  $\varepsilon$ . This definition is chosen due to the main result of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem, which decomposes a first-order proof into an upper segment in the falsifier calculus  $\text{SKSg}\varepsilon$  and a lower segment in  $\{\exists, \text{qc}\downarrow\}$ . The Falsifier Decomposition Theorem will be the main subject of the following subsection, Subsection 4.2. The inference rules of the falsifier calculus  $\text{SKSg}\varepsilon$  are displayed in Figure 3, after Corollary 4.15, which demonstrates that the open-deduction system with inference rules  $\text{SKSg}\varepsilon \cup \{\exists, \text{qc}\downarrow\}$  forms a complete proof system for first-order predicate logic.

**Definition 4.3.** The *falsifier calculus*  $\text{SKSg}\varepsilon$  is given by  $\text{SKSg}\varepsilon = \text{SKSg}_P \cup \{\varepsilon, \forall, =_\forall\}$ .

## Inference Rule Decompositions

As was discussed in Subsection 3.3 and demonstrated in Lemma 3.12, instances of the identity rule  $i\downarrow$  may be decomposed into derivations including the quantifier-shift  $r1\downarrow$ , which plays a central role in decomposing the part of the identity rule  $i\downarrow$  which acts on quantifiers. Since the falsifier rule  $\varepsilon$  generalises the  $r1\downarrow$  quantifier-shift, it is natural to consider how the falsifier rule  $\varepsilon$  might inform the decomposition of instances of the identity rule  $i\downarrow$ . An instance of the identity rule  $i\downarrow$  on a quantified formula  $\forall x A(x)$  may be decomposed into an instance of the falsifier rule  $\varepsilon$ , an instance of the existential witnessing rule  $\exists$  and an instance of the identity rule  $i\downarrow$  on the formula  $A(x)$ :

$$i\downarrow \frac{t}{\forall x A(x) \vee \exists x \bar{A}(x)} \rightarrow \varepsilon \frac{\begin{array}{c} = \frac{t}{\forall x \boxed{i\downarrow \frac{t}{A(x) \vee \bar{A}(x)}}} \\ \boxed{\forall x A(x) \vee \exists \boxed{\varepsilon_y \frac{\bar{A}(\varepsilon_y \bar{A}(y))}{\exists x \bar{A}(x)}}} \end{array}}{}$$

Observe that it follows that existential quantifiers introduced by instances of the identity rule  $i\downarrow$  may be understood to be witnessed by  $\varepsilon$ -terms.

The following proposition provides a decomposition of instances of the identity rule  $i\downarrow$  into derivations including the falsifier rule  $\varepsilon$ , which parallels the decomposition of instances of  $i\downarrow$  into derivations including the quantifier-shift  $r1\downarrow$  given in Lemma 3.12.

**Proposition 4.4.** *For every instance of the rule  $i\downarrow$  of the form*

$$i\downarrow \frac{t}{A \vee \bar{A}}$$

*there exists a derivation*

$$\frac{t}{\parallel \{\text{ai}\downarrow, s, \exists, \varepsilon, =_P\}} \frac{}{A \vee \bar{A}}$$

*which is of size  $O(|A|^2)$  and  $\varepsilon$ -size  $O(|A|^2|A|_\varepsilon^2)$ .*

*Proof.* I present derivations for decomposing  $i\downarrow$  rule instances by structural induction.

For all formulae  $A$ , let  $D(A)$  denote the derivation in  $\{\text{ai}\downarrow, s, \exists, \varepsilon, =_P\}$  obtained by decomposing the rule instance  $i\downarrow \frac{t}{A \vee \bar{A}}$ . For all formulae  $A$  and  $B$ , all atomic formulae  $a$  and all variables  $x$ , we have

$$D(A \vee B) = \frac{= \frac{t}{\boxed{\boxed{D(A) \parallel \frac{t}{A \vee \bar{A}}} \wedge \boxed{D(B) \parallel \frac{t}{B \vee \bar{B}}}}} \parallel_{\{s\}}}{(A \vee B) \vee (\bar{A} \wedge \bar{B})}$$

so that  $|D(A \vee B)| \leq k_1|A \vee B| + |D(A)| + |D(B)|$  for some constant  $k_1$  and the largest  $\varepsilon$ -term to occur in  $D(A \vee B)$  is the largest  $\varepsilon$ -term to occur in either  $D(A)$  or  $D(B)$ .



$$\begin{aligned}
D(A \wedge B) &= \overline{\overline{\begin{array}{c} \text{t} \\ D(A) \parallel \\ A \vee \bar{A} \end{array} \wedge \begin{array}{c} \text{t} \\ D(B) \parallel \\ B \vee \bar{B} \end{array}}} \\
&\quad \parallel_{\{s\}} \\
&\quad (A \wedge B) \vee (\bar{A} \vee \bar{B})
\end{aligned}$$

so that  $|D(A \wedge B)| \leq k_2|A \wedge B| + |D(A)| + |D(B)|$  for some constant  $k_2$  and the largest  $\varepsilon$ -term to occur in  $D(A \wedge B)$  is the largest  $\varepsilon$ -term to occur in either  $D(A)$  or  $D(B)$ .

$$\begin{aligned}
D(\forall x A(x)) &= \varepsilon \overline{\overline{\begin{array}{c} \forall x \begin{array}{c} \text{t} \\ D(A(x)) \parallel \\ A(x) \vee \bar{A}(x) \end{array} \\ \forall x A(x) \vee \begin{array}{c} \overline{A(\varepsilon_y \bar{A}(y))} \\ \exists x \bar{A}(x) \end{array} \end{array}}}
\end{aligned}$$

where  $y$  is some variable term that is free for  $x$  in  $A(x)$ , so that  $|D(\forall x A(x))| \leq 3|\forall x A(x)| + |D(A(x))| + 2$  and the largest  $\varepsilon$ -term to occur in  $D(\forall x A(x))$  is the largest  $\varepsilon$ -term to occur in either  $D(A(x))$  or  $\bar{A}(\varepsilon_y \bar{A}(y))$ . Note that the largest  $\varepsilon$ -term to occur in  $\bar{A}(\varepsilon_y \bar{A}(y))$  is of  $\varepsilon$ -size at most  $|\bar{A}(\varepsilon_y \bar{A}(y))|_\varepsilon \leq |A(x)|_\varepsilon^2$ .

$$\begin{aligned}
D(\exists x A(x)) &= \varepsilon \overline{\overline{\begin{array}{c} \forall x \begin{array}{c} \text{t} \\ D(A(x)) \parallel \\ A(x) \vee \bar{A}(x) \end{array} \\ \begin{array}{c} \overline{A(\varepsilon_y A(y))} \\ \exists x A(x) \end{array} \vee \forall x \bar{A}(x) \end{array}}}
\end{aligned}$$

where  $y$  is some variable term that is free for  $x$  in  $A(x)$ , so that  $|D(\exists x A(x))| \leq 3|\exists x A(x)| + |D(A(x))| + 2$  and the largest  $\varepsilon$ -term to occur in  $D(\exists x A(x))$  is the largest  $\varepsilon$ -term to occur in either  $D(A(x))$  or  $A(\varepsilon_y A(y))$ . Note that the largest  $\varepsilon$ -term to occur in  $A(\varepsilon_y A(y))$  is of  $\varepsilon$ -size at most  $|A(\varepsilon_y A(y))|_\varepsilon \leq |A(x)|_\varepsilon^2$ .

$$D(\text{t}) = \frac{\text{t}}{\text{t} \vee \text{f}}$$

so that  $|D(\text{t})| = 4$  and no  $\varepsilon$ -terms occur in  $D(\text{t})$ .

$$D(\text{f}) = \frac{\text{t}}{\text{f} \vee \text{t}}$$

so that  $|D(\text{f})| = 4$  and no  $\varepsilon$ -terms occur in  $D(\text{f})$ .

$$D(a) = \text{ai}\downarrow \frac{\text{t}}{a \vee \bar{a}}$$

so that  $|D(a)| = 2|a| + 2$  and the largest  $\varepsilon$ -term to occur in  $D(a)$  is the largest  $\varepsilon$ -term to occur in  $a$ .

It follows by appropriate choice of constant  $K$  that for all formulae  $A$ ,  $|D(A)| \leq K|A|^2$  and that the largest  $\varepsilon$ -term to occur in  $D(A)$  is of  $\varepsilon$ -size at most  $|A|_\varepsilon^2$ . By Lemma 2.33, it follows that  $|D(A)|_\varepsilon \leq |D(A)||A|_\varepsilon^2$ . The result follows.  $\square$

Observe that, unlike the sets of inference rules previously defined (see Proposition 3.3), the falsifier calculus  $\mathbf{SKSg}\varepsilon$  is not closed under dual inference rules, since the inference rule  $\bar{\varepsilon}$  which is dual to the falsifier rule  $\varepsilon$ , given by

$$\bar{\varepsilon} \frac{\exists x A(x) \wedge B(\varepsilon_y A(y))}{\exists x (A(x) \wedge B(x))}$$

where  $y$  is a variable such that  $y$  is free for  $x$  in  $A(x)$  and  $\varepsilon_y A(y)$  is free for  $x$  in  $B(x)$ , is not included in the falsifier calculus  $\mathbf{SKSg}\varepsilon$ . This is because the inference rule  $\bar{\varepsilon}$  is not directly relevant to Theorem 4.13 the Falsifier Decomposition Theorem and, as such, its proof theory will not be explored in detail in this thesis.

However, it follows from Proposition 4.4 above, by taking the duals of the derivations in that proposition, that the inference rule  $\bar{\varepsilon}$  may be used to provide a novel decomposition result for instances of the cut rule  $i\uparrow$ , as follows.

**Proposition 4.5.** *For every instance of the rule  $i\uparrow$  of the form*

$$i\uparrow \frac{A \wedge \bar{A}}{f}$$

*there exists a derivation*

$$\frac{A \wedge \bar{A}}{\parallel \{\mathbf{ai}\uparrow, \mathbf{s}, \mathbf{\forall}, \mathbf{\bar{\varepsilon}}, =_P\}}{f}$$

*which is of size  $O(|A|^2)$  and  $\varepsilon$ -size  $O(|A|^2|A|_\varepsilon^2)$ .*

In particular, an instance of the cut rule  $i\uparrow$  on a quantified formula  $\forall x A(x)$  may be decomposed into an instance of the dual of the falsifier rule  $\bar{\varepsilon}$ , an instance of the universal instantiation rule  $\forall$  and an instance of the cut rule  $i\uparrow$  on the formula  $A(x)$ :

$$i\uparrow \frac{\forall x A(x) \wedge \exists x \bar{A}(x)}{f} \rightarrow \bar{\varepsilon} \frac{\boxed{\forall \frac{\forall x A(x)}{A(\varepsilon_y \bar{A}(y))}} \wedge \exists x \bar{A}(x)}{\exists x \boxed{i\uparrow \frac{A(x) \wedge \bar{A}(x)}{f}}}$$

Observe that this resembles the decomposition of instances of the cut rule  $i\uparrow$  using the quantifier-shift  $r1\uparrow$ , of which the inference rule  $\bar{\varepsilon}$  is a generalisation, given by Lemma 3.12 and described in Subsection 3.3:

$$i\uparrow \frac{\forall x A(x) \wedge \exists x \bar{A}(x)}{f} \rightarrow r1\uparrow \frac{\boxed{= \frac{\forall x A(x)}{\forall y A(y)} \wedge \exists x \bar{A}(x)}}{\exists x \boxed{i\uparrow \frac{\boxed{\forall \frac{\forall y A(y)}{A(x)}} \wedge \bar{A}(x)}}{f}}{f}$$

However, in those decompositions, the instance of the universal instantiation rule  $\forall$  is instantiated by the variable corresponding to the existential quantifier in the instance of the cut rule  $i\uparrow$ . In the decomposition using the rule  $\bar{\varepsilon}$ , the instance of the universal instantiation rule  $\forall$  is instantiated by an  $\varepsilon$ -term, suggesting that this decomposition may be of a different nature to decompositions of the cut rule  $i\uparrow$  studied previously in the deep-inference literature, and may be able to provide a unique perspective on cut elimination. Such an investigation has fallen outside the scope of this thesis, but provides an interesting avenue for further research.

## Epsilon-Closed Derivations

I will now introduce the notion of *epsilon-closed* derivations in the falsifier calculus  $\text{SKSg}\varepsilon$ , which are derivations which meet certain conditions on how  $\varepsilon$ -terms are introduced into the derivation and admit a simple lemma bounding their  $\varepsilon$ -size in terms of their size (see Lemma 4.12 below). I first introduce the following definitions which will be necessary for providing the definition of epsilon-closed derivations.

The following definition may be used to distinguish  $\varepsilon$ -terms which occur in a formula and contain free variables that are not bound by quantifiers or epsilon symbols in that formula.

**Definition 4.6.** If  $x$  occurs free in a formula  $A(x)$  and  $t(y)$  is a term such that  $y$  occurs free in  $t(y)$  and  $t(y)$  is free for  $x$  in  $A(x)$ , then  $t(y)$  is said to *occur with  $y$  free in  $A(t(y))$* .

**Example 4.7.** Let  $A(y)$ ,  $B(y)$  and  $C(y)$  be formulae in which  $y$  occurs free. Then  $y$  occurs free in the  $\varepsilon$ -term  $\varepsilon_x C(y)$  and  $\varepsilon_x C(y)$  occurs in the formula  $D(y)$  given by  $A(y) \wedge \exists y B(\varepsilon_x C(y))$ , but does not occur with  $y$  free in  $D(y)$ , since all occurrences of  $y$  in  $\varepsilon_x C(y)$  are bound by an existential quantifier in  $D(y)$ . Conversely,  $\varepsilon_x C(y)$  occurs with  $y$  free in the formula  $D(\varepsilon_x C(y))$  given by  $A(\varepsilon_x C(y)) \wedge \exists y B(\varepsilon_x C(y))$  since some occurrences of  $y$  in  $A(\varepsilon_x C(y))$  occur free in  $D(\varepsilon_x C(y))$ .

The following definition introduces the notion of a term being *constructed by* an occurrence of the universal instantiation rule  $\forall$  or the falsifier rule  $\varepsilon$ .

**Definition 4.8.** For a given instance of the universal instantiation rule  $\forall$  of the form

$$\frac{\forall x A(x)}{A(t)}$$

for all terms  $s(x)$  which occur with  $x$  free in  $A(x)$ , the term  $s(t)$  is said to be *constructed by* the instance of  $\forall$ .

Likewise, for a given instance of the falsifier rule  $\varepsilon$  of the form

$$\frac{\forall x (A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))}$$

for all terms  $s(x)$  which occur with  $x$  free in  $B(x)$ , the term  $s(\varepsilon_y \bar{A}(y))$  is said to be *constructed by* the instance of  $\varepsilon$ .

I now introduce the notion of an *epsilon-closed* derivation, which is a derivation in which all  $\varepsilon$ -terms are constructed by an occurrence of the falsifier rule  $\varepsilon$  or the universal instantiation rule  $\forall$ , which is “closed” in the sense that there exists an ordering of occurrences of  $\varepsilon$  and  $\forall$  in the derivation such that every  $\varepsilon$ -term constructed by a given rule occurrence is built from  $\varepsilon$ -terms which are constructed by earlier rule occurrences. This property will be used to provide bounds for the  $\varepsilon$ -size of derivations in the falsifier calculus  $\text{SKSg}\varepsilon$  (see Lemma 4.12 below).

**Definition 4.9.** A derivation  $\phi$  is said to be *epsilon-closed* if every  $\varepsilon$ -term which occurs in  $\phi$  is constructed by some occurrence of the  $\varepsilon$  rule or the  $\forall$  rule in  $\phi$  and there exists an *epsilon closure ordering*  $\preceq_\varepsilon$  for  $\phi$ , which is a total order on the occurrences of  $\varepsilon$  and  $\forall$  in  $\phi$  such that for every occurrence  $\chi$  of  $\varepsilon$  or  $\forall$  in  $\phi$  of the form  $\rho \frac{A}{B}$ :

1. every  $\varepsilon$ -term which occurs in  $A$  is constructed by some occurrence  $\omega$  of  $\varepsilon$  or  $\forall$  in  $\phi$  such that  $\omega \prec_\varepsilon \chi$
2. if  $\rho$  is  $\forall$ , every  $\varepsilon$ -term which occurs in the term that instantiates  $\chi$  is constructed by some occurrence  $\omega$  of  $\varepsilon$  or  $\forall$  in  $\phi$  such that  $\omega \prec_\varepsilon \chi$

**Example 4.10.** The following is an epsilon-closed derivation in  $\text{SKSg}_\varepsilon$

$$\varepsilon \frac{\forall x(A(x) \vee B(x))}{\boxed{\forall \frac{\forall x A(x)}{A(\varepsilon_y \bar{A}(y))} \vee B(\varepsilon_y \bar{A}(y))}}$$

since an ordering for which the occurrence of  $\varepsilon$  in the derivation precedes the occurrence of  $\forall$  is an epsilon closure ordering.

**Example 4.11.** The following is a derivation in  $\text{SKSg}_\varepsilon$  which is not epsilon-closed

$$\boxed{\forall \frac{\forall x A(x)}{A(\varepsilon_y C)} \wedge \forall \frac{\forall x B(x)}{B(\varepsilon_y C)}}$$

since, although every  $\varepsilon$ -term which is constructed by an occurrence of  $\forall$  in the derivation is constructed by another occurrence of  $\forall$ , there is no way of ordering the two occurrences of  $\forall$  such that condition (1) of Definition 4.9 is met to form an epsilon closure ordering.

The following lemma will be used to establish bounds for the  $\varepsilon$ -size of epsilon-closed derivations in the falsifier calculus  $\text{SKSg}_\varepsilon$  in terms of their size. In particular, it provides an elementary bound for the  $\varepsilon$ -size of an epsilon-closed derivation in the falsifier calculus  $\text{SKSg}_\varepsilon$  in terms of its size.

**Lemma 4.12.** *Let  $\phi$  be an epsilon-closed derivation in  $\text{SKSg}_\varepsilon$ . Then*

$$|\phi|_\varepsilon = O(\exp(\exp |\phi| \ln |\phi|))$$

*Proof.* If  $\phi$  is epsilon-free, then  $|\phi|_\varepsilon = |\phi| = O(\exp(\exp |\phi| \ln |\phi|))$ . Otherwise, since  $\phi$  is epsilon-closed, there exists an epsilon closure ordering  $\prec_\varepsilon$  for  $\phi$ . For any occurrence  $\chi$  of  $\varepsilon$  or  $\forall$  in  $\phi$  of the form  $\rho \frac{A}{B}$ , I will show by induction on  $n$ , the total number of occurrences  $\omega$  of  $\varepsilon$  and  $\forall$  in  $\phi$  such that  $\omega \prec_\varepsilon \chi$ , that every  $\varepsilon$ -term  $\varepsilon_x C$  which is constructed by  $\chi$  satisfies  $|\varepsilon_x C|_\varepsilon \leq |\phi|^{2^{n+1}-1}$ .

Consider the case  $n = 0$ , when  $\phi$  does not contain any occurrences  $\omega$  of  $\varepsilon$  or  $\forall$  such that  $\omega \prec_\varepsilon \chi$ . If  $\rho$  is  $\forall$ , by conditions (1) and (2) of Definition 4.9,  $A$  and  $B$  must be epsilon-free so that the claim is vacuously true. Otherwise, if  $\chi$  is an occurrence of  $\varepsilon$  of the form

$$\varepsilon \frac{\forall x(C(x) \vee D(x))}{\forall x C(x) \vee D(\varepsilon_y \bar{C}(y))}$$

then, by condition (1) of Definition 4.9, since  $\phi$  does not contain any occurrences  $\omega$  of  $\varepsilon$  or  $\forall$  such that  $\omega \prec_\varepsilon \chi$ ,  $C(x)$  and  $D(x)$  must be epsilon-free. Thus the only  $\varepsilon$ -term which may be constructed by  $\chi$  is  $\varepsilon_y \bar{C}(y)$ , which satisfies  $|\varepsilon_y \bar{C}(y)|_\varepsilon = |C(x)|_\varepsilon = |C(x)| \leq |\phi|$ . Hence the claim is true for  $n = 0$ .

Now suppose that the claim holds true for all  $n \leq k$  for some  $k \in \mathbb{N}$  and consider the case  $n = k + 1$ , when  $\phi$  contains  $k + 1$  occurrences  $\omega$  of  $\varepsilon$  or  $\forall$  such that  $\omega \prec_\varepsilon \chi$ . By the induction hypothesis, every  $\varepsilon$ -term  $\varepsilon_x C$  which is constructed by an occurrence  $\omega$  of  $\varepsilon$  or  $\forall$  in  $\phi$  such that  $\omega \prec_\varepsilon \chi$  satisfies  $|\varepsilon_x C|_\varepsilon \leq |\phi|^{2^{k+1}-1}$ .

If  $\chi$  is an occurrence of  $\varepsilon$  in  $\phi$  of the form

$$\varepsilon \frac{\forall x(C(x) \vee D(x))}{\forall x C(x) \vee D(\varepsilon_y \bar{C}(y))}$$

then, by condition (1) of Definition 4.9, every  $\varepsilon$ -term which occurs in  $C(x)$  or  $D(x)$  is constructed by some occurrence  $\omega$  of  $\varepsilon$  or  $\forall$  in  $\phi$  such that  $\omega \prec_\varepsilon \chi$  and hence has  $\varepsilon$ -size

at most  $|\phi|^{2^{k+1}-1}$ . Thus, by Lemma 2.33,  $|\varepsilon_y \overline{C}(y)|_\varepsilon \leq |C(x)| |\phi|^{2^{k+1}-1} \leq |\phi|^{2^{k+1}}$ . Since every  $\varepsilon$ -term which occurs in  $D(x)$  has  $\varepsilon$ -size at most  $|\phi|^{2^{k+1}-1}$ , the largest possible  $\varepsilon$ -term constructed by  $\chi$  has  $\varepsilon$ -size at most  $|\varepsilon_y \overline{C}(y)|_\varepsilon |\phi|^{2^{k+1}-1} \leq (|\phi|^{2^{k+1}})(|\phi|^{2^{k+1}-1}) = |\phi|^{2^{k+2}-1}$ .

Otherwise, if  $\chi$  is an occurrence of  $\forall$  in  $\phi$  of the form

$$\forall \frac{\forall x C(x)}{C(t)}$$

then, by conditions (1) and (2) of Definition 4.9, every  $\varepsilon$ -term which occurs in  $C(x)$  or  $t$  is constructed by some occurrence  $\omega$  of  $\varepsilon$  or  $\forall$  in  $\phi$  such that  $\omega \prec_\varepsilon \chi$  and hence has  $\varepsilon$ -size at most  $|\phi|^{2^{k+1}-1}$ . Thus, by Lemma 2.33,  $|t|_\varepsilon \leq |t| |\phi|^{2^{k+1}-1} \leq |\phi|^{2^{k+1}}$ . Since every  $\varepsilon$ -term which occurs in  $C(x)$  has  $\varepsilon$ -size at most  $|\phi|^{2^{k+1}-1}$ , the largest possible  $\varepsilon$ -term constructed by  $\chi$  has  $\varepsilon$ -size at most  $|t|_\varepsilon |\phi|^{2^{k+1}-1} \leq (|\phi|^{2^{k+1}})(|\phi|^{2^{k+1}-1}) = |\phi|^{2^{k+2}-1}$ .

The claim is therefore true for  $n = k + 1$  and, by induction, holds for all  $n \in \mathbb{N}$ . Since  $\phi$  contains at most  $|\phi|$  total occurrences of  $\varepsilon$  and  $\forall$ , it follows that the largest  $\varepsilon$ -term  $\varepsilon_x A$  which may occur in  $\phi$  satisfies  $|\varepsilon_x A|_\varepsilon \leq |\phi|^{2^{|\phi|}-1}$ . Thus, by Lemma 2.33,  $|\phi|_\varepsilon \leq |\phi| |\varepsilon_x A|_\varepsilon \leq |\phi|^{2^{|\phi|}} = O(\exp(\exp |\phi| \ln |\phi|))$ , as required.  $\square$

## 4.2 The Falsifier Decomposition Theorem

I will now explore how the falsifier rule  $\varepsilon$  and the falsifier calculus  $\mathbf{SKSg}\varepsilon$  can be used to attain termination of the procedure described in Subsection 3.4 for extracting the case analyses contained in quantifier contraction rules from proofs, which is non-terminating in the standard epsilon-free syntax of predicate logic. This procedure permutes occurrences of the existential contraction rule  $\mathbf{qc}\downarrow$  down a proof and permutes occurrences of the universal cocontraction rule  $\mathbf{qc}\uparrow$  up a proof. The procedure is non-terminating in the standard epsilon-free syntax of predicate logic since each phase of permuting occurrences of either  $\mathbf{qc}\downarrow$  or  $\mathbf{qc}\uparrow$  introduces occurrences of the dual rule into the proof. The decomposition theorem resulting from the termination of this procedure in the falsifier calculus  $\mathbf{SKSg}\varepsilon$  is the central theorem of this thesis, Theorem 4.13 the Falsifier Decomposition Theorem, which will be stated below.

Recall transformation (2) from Subsection 3.4, which describes permuting an occurrence of  $\mathbf{qc}\uparrow$  up through an occurrence of  $\mathbf{r1}\downarrow$  and results in the non-termination of the procedure due to the introduction of an occurrence of  $\mathbf{qc}\downarrow$  into the proof:

$$\begin{array}{c} \mathbf{r1}\downarrow \frac{\forall x(A \vee \exists yB)}{\boxed{\mathbf{qc}\uparrow \frac{\forall xA}{\forall xA \wedge \forall xA} \vee \exists yB}} \end{array} \rightarrow \begin{array}{c} \mathbf{qc}\uparrow \frac{\forall x(A \vee \exists yB)}{\boxed{\mathbf{r1}\downarrow \frac{\forall x(A \vee \exists yB)}{\forall xA \vee \exists yB} \wedge \mathbf{r1}\downarrow \frac{\forall x(A \vee \exists yB)}{\forall xA \vee \exists yB}}} \\ \mathbf{s} \frac{\boxed{\mathbf{r1}\downarrow \frac{\forall x(A \vee \exists yB)}{\forall xA \vee \exists yB} \wedge \mathbf{r1}\downarrow \frac{\forall x(A \vee \exists yB)}{\forall xA \vee \exists yB}}{\boxed{\mathbf{s} \frac{(\forall xA \vee \exists yB) \wedge \forall xA}{(\forall xA \wedge \forall xA) \vee \exists yB} \vee \exists yB}} \\ = \boxed{(\forall xA \wedge \forall xA) \vee \mathbf{qc}\downarrow \frac{\exists yB \vee \exists yB}{\exists yB}} \end{array}$$

Observe that the two occurrences of  $\mathbf{r1}\downarrow$  in the derivation resulting from the transformation share the same premise  $\forall x(A \vee \exists yB)$  and conclusion  $\forall xA \vee \exists yB$ , where the existential quantifiers  $\exists y$  in the premises of the occurrences of  $\mathbf{r1}\downarrow$  result from the occurrence of  $\mathbf{qc}\uparrow$  at the top of the derivation so that they share the same witness. Therefore the existential quantifiers  $\exists y$  in the conclusion  $\forall xA \vee \exists yB$  of the occurrences of  $\mathbf{r1}\downarrow$  also share the same witness. Consequently, the highlighted occurrence of the existential contraction rule  $\mathbf{qc}\downarrow$  on

the formula  $\exists yB$  which is introduced into the proof represents a superfluous case analysis – it is a case analysis on two identical witnesses. This is due to the standard syntax of the predicate calculus being unable to express that the two witnesses are equal. However, by replacing the occurrences of  $r1\downarrow$  in the derivation with occurrences of the falsifier rule  $\varepsilon$ , the more expressive syntax of the epsilon-calculus is able to express that the two witnesses are identical and hence an occurrence of the regular contraction rule  $c\downarrow$  on a non-existentially quantified formula may be introduced in place of the occurrence of the existential contraction rule  $qc\downarrow$ , yielding termination of the procedure. In the falsifier calculus, by permuting occurrences of existential rules down a proof, we can obtain an explicit disjunction of witnesses for each existential quantifier in the proof, with  $\varepsilon$ -terms generated as witnesses when permuting down through an occurrence of the falsifier rule  $\varepsilon$ . The reduction corresponding to (2) for permuting an occurrence of  $qc\uparrow$  up through an occurrence of  $r1\downarrow$  in the falsifier calculus then takes the following form:

$$\begin{array}{c}
\boxed{\forall x \left( A \vee \boxed{\boxed{B' \parallel \{\exists, qc\downarrow\}} \exists y B(y)} \right)} \\
\downarrow r1\downarrow \\
\boxed{\boxed{qc\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}} \vee \exists y B(y)} \\
\downarrow \\
\forall x (A \vee B') \\
qc\uparrow \frac{}{\boxed{\boxed{\boxed{\frac{\varepsilon \frac{\forall x (A \vee B')}{\forall x A \vee B'[\varepsilon_y \bar{A}(y)/x]} \wedge \frac{\varepsilon \frac{\forall x (A \vee B')}{\forall x A \vee B'[\varepsilon_y \bar{A}(y)/x]}}{s} \frac{\forall x A \wedge (\forall x A \vee B'[\varepsilon_y \bar{A}(y)/x])}{(\forall x A \wedge \forall x A) \vee B'[\varepsilon_y \bar{A}(y)/x]} \vee B'[\varepsilon_y \bar{A}(y)/x]} \\
= \\
\boxed{(\forall x A \wedge \forall x A) \vee \boxed{\boxed{\boxed{B'[\varepsilon_y \bar{A}(y)/x] \vee B'[\varepsilon_y \bar{A}(y)/x]} \downarrow c\downarrow B'[\varepsilon_y \bar{A}(y)/x]} \parallel \{\exists, qc\downarrow\} \exists y B(y)}}}
\end{array} \tag{3}$$

where  $B'$  is the formula  $B(t_1) \vee \dots \vee B(t_n)$  for some terms  $t_1, \dots, t_n$ . Observe that, since the syntax of the epsilon-calculus is able to express that the two formulae  $B'[\varepsilon_y \bar{A}(y)/x]$  in the premise of the highlighted occurrence of the contraction rule  $c\downarrow$  are identical, the contraction acts on a non-existentially quantified formula so that no occurrence of  $qc\downarrow$  is introduced and there is no superfluous case analysis present.

In this manner, we can extract all case analyses contained within occurrences of the existential contraction rule  $qc\downarrow$  and the universal cocontraction rule  $qc\uparrow$  from a first-order proof by translating it into the falsifier calculus and performing the appropriate permutations. We first permute all occurrences of  $qc\downarrow$  down to the bottom of the proof, potentially introducing further occurrences of  $qc\uparrow$  into the proof in doing so. We then permute all occurrences of the existential witnessing rule  $\exists$  down to the bottom of the proof, assigning an explicit witness to each existential quantifier in the proof in doing so, and introducing  $\varepsilon$ -terms into the proof when permuting occurrences of  $\exists$  down through occurrences of the falsifier rule  $\varepsilon$ . Finally, we permute all occurrences of  $qc\uparrow$  up the proof until they are elim-

inated. This does not introduce any further occurrences of  $\mathbf{qc}\downarrow$  into the proof, as is the case for the permutations using the standard epsilon-free syntax of the predicate calculus described in Subsection 3.4, since all existential quantifiers in the proof have been assigned explicit witnesses which possibly include  $\varepsilon$ -terms. The procedure therefore terminates, as desired. Furthermore, this procedure only increases the size of the proof elementarily, since each of the three phases of rule permutations are of elementary complexity.

The main result of this thesis, Theorem 4.13 the *Falsifier Decomposition Theorem*, is the resultant decomposition theorem arising from this procedure. The proofs resulting from this procedure are in a normal form which I call *falsifier normal form*, which is a factorisation of the proof into an upper segment in the falsifier calculus  $\mathbf{SKSg}\varepsilon$  and a lower segment of existential witnesses and existential contractions in  $\{\exists, \mathbf{qc}\downarrow\}$ . I call the weakly existential-free formula which connects the two segments a *falsifier disjunction* for the conclusion of the proof, which may be seen as an analogue of a Herbrand disjunction in the language of the epsilon-calculus.

The Falsifier Decomposition Theorem for decomposing epsilon-free proofs in  $\mathbf{SKSg1}$  is stated as follows.

**Theorem 4.13** (The Falsifier Decomposition Theorem). *For every epsilon-free proof  $\phi$  with conclusion  $A$  in  $\mathbf{SKSg1}$ , there exists a proof of the form*

$$\begin{array}{c} \phi' \\ \parallel_{\mathbf{SKSg}\varepsilon} \\ A' \\ \parallel_{\{\exists\}} \\ A'' \\ \parallel_{\{\mathbf{qc}\downarrow\}} \\ A \end{array} \quad (4)$$

such that the following elementary bounds hold

$$\begin{aligned} |\phi'| &= \exp^8(O(|\phi|^2 \ln |\phi|)) \\ |A'| &= \exp^5(O(|\phi|^2 \ln |\phi|)) \\ |\phi'|_\varepsilon &= \exp^{10}(O(|\phi|^2 \ln |\phi|)) \\ |A'|_\varepsilon &= \exp^{10}(O(|\phi|^2 \ln |\phi|)) \end{aligned}$$

Furthermore, if  $\phi$  is cut-free then  $\phi'$  may be chosen to be cut-free.

It is expected that smaller bounds exist for the sizes and  $\varepsilon$ -sizes of  $\phi'$  and  $A'$  than those stated above, but the present bounds have been chosen for the sake of exposition of the complexity assessments and to demonstrate elementary complexity of proofs arising from the decomposition. I will provide a formal proof of the Falsifier Decomposition Theorem in the following section, Section 5.

The Falsifier Decomposition Theorem gives rise to the notion of *falsifier normal form* and *falsifier disjunctions*, as follows.

**Definition 4.14.** The normal form for proofs given by (4) is called *falsifier normal form* and the formula  $A'$  is called a *falsifier disjunction* for  $A$ .

The Falsifier Decomposition Theorem is analogous to Herbrand's Theorem in that the case analyses contained within occurrences of the existential contraction rule  $\mathbf{qc}\downarrow$  and universal cocontraction rule  $\mathbf{qc}\uparrow$  are extracted from the proof. However, unlike Herbrand's Theorem, the Falsifier Decomposition Theorem does not result in non-elementary blowups, since occurrences of the  $\mathbf{r1}\downarrow$  quantifier-shift in the proof are left intact in the form of occurrences of the falsifier rule  $\varepsilon$  which introduce  $\varepsilon$ -terms into the proof. The non-elementary difference in complexity between the two theorems may be seen as resulting from their difference in constructivity, since  $\varepsilon$ -terms represent elements which are drawn from the domain

non-constructively. Further comparison between the Falsifier Decomposition Theorem and Herbrand's Theorem is given in Subsection 4.4.

Examples of proofs in falsifier normal form are presented and discussed in Subsection 4.3 and a comparison of the Falsifier Decomposition Theorem and the falsifier calculus  $\mathbf{SKSg}_\varepsilon$  with existing proof systems and proof interpretations for first-order predicate logic may be found in Subsection 4.4. I will now note some of the consequences of Theorem 4.13 the Falsifier Decomposition Theorem.

It follows from the Falsifier Decomposition Theorem that the open-deduction proof system with rules  $\mathbf{SKSg}_\varepsilon \cup \{\exists, \text{qc}\downarrow\}$  forms a cut-free complete proof system for epsilon-free first-order predicate logic and admits non-elementarily smaller cut-free proofs than the sequent calculus  $\mathbf{LK}$  for certain classes of theorems, as follows.

**Corollary 4.15.** *Every valid epsilon-free formula has a proof in  $(\mathbf{SKSg}_\varepsilon \setminus \{\text{i}\uparrow\}) \cup \{\exists, \text{qc}\downarrow\}$  and there is no elementary function bounding the size of the smallest cut-free  $\mathbf{LK}$  proof of a formula in terms of the size or  $\varepsilon$ -size of its smallest proof in  $(\mathbf{SKSg}_\varepsilon \setminus \{\text{i}\uparrow\}) \cup \{\exists, \text{qc}\downarrow\}$ .*

*Proof.* By Theorem 3.3 of [2], there is no elementary function bounding the size of the smallest cut-free  $\mathbf{LK}$  proof of a formula in terms of the size of its smallest cut-free  $\mathbf{LK}_{\text{shift}}$  proof, where  $\mathbf{LK}_{\text{shift}}$  is the system presented in [2]. By Lemma 3.8,  $\mathbf{SKSg1} \setminus \{\text{i}\uparrow\}$  polynomially simulates cut-free  $\mathbf{LK}_{\text{shift}}$  so that there is no elementary function bounding the size of the smallest cut-free  $\mathbf{LK}$  proof of a formula in terms of the size of its smallest  $\mathbf{SKSg1} \setminus \{\text{i}\uparrow\}$  proof. The result follows by Theorem 4.13.  $\square$

The inference rules of the proof system  $\mathbf{SKSg}_\varepsilon \cup \{\exists, \text{qc}\downarrow\}$  are displayed in Figure 3.

It also follows from the Falsifier Decomposition Theorem that there exist classes of first-order theorems which admit non-elementarily smaller falsifier disjunctions than Herbrand disjunctions, as follows.

**Corollary 4.16.** *There is no elementary function bounding the size of the smallest Herbrand disjunction of a valid epsilon-free formula in terms of the size or  $\varepsilon$ -size of its smallest falsifier disjunction.*

*Proof.* By Theorem 3.3 of [2], there is no elementary function bounding the size of the smallest cut-free  $\mathbf{LK}$  proof of a formula in terms of the size of its smallest cut-free  $\mathbf{LK}_{\text{shift}}$  proof, where  $\mathbf{LK}_{\text{shift}}$  is the system presented in [2]. For any Herbrand disjunction  $A'$  for a valid formula  $A$ , the smallest cut-free  $\mathbf{LK}$  proof of  $A$  is of size at most  $O(\exp |A'|)$  since exponentially-sized cut-free  $\mathbf{LK}$  proofs exist for all propositional tautologies. It follows that there is no elementary function bounding the size of the smallest Herbrand disjunction for a formula in terms of the size of its smallest cut-free  $\mathbf{LK}_{\text{shift}}$  proof. By Lemma 3.8,  $\mathbf{SKSg1} \setminus \{\text{i}\uparrow, \forall\}$  polynomially simulates cut-free  $\mathbf{LK}_{\text{shift}}$  so that there is no elementary function bounding the size of the smallest Herbrand disjunction for a formula in terms of the size of its smallest  $\mathbf{SKSg1} \setminus \{\text{i}\uparrow, \forall\}$  proof. The result follows by Theorem 4.13.  $\square$

### 4.3 Examples

I will now provide some examples of proofs in falsifier normal form to provide insight into Theorem 4.13 the Falsifier Decomposition Theorem and compare them with proofs in other systems.

The drinker's paradox provides a simple yet elucidating example of the Falsifier Decomposition Theorem and its complexity.



The inference rules of  $\text{SKSg}\varepsilon$ :

$\text{i}\downarrow \frac{t}{A \vee \bar{A}}$	$\text{w}\downarrow \frac{f}{A}$	$\text{c}\downarrow \frac{A \vee A}{A}$	$\text{s} \frac{A \wedge (B \vee C)}{(A \wedge B) \vee C}$	$\forall \frac{\forall x A(x)}{A(t)}$
$\text{i}\uparrow \frac{A \wedge \bar{A}}{f}$	$\text{w}\uparrow \frac{A}{t}$	$\text{c}\uparrow \frac{A}{A \wedge A}$	$\text{m} \frac{(A \wedge B) \vee (C \wedge D)}{(A \vee C) \wedge (B \vee D)}$	$\varepsilon \frac{\forall x(A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))}$

where the formula  $A$  is weakly quantifier-free in the inference rules  $\text{i}\downarrow, \text{i}\uparrow, \text{w}\downarrow, \text{w}\uparrow, \text{c}\downarrow$  and  $\text{c}\uparrow$ ,  $t$  is free for  $x$  in  $A(x)$  in the inference rule  $\forall$ ,  $y$  is free for  $x$  in  $A(x)$  in the inference rule  $\varepsilon$  and  $\varepsilon_y \bar{A}(y)$  is free for  $x$  in  $B(x)$  in the inference rule  $\varepsilon$ .

The equality rules of  $\text{SKSg}\varepsilon$ :

$A \vee f =_{\mathbf{P}} A$	$f \wedge f =_{\mathbf{P}} f$	$t \vee t =_{\mathbf{P}} t$	$A \wedge t =_{\mathbf{P}} A$
$A \vee B =_{\mathbf{P}} B \vee A$	$(A \vee B) \vee C =_{\mathbf{P}} A \vee (B \vee C)$		
$A \wedge B =_{\mathbf{P}} B \wedge A$	$(A \wedge B) \wedge C =_{\mathbf{P}} A \wedge (B \wedge C)$		
$A =_{\forall} \forall x A$	$\forall x \forall y A =_{\forall} \forall y \forall x A$	$\forall x A(x) =_{\forall} \forall y A(y)$	

where  $x$  does not occur free in  $A$  in  $A =_{\forall} \forall x A$  and  $y$  is free for  $x$  in  $A(x)$  in  $\forall x A(x) =_{\forall} \forall y A(y)$ .

The inference rules of  $\{\exists, \text{qc}\downarrow\}$ :

$\exists \frac{A(t)}{\exists x A(x)}$	$\text{qc}\downarrow \frac{\exists x A \vee \exists x A}{\exists x A}$
---------------------------------------	--

where  $t$  is free for  $x$  in  $A(x)$  in the inference rule  $\exists$ .

**Figure 3:** The inference rules of  $\text{SKSg}\varepsilon \cup \{\exists, \text{qc}\downarrow\}$

**Example 4.17.** The following is a proof of the drinker's paradox in falsifier normal form.

$$\begin{array}{c}
= \frac{\mathbf{t}}{\forall x \frac{\mathbf{i}\downarrow \frac{\mathbf{t}}{P(x) \vee \overline{P}(x)}}{\varepsilon \frac{\forall x P(x) \vee \overline{P}(\varepsilon_y \overline{P}(y))}{\exists y (\forall x P(x) \vee \overline{P}(y))}}}
\end{array}$$

In this example, the falsifier disjunction for the formula  $\exists y(\forall x P(x) \vee \overline{P}(y))$  is  $\forall x P(x) \vee \overline{P}(\varepsilon_y \overline{P}(y))$ . The smallest Herbrand disjunction for the formula is  $\forall x_1 \forall x_2 (P(x_1) \vee \overline{P}(c) \vee P(x_2) \vee \overline{P}(x_1))$ , reflective of the compression seen in falsifier disjunctions over Herbrand disjunctions.

Indeed, for comparison, the following is the smallest proof in Herbrand normal form for the formula in **SKSg1**.

$$\begin{array}{c}
\forall x_1 \forall x_2 \frac{= \frac{\mathbf{i}\downarrow \frac{\mathbf{t}}{P(x_1) \vee \overline{P}(x_1)}}{P(x_1) \vee \frac{\mathbf{w}\downarrow \frac{\mathbf{f}}{\overline{P}(c)}} \vee \frac{\mathbf{w}\downarrow \frac{\mathbf{f}}{P(x_2)} \vee \overline{P}(x_1)}}{\parallel \{\exists\} \\
\exists y_1 \forall x_1 \exists y_2 \forall x_2 (P(x_1) \vee \overline{P}(y_1) \vee P(x_2) \vee \overline{P}(y_2)) \\
\parallel \{\mathbf{r1}\downarrow, \mathbf{r3}\downarrow\} \\
\mathbf{qc}\downarrow \frac{\exists y_1 (\forall x_1 P(x_1) \vee \overline{P}(y_1)) \vee \exists y_2 (\forall x_2 P(x_2) \vee \overline{P}(y_2))}{\exists y (\forall x P(x) \vee \overline{P}(y))}
\end{array}$$

The following is another proof for the formula in **SKSg1** which uses quantifier-shifts.

$$\begin{array}{c}
\mathbf{i}\downarrow \frac{\mathbf{t}}{\forall x P(x) \vee \exists y \overline{P}(y)} \\
\mathbf{r2}\uparrow \frac{\forall x P(x) \vee \exists y \overline{P}(y)}{\exists y (\forall x P(x) \vee \overline{P}(y))}
\end{array}$$

Finally, I will provide a proof of the drinker's paradox in the traditional epsilon-calculus for comparison. In the traditional epsilon-calculus, quantifiers are encoded by  $\varepsilon$ -terms using the logical equivalences  $\exists x A(x) \equiv A(\varepsilon_x A(x))$  and  $\forall x A(x) \equiv A(\varepsilon_x \overline{A}(x))$  and  $\varepsilon$ -terms are introduced into a proof by critical axioms, which may be represented in deep inference by the inference rule **CA** given by

$$\mathbf{CA} \frac{A(t)}{A(\varepsilon_x A(x))}$$

where  $\varepsilon_x A(x)$  and  $t$  are free for  $x$  in  $A(x)$ . The proof is given by

$$\mathbf{CA} \frac{\mathbf{i}\downarrow \frac{\mathbf{t}}{P(\varepsilon_x \overline{P}(x)) \vee \overline{P}(\varepsilon_x \overline{P}(x))}}{P((\varepsilon_x \overline{P}(x))) \vee \overline{P}(\varepsilon_y (P((\varepsilon_x \overline{P}(x))) \vee \overline{P}(y)))}$$

The following is an example of a proof in falsifier normal form of a formula which contains a conjunction and may be considered as a generalisation of the drinker's paradox.

$$\begin{array}{c}
= \\
\boxed{\begin{array}{c}
\forall x \forall y \quad \epsilon \quad \boxed{\forall x' \quad \boxed{\frac{t}{P(x') \vee \overline{P}(x')}}} \quad \wedge \quad \epsilon \quad \boxed{\forall y' \quad \boxed{\frac{t}{Q(y') \vee \overline{Q}(y')}}} \\
\frac{\boxed{\frac{\forall x' P(x')}{P(x)}} \vee \overline{P}(\epsilon_z \overline{P}(z))}{\vee} \quad \frac{\boxed{\frac{\forall y' Q(y')}{Q(y)}} \vee \overline{Q}(\epsilon_w \overline{Q}(w))}{\vee}
\end{array}} \\
\exists \\
\frac{\exists v \forall x \forall y [(P(x) \vee \overline{P}(\epsilon_z \overline{P}(z))) \wedge (Q(y) \vee \overline{Q}(v))]}{\exists v \exists v \forall x \forall y [(P(x) \vee \overline{P}(u)) \wedge (Q(y) \vee \overline{Q}(v))]}
\end{array}$$

For comparison, the following is a proof for the formula in **SKSg1** which uses quantifier-shifts.

The following is an example of a proof in falsifier normal form for which the corresponding falsifier disjunction contains a nested  $\varepsilon$ -term.

$$\begin{array}{c} \text{=} \\ \boxed{\begin{array}{c} \text{t} \\ \boxed{\forall x \begin{array}{c} \forall y \begin{array}{c} \forall x' \begin{array}{c} \forall y' \begin{array}{c} \text{i}\downarrow \frac{\text{t}}{P(x', y') \vee \overline{P}(x', y')} \\ \varepsilon \frac{}{\forall y' P(x', y') \vee \overline{P}(x', \varepsilon_z \overline{P}(x', z))} \end{array} \\ \varepsilon \frac{}{\forall y' P(\varepsilon_w P(w, \varepsilon_z \overline{P}(w, z)), y') \vee \forall x' \overline{P}(x', \varepsilon_z \overline{P}(x', z))} \\ \exists \frac{}{\exists v \forall y (P(\varepsilon_w P(w, \varepsilon_z \overline{P}(w, z)), y) \vee \overline{P}(x, v))} \end{array} \end{array} \end{array} \end{array} \\ \boxed{\exists u \forall x \exists v \forall y (P(u, y) \vee \overline{P}(x, v))} \end{array}$$

For comparison, the following is a proof for the formula in **SKSg1** which uses quantifier-shifts.

66

## 4.4 Comparison With Existing Proof Systems and Interpretations

I will now explore some of the connections between the falsifier calculus  $\text{SKSg}\varepsilon$  and Theorem 4.13 the Falsifier Decomposition Theorem with existing proof systems and proof interpretations for first-order predicate logic.

### Herbrand's Theorem

As discussed in Subsections 1.3 and 3.4, the general version of Herbrand's Theorem may be stated in deep inference [21] in the form of a decomposition theorem for proofs in first-order predicate logic, that every valid epsilon-free formula  $A$  has a proof of the form

$$\forall x_1 \dots \forall x_n \left[ \begin{array}{c} \prod \text{Propositional rules (SKSg}_P\text{)} \\ A' \end{array} \right] \begin{array}{c} \parallel \{\exists\} \\ A'' \\ \parallel \{r1\downarrow, r2\downarrow, r3\downarrow, r4\downarrow\} \\ A''' \\ \parallel \{\text{qc}\downarrow\} \\ A \end{array}$$

where  $A''$  is in prenex normal form and the formula  $\forall x_1 \dots \forall x_n A'$  is a Herbrand disjunction for  $A$ . This is in contrast to many statements of Herbrand's Theorem to be found in the literature, which are often restricted only to certain classes of formulae (such as existential formulae or  $\forall\exists$ -formulae) and do not admit direct statements in the formalism of the sequent calculus, instead being stated as metatheorems.

The resemblance of Herbrand's Theorem to Theorem 4.13 the Falsifier Decomposition Theorem is immediately apparent: both theorems provide a factorisation of proofs in first-order predicate logic into a form in which all occurrences of the existential contraction rule  $\text{qc}\downarrow$  are contained in the lowermost segment, representing the extraction of all case analyses contained within such rules from the proof. However, there is no elementary bound on the size of the smallest proof in Herbrand normal form for a formula in terms of the size of its smallest proof in  $\text{SKSg}1$  or Gentzen's sequent calculus **LK**, whereas Theorem 4.13 the Falsifier Decomposition Theorem provides elementary bounds for the size and  $\varepsilon$ -size of the smallest proof of a formula in falsifier normal form in terms of the size of its smallest proof in  $\text{SKSg}1$ . This is a consequence of the occurrences of the falsifier rule  $\varepsilon$  contained in the upper segment of a proof in falsifier normal form, which are introduced into the proof during the proof of the Falsifier Decomposition Theorem by substituting them for equivalent occurrences of the quantifier-shift  $r1\downarrow$  and introduce  $\varepsilon$ -terms into the proof. It follows that in a proof in falsifier normal form, the propositional content and the content induced by quantifier-shifts both occur in the upper segment of the proof in  $\text{SKSg}\varepsilon$ . In a proof in Herbrand normal form, on the other hand, all occurrences of quantifier-shift rules are contained in one of the lower segments of the proof so that the propositional segment in  $\text{SKSg}_P$  and quantifier-shift segment in  $\{r1\downarrow, r2\downarrow, r3\downarrow, r4\downarrow\}$  are separate.

The Falsifier Decomposition Theorem may be understood as a non-constructive analogue to Herbrand's Theorem since the  $\varepsilon$ -terms contained in a proof in falsifier normal form represent elements which are drawn from the domain non-constructively by the semantics of the epsilon-calculus, whereas the ground terms in a proof in Herbrand normal form represent the information in the proof in a purely constructive manner. The non-elementary difference in complexity between the two theorems may be seen as arising from this difference in constructivity. The elementary bounds on the size and  $\varepsilon$ -size of the smallest proof of a formula in falsifier normal form in terms of the size of its smallest proof in  $\text{SKSg}1$  also suggests that, unlike Herbrand's Theorem, the Falsifier Decomposition Theorem does not correspond to the elimination of all quantifier cuts from a first-order proof. Cut elimination for first-order

predicate logic is of non-elementary complexity [31], suggesting that the elimination of  $\varepsilon$ -terms from a proof in falsifier normal form is intimately related to cut elimination, as is the case in the traditional epsilon-calculus.

Furthermore, Herbrand's Theorem may be understood as a kind of reduction of proofs in first-order predicate logic to proofs in propositional classical logic. Since propositional classical logic is decidable and first-order predicate logic is undecidable, Herbrand's Theorem can be seen to isolate the undecidable content contained in a first-order proof in the lower segments of a proof in Herbrand normal form, in particular in the lowermost segment of existential contractions in  $\{\mathbf{qc}\downarrow\}$ . The decidability of provability in the falsifier calculus  $\mathbf{SKSg}\varepsilon$  remains an open question, which may yield a finer insight into the source of undecidability in first-order predicate logic.

Universal and existential quantifiers both play similar roles in Herbrand's Theorem and Theorem 4.13 the Falsifier Decomposition Theorem, in that universal quantifiers appear in the upper segments of the corresponding decomposition theorems and existential quantifiers appear in the lower segments, in an existential witnessing segment in  $\{\exists\}$  and an existential contraction segment in  $\{\mathbf{qc}\downarrow\}$ . This suggests that the asymmetric roles of universal and existential quantifiers in these theorems may be fundamental to their nature in the normalisation theory of first-order proofs. It follows from Herbrand's Theorem that every valid existential formula  $A$  has a proof of the form

$$\begin{array}{c} \prod \text{Propositional rules (SKSg}_P\text{)} \\ A' \\ \prod \{\exists\} \\ A'' \\ \prod \{\mathbf{qc}\downarrow\} \\ A \end{array}$$

which is both in Herbrand normal form and falsifier normal form so that the two theorems intersect in the universal-free setting. However, the inclusion of universal quantifiers and occurrences of the falsifier rule  $\varepsilon$  in the upper segment of such a proof of an existential formula can provide speedups (see, e.g., Example 6.7 below) so that the Falsifier Decomposition Theorem for existential formulae is not equivalent to Herbrand's Theorem.

Herbrand's Theorem is part of a broader family of theorems which provide independence and separation of the various elements of proofs in first-order predicate logic, such as Gentzen's sharpened *Hauptsatz* [41] and the epsilon theorems (see [75] and the discussion below). The deep-inference methodology has also provided new decomposition theorems for proofs in first-order predicate logic [20, 62], which have recently been applied in combinatorial proof theory [62]. Whereas the former class of theorems generally separate the propositional and first-order content of a proof, the latter class of decomposition theorems in deep inference stratify a proof in first-order predicate logic into independent segments comprised only of certain inference rules so that the separation is more syntactic in nature. As Theorem 4.13 the Falsifier Decomposition Theorem does not fully separate the propositional and first-order content of a proof, it provides an entirely new decomposition theorem and a novel insight into the normalisation theory of proofs in first-order predicate logic. Investigating methods of translating between proofs in falsifier normal form and other normal forms for first-order proofs provides an interesting avenue for further research. In particular, a procedure for extracting a proof in Herbrand normal form from a proof in falsifier normal form could potentially yield new insights into the structure of Herbrand disjunctions and into how falsifier disjunctions unfold into Herbrand disjunctions, providing a better understanding of how falsifier disjunctions serve as a non-elementarily more compact means of representing Herbrand disjunctions. It is expected that such a normalisation procedure may also provide a new perspective on the relationship between Herbrand's Theorem and cut elimination.

## The Traditional Epsilon-Calculus

The epsilon-calculus was initially introduced by Hilbert [57] in the early part of the twentieth century as part of his eponymous program attempting to prove the consistency of mathematics. The intention was to encode all quantifier information contained in a proof in  $\varepsilon$ -terms representing “ideal elements” using the logical equivalences  $\exists x A(x) \equiv A(\varepsilon_x A(x))$  and  $\forall x A(x) \equiv A(\varepsilon_x \bar{A}(x))$  and then to provide a procedure for eliminating all  $\varepsilon$ -terms from a proof, thereby reducing the proof to propositional form and demonstrating consistency of the system. Proof theory at this time was primarily focused on attaining completeness and consistency of proof systems, whereas modern research places more emphasis on complexity and normalisation, as reflected in the design of the falsifier calculus  $\text{SKSg}\varepsilon$ .

In the traditional epsilon-calculus,  $\varepsilon$ -terms are introduced into a proof using critical axioms, which may be represented in deep inference by the inference rule **CA** given by

$$\mathbf{CA} \frac{A(t)}{A(\varepsilon_x A(x))}$$

where  $\varepsilon_x A(x)$  and  $t$  are free for  $x$  in  $A(x)$ . Observe that **CA** is equivalent to the existential witnessing rule  $\exists$ , given by

$$\exists \frac{A(t)}{\exists x A(x)}$$

where  $t$  is free for  $x$  in  $A(x)$ , under the encoding of quantifiers by  $\varepsilon$ -terms since  $A(\varepsilon_x A(x)) \equiv \exists x A(x)$ . In the falsifier calculus  $\text{SKSg}\varepsilon$ , on the other hand,  $\varepsilon$ -terms are introduced into a proof using the falsifier rule  $\varepsilon$ , given by

$$\varepsilon \frac{\forall x(A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))}$$

where  $y$  is free for  $x$  in  $A(x)$  and  $\varepsilon_y \bar{A}(y)$  is free for  $x$  in  $B(x)$ . Both proof systems admit non-elementarily smaller cut-free proofs of certain theorems than the traditional sequent calculus [17], related to the non-elementary compression of cut-free proofs yielded by the extension of the sequent calculus by quantifier-shifts [2]. In the case of the falsifier calculus, this is because the open deduction formalism allows for the decomposition of instances of most of the quantifier-shift rules into instances of other rules (see Proposition 3.16) and because the falsifier rule  $\varepsilon$  generalises the quantifier-shift  $r1\downarrow$  and may be used to remove occurrences of the quantifier-shift  $r1\uparrow$  from a proof (see Lemma 5.5). In the traditional epsilon-calculus, the non-elementary compression of cut-free proofs results from the interaction between  $\varepsilon$ -terms and critical axioms, which simulate quantifier-shifts in the presence of the encodings of quantifiers by  $\varepsilon$ -terms. Consider, for instance, Example 4.17, in which a three-line proof of the  $\varepsilon$ -encoded form of the drinker’s paradox is given using a critical axiom. Furthermore,  $\varepsilon$ -terms play fundamentally different roles in the traditional epsilon-calculus and the falsifier calculus. In the traditional epsilon-calculus,  $\varepsilon$ -terms are used to encode quantifiers so that all quantifier information in a proof is represented by the  $\varepsilon$ -terms, which can result in unwieldy translated formulae in the epsilon-calculus (see Remark 2.2 for an example). In the falsifier calculus, however,  $\varepsilon$ -terms are present alongside quantifiers so that  $\varepsilon$ -terms are strictly used to represent elements which are drawn from the domain non-constructively. The falsifier calculus also does not use the cumbersome encodings of quantifiers by  $\varepsilon$ -terms, providing more legible proofs than the traditional epsilon-calculus. Furthermore, in an epsilon-closed proof in  $\text{SKSg}\varepsilon$ , such as those resulting from the procedure presented in Section 5 for transforming an epsilon-free proof in  $\text{SKSg}1$  into falsifier normal form, all  $\varepsilon$ -terms in the proof are also comprised only of  $\varepsilon$ -terms which are constructed by occurrences of the falsifier rule  $\varepsilon$  in the proof.

Central to the proof theory of the traditional epsilon-calculus are the first epsilon theorem and second epsilon theorem (see [75]), which respectively establish the conservativity of

extending propositional classical logic and first-order predicate logic by critical axioms. This is demonstrated by eliminating critical axioms from a proof through a procedure known as *epsilon substitution*. The extended first epsilon theorem [14, 75] further establishes that for any quantifier-free, epsilon-free formula  $A(x_1, \dots, x_n)$  and  $\varepsilon$ -terms  $\varepsilon_{x_1}B_1, \dots, \varepsilon_{x_n}B_n$  which are respectively free for  $x_1, \dots, x_n$  in  $A(x_1, \dots, x_n)$ , if  $A(\varepsilon_{x_1}B_1, \dots, \varepsilon_{x_n}B_n)$  is provable in the traditional epsilon-calculus, then there exist epsilon-free terms  $t_j^i$  for  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, m\}$  such that

$$\bigvee_{i=1}^{i=n} A(t_1^i, \dots, t_m^i)$$

is a propositional tautology. Herbrand's Theorem for existential formulae is an immediate corollary of this since existential quantifiers may be encoded by  $\varepsilon$ -terms using the logical equivalence  $\exists x A(x) \equiv A(\varepsilon_x A(x))$ . A natural question is whether an analogue of the extended first epsilon theorem holds for the falsifier calculus  $\text{SKSg}\varepsilon$ , which is related to the problem of providing a procedure for extracting proofs in Herbrand normal form from proofs in falsifier normal form. The traditional epsilon-calculus provides an efficient method for extracting Herbrand disjunctions from proofs [14], suggesting that such a procedure for the falsifier calculus  $\text{SKSg}\varepsilon$  may also have complexity benefits over traditional approaches to deriving Herbrand disjunctions, such as Buss-style proofs of Herbrand's Theorem. Theorem 4.13 the Falsifier Decomposition Theorem also resembles the epsilon theorems and may in a certain sense be understood as a variant of an epsilon theorem.

### The No-Counterexample Interpretation

In [65, 66], Kreisel introduced the *no-counterexample interpretation* of first-order predicate logic and Peano arithmetic. The no-counterexample interpretation uses terms constructed from case distinction functions to provide interpretations of proofs in these logics. The correctness of the no-counterexample interpretation is usually proved using epsilon substitution.

Following an overview due to Kohlenbach [64], the no-counterexample interpretation for existential formulae in first-order predicate logic may be stated as follows. For every valid epsilon-free formula of the form

$$\exists x_1 \dots \exists x_n A(x_1, \dots, x_n)$$

where  $A(x_1, \dots, x_n)$  is quantifier-free, by Herbrand's Theorem, there exist epsilon-free terms  $t_1^1, \dots, t_{m_1}^1, \dots, t_1^n, \dots, t_{m_n}^n$  such that

$$\bigvee_{i_1=1}^{i_1=m_1} \dots \bigvee_{i_n=1}^{i_n=m_n} A(t_{i_1}^1, \dots, t_{i_n}^n)$$

is valid. It follows that there exist case distinction functions

$$\Phi_1 = \begin{cases} t_1^1 & \text{if } \bigvee_{i_2=1}^{i_2=m_2} \dots \bigvee_{i_n=1}^{i_n=m_n} A(t_1^1, t_{i_2}^2, \dots, t_{i_n}^n) \\ t_2^1 & \text{if } \neg(\text{case 1}) \wedge \bigvee_{i_2=1}^{i_2=m_2} \dots \bigvee_{i_n=1}^{i_n=m_n} A(t_2^1, t_{i_2}^2, \dots, t_{i_n}^n) \\ \vdots & \\ \vdots & \end{cases}$$

$$\Phi_n = \begin{cases} t_1^n & \text{if } A(\Phi_1, \dots, \Phi_{n-1}, t_1^n) \\ t_2^n & \text{if } \neg(\text{case 1}) \wedge A(\Phi_1, \dots, \Phi_{n-1}, t_2^n) \\ \vdots & \end{cases}$$

such that

$$A(\Phi_1, \dots, \Phi_n)$$

is valid in an appropriate extension of first-order predicate logic. In such a case,  $\Phi_1, \dots, \Phi_n$  are said to *satisfy the no-counterexample interpretation of  $\exists x_1 \dots \exists x_n A(x_1, \dots, x_n)$* .

The  $\varepsilon$ -terms which occur in falsifier disjunctions closely resemble the case distinction functions used in the no-counterexample interpretation, suggesting that there may be some connection between Theorem 4.13 the Falsifier Decomposition Theorem and this interpretation. Developing a method of epsilon substitution for the falsifier calculus  $\text{SKSg}\varepsilon$  will likely yield insight into how the  $\varepsilon$ -terms which occur in a falsifier disjunction relate to the case distinction functions used in the no-counterexample interpretation, providing an insight into the connection between the two interpretations.

Another notable proof interpretation for first-order predicate logic which is defined using case distinction functions is Shoenfield's variant of Gödel's functional interpretation [42, 80] (see also the related recent work due to Afshari, Hetzl and Leigh [1]). Gödel's *functional interpretation* [50], sometimes called the *Dialectica interpretation* or *System T*, represents proofs as typed lambda terms with base type of the domain. In [80], Shoenfield introduced an extension of the functional interpretation by case distinction functions in order to provide a soundness proof for Peano arithmetic. This interpretation has been employed by Gerhardy and Kohlenbach to provide a means of extracting Herbrand disjunctions from sequent-calculus proofs for first-order predicate logic [42]. The case distinction functions in this interpretation arise from contraction rules in the corresponding sequent-calculus proof, but also resemble the  $\varepsilon$ -terms which occur in falsifier disjunctions due to the semantics of the epsilon-calculus. Once again, formalising the connection between Herbrand's Theorem and Theorem 4.13 the Falsifier Decomposition Theorem will likely reveal whether there is any connection between these case distinction functions and the  $\varepsilon$ -terms which occur in falsifier disjunctions.



## 5 Proof of The Falsifier Decomposition Theorem

The problems are solved, not by giving new information,  
but by arranging what we have known since long.

*L. Wittgenstein*

In this section, I will prove Theorem 4.13 the Falsifier Decomposition Theorem by formalising the three-phase procedure described in Subsection 4.2. The procedure will transform an epsilon-free proof in  $\text{SKSg1}$  into a proof in falsifier normal form as follows. To begin, the proof is first transformed into a proof in  $(\text{SKS1} \setminus \{\text{qc}\uparrow\}) \cup \{\text{c}\uparrow\}$  (see Lemma 5.1) so that the proof is in an appropriate form for performing the necessary rule permutations. Then, in the first phase of the procedure, occurrences of the existential contraction rule  $\text{qc}\downarrow$  are permuted down to the bottom of the proof (see Lemmas 5.2 and 5.3). In the second phase of the procedure, occurrences of the  $\text{r1}\downarrow$  quantifier-shift in the proof are replaced with equivalent occurrences of the falsifier rule  $\varepsilon$  so that occurrences of the existential witnessing rule  $\exists$  may then be permuted down the proof to separate the proof into an upper segment of weakly existential-free formulae and a lower segment in  $\{\exists, \text{qc}\downarrow\}$  (see Lemmas 5.4 and 5.5). In the third phase of the procedure, occurrences of the universal cocontraction rule  $\text{qc}\uparrow$  are permuted up the proof until they are eliminated (see Lemmas 5.6, 5.7 and 5.8), yielding a decomposed proof in falsifier normal form, as desired. The procedure is structured as follows:

$$\begin{array}{ccccc}
 \prod_A \text{SKSg1} & \xrightarrow[\text{Subsection 5.1}]{\text{Decompose}} & \prod_A (\text{SKS1} \setminus \{\text{qc}\uparrow\}) \cup \{\text{c}\uparrow\} & \xrightarrow[\text{Subsection 5.2}]{\text{Phase 1}} & \prod_{A'} \text{SKS1} \setminus \{\text{qc}\downarrow\} \\
 & & & & \prod_A \{\text{qc}\downarrow\} \xrightarrow[\text{Subsection 5.3}]{\text{Phase 2}} \\
 & & \prod_{A''} (\text{SKS1} \setminus \{\text{qc}\downarrow, \text{r1}\downarrow, \text{r1}\uparrow, \exists, =\exists\}) \cup \{\varepsilon\} & & \prod_{A'''} \text{SKSg}\varepsilon \\
 & & \prod_{A'} \{\exists\} & \xrightarrow[\text{Subsection 5.4}]{\text{Phase 3}} & \prod_{A'} \{\exists\} \\
 & & \prod_A \{\text{qc}\downarrow\} & & \prod_A \{\text{qc}\downarrow\}
 \end{array}$$

I separate the proof of Theorem 4.13 into three lemmas (Lemmas 5.3, 5.5 and 5.8), corresponding to each of the three phases of the procedure, along with further lemmas for each phase which provide local reductions for permuting occurrences of the appropriate inference rules up or down the proof. Each lemma also includes further information concerning the complexity and structure of the relevant transformations, which ensure that the complexity bounds of Theorem 4.13 are met. The transformations presented are not optimal, but are sufficient to obtain elementary bounds for the size and  $\varepsilon$ -size of the proof in falsifier normal form obtained by applying the procedure to an epsilon-free proof in  $\text{SKSg1}$ . Some optimisations for the procedure will be presented in Subsection 5.6, after the proof of Theorem 4.13 has been presented.

In Subsection 5.1 I provide a lemma for decomposing occurrences of certain inference rules in epsilon-free proofs in  $\text{SKSg1}$  into occurrences of inference rules in  $(\text{SKS1} \setminus \{\text{qc}\uparrow\}) \cup \{\text{c}\uparrow\}$ , which will be used to transform the proof into a form which makes the necessary rule permutations possible when transforming the proof into falsifier normal form. In Subsection 5.2 I provide lemmas for Phase 1 of the procedure, which permutes occurrences of the existential contraction rule  $\text{qc}\downarrow$  down to the bottom of the proof. In Subsection 5.3 I provide lemmas for Phase 2 of the procedure, which permutes occurrences of the existential witnessing rule  $\exists$  down to the bottom of the proof. In Subsection 5.4 I provide lemmas for Phase 3 of the procedure, which eliminates occurrences of the universal cocontraction rule  $\text{qc}\uparrow$  from the proof by permuting them up the proof. In Subsection 5.5 I provide the final

proof of Theorem 4.13 the Falsifier Decomposition Theorem, by combining the lemmas in the previous subsections. Finally, in Subsection 5.6 I describe some alterations which may be made to the procedure to optimise the complexity of the final proof in falsifier normal form and retain more of the structure of the initial proof prior to the procedure.

## 5.1 Rule Decomposition

The following lemma is a variant of Lemma 3.17 which provides a decomposition of derivations in  $\text{SKSg1}$  to derivations in  $(\text{SKS1} \setminus \{\text{qc}\uparrow\}) \cup \{\text{c}\uparrow\}$ . Prior to performing the necessary rule permutations for transforming an epsilon-free proof in  $\text{SKSg1}$  into falsifier normal form, it will first be transformed into a proof in  $(\text{SKS1} \setminus \{\text{qc}\uparrow\}) \cup \{\text{c}\uparrow\}$  using this lemma to make the proof amenable to rule permutations. This is necessary to isolate occurrences of the existential contraction rule  $\text{qc}\downarrow$  which may be contained in larger occurrences of the contraction rule  $\text{c}\downarrow$ . This lemma differs from Lemma 3.17 in that the decomposed proof may contain occurrences of  $\text{c}\uparrow$ . This is for convenience and efficiency during the first phase of the procedure (see Subsection 5.2 below), when occurrences of the existential contraction rule  $\text{qc}\downarrow$  are permuted down the proof, since decomposing occurrences of  $\text{c}\uparrow$  in the proof may introduce occurrences of vacuous  $=\exists$  rules into the proof which result in greater complexity when performing the permutations. As such, occurrences of  $\text{c}\uparrow$  in the proof are decomposed using Lemma 3.13 after the first phase of the procedure.

**Lemma 5.1.** *For every derivation  $\phi$  with premise  $A$  and conclusion  $B$  in  $\text{SKSg1}$ , there exists a proof  $\phi'$  with premise  $A$  and conclusion  $B$  in  $(\text{SKS1} \setminus \{\text{qc}\uparrow\}) \cup \{\text{c}\uparrow\}$  such that  $|\phi'| = O(|\phi|^2)$ . Furthermore, if  $\phi$  is epsilon-free then  $\phi'$  may be chosen to be epsilon-free and if  $\phi$  is cut-free then  $\phi'$  may be chosen to be cut-free.*

*Proof.* By replacing all occurrences of inference rules in  $\text{SKSg1} \setminus (\text{SKS1} \cup \{\text{c}\uparrow\})$  in  $\phi$  with derivations in  $\text{SKS1}$  using sequential composition, by Lemmas 3.12, 3.13, 3.14 and 3.15 and Proposition 3.16.  $\square$

## 5.2 Phase 1

In this subsection I will formalise the first phase of the proof of Theorem 4.13 the Falsifier Decomposition Theorem, in which occurrences of the existential contraction rule  $\text{qc}\downarrow$  are permuted down to the bottom of the proof.

The following lemma provides reduction rules for permuting occurrences of the existential contraction rule  $\text{qc}\downarrow$  down through occurrences of other inference rules in  $(\text{SKS1} \setminus \{\text{qc}\uparrow\}) \cup \{\text{c}\uparrow\}$ , which will be used during the first phase of the procedure (see Lemma 5.3 below).

**Lemma 5.2.** *For every inference rule  $\rho \in (\text{SKS1} \setminus \{\text{qc}\downarrow, \text{qc}\uparrow\}) \cup \{\text{c}\uparrow\}$  and every derivation  $\phi$  of the form*

$$\rho \frac{K \left\{ \frac{\text{qc}\downarrow \frac{\exists x A \vee \exists x A}{\exists x A}}{\exists x A} \right\}}{B}$$

where  $K\{\}$  is a formula context, there exists a derivation of the form

$$\begin{array}{c} K\{\exists x A \vee \exists x A\} \\ \phi' \parallel (\text{SKS1} \setminus \{\text{qc}\downarrow, \text{qc}\uparrow, \text{ai}\uparrow\})\{\text{c}\uparrow\} \\ B' \\ \parallel \{\text{qc}\downarrow\} \\ B \end{array}$$

such that  $|\phi'| \leq k|\phi|^2$  for some constant  $k$ . Furthermore, if  $\phi$  is epsilon-free then  $\phi'$  may be chosen to be epsilon-free.

*Proof.* I present transformations for each possible inference rule  $\rho$ .

1.

$$\begin{array}{ccc} \boxed{\boxed{\text{qc}\downarrow \frac{\exists x A \vee \exists x A}{\exists x A}} \wedge C} & \rightarrow & \boxed{\boxed{\boxed{\exists x A \vee \exists x A} \wedge \boxed{\text{c}\uparrow \frac{C}{C \wedge C}}} \\ \text{r1}\uparrow & & \text{s} \\ \exists x(A \wedge C) & & (\exists x A \vee (\exists x A \wedge C)) \wedge C \\ & & \text{s} \\ & & \boxed{\boxed{\text{r1}\uparrow \frac{\exists x A \wedge C}{\exists x(A \wedge C)}} \vee \boxed{\text{r1}\uparrow \frac{\exists x A \wedge C}{\exists x(A \wedge C)}} \\ & & \text{qc}\downarrow \\ & & \exists x(A \wedge C) \end{array}$$

2.

$$\text{qc}\downarrow \frac{\frac{\exists xA \vee \exists xA}{= \exists xA}}{A} \quad \rightarrow \quad \text{c}\downarrow \frac{\boxed{\frac{= \exists xA}{A}} \vee \boxed{\frac{= \exists xA}{A}}}{A}$$

where the occurrence of  $\mathbf{c}\downarrow$  in the derivation above is replaced with a derivation in  $\{\mathbf{ac}\downarrow, \mathbf{m}, \mathbf{qc}\downarrow, \forall, \mathbf{p}, =_{\mathbf{v}}\}$  using Lemma 3.13 and sequential composition.

3.

$$\frac{\frac{\forall y J \left\{ \frac{\text{qc} \downarrow \exists x A \vee \exists x A}{\exists x A} \right\}}{J\{\exists x A\}[t/y]}}{\forall} \quad \rightarrow \quad \frac{\forall y J\{\exists x A \vee \exists x A\}}{\forall} \frac{J \left\{ \frac{\text{qc} \downarrow \exists x A \vee \exists x A}{\exists x A} \right\} [t/y]}{\forall}$$

4.

$$\frac{\exists \frac{K \left\{ \frac{\text{qc} \downarrow \frac{\exists x A \vee \exists x A}{\exists x A}}{\text{qc} \downarrow} \right\} [t/y]}{\exists y K \{ \exists x A \}}}{\exists \frac{K \{ \exists x A \vee \exists x A \} [t/y]}{\exists y K \left\{ \frac{\text{qc} \downarrow \frac{\exists x A \vee \exists x A}{\exists x A}}{\text{qc} \downarrow} \right\}}}} \rightarrow$$

5.

$$\begin{array}{c}
\boxed{K \left\{ \frac{\text{qc} \downarrow \exists x A \vee \exists x A}{\exists x A} \right\}} \\
\text{c}\Uparrow \quad \boxed{\boxed{\phantom{K \{ \exists x A \} \wedge K \{ \exists x A \}}}} \\
K \{ \exists x A \} \wedge K \{ \exists x A \} \\
\downarrow \\
K \{ \exists x A \vee \exists x A \} \\
\text{c}\Uparrow \quad \boxed{K \left\{ \frac{\text{qc} \downarrow \exists x A \vee \exists x A}{\exists x A} \right\} \wedge K \left\{ \frac{\text{qc} \downarrow \exists x A \vee \exists x A}{\exists x A} \right\}}
\end{array}$$

6.

$$\frac{\text{qc}\downarrow \frac{\exists x A \vee \exists x A}{\exists x A}}{\exists y(A[y/x])} \rightarrow \frac{\frac{\frac{\exists x A}{\exists y(A[y/x])}}{\text{qc}\downarrow} \vee \frac{\frac{\exists x A}{\exists y(A[y/x])}}{\text{qc}\downarrow}}{\exists y(A[y/x])}$$

7.

$$\frac{s \overline{J \left\{ \frac{\text{qc} \downarrow \exists x A \vee \exists x A}{\exists x A} \right\} \wedge (C \vee D)}}{(J\{\exists x A\} \wedge C) \vee D} \rightarrow \frac{s \overline{J\{\exists x A \vee \exists x A\} \wedge (C \vee D)}}{J \left\{ \frac{\text{qc} \downarrow \exists x A \vee \exists x A}{\exists x A} \right\} \wedge C \vee D}$$

The remaining cases for  $s, m, r1\downarrow, r1\uparrow$  and the remaining  $=$  rules are similar to case (7) above in that they are immediately verified.  $\square$

I will now provide a lemma which corresponds to the first phase of the procedure, which permutes occurrences of the existential contraction rule  $qc\downarrow$  down to the bottom of an epsilon-free proof in  $(SKS1 \setminus \{qc\uparrow\}) \cup \{c\uparrow\}$ , using Lemma 5.2 to locally rewrite subderivations of the proof when performing the permutations.

As described in the previous subsection, Subsection 5.1, occurrences of the cocontraction rule  $c\uparrow$  in the proof are not decomposed prior to this phase of the procedure since, by Lemma 3.13, this would introduce further occurrences of vacuous  $=\exists$  rules into the proof. By case (2) in the proof of Lemma 5.2 above, such occurrences of vacuous  $=\exists$  rules may introduce further occurrences of  $qc\downarrow$  into the proof when permuting occurrences of  $qc\downarrow$  down through them, resulting in greater complexity during the procedure. After all occurrences of  $qc\downarrow$  have been permuted down to the bottom of the proof, occurrences of  $c\uparrow$  in the proof are then decomposed using Lemma 3.13.

**Lemma 5.3** (Phase 1, Existential Contraction Extraction). *For every epsilon-free proof  $\phi$  with conclusion  $A$  in  $(SKS1 \setminus \{qc\uparrow\}) \cup \{c\uparrow\}$ , there exists an epsilon-free proof of the form*

$$\begin{array}{c} \phi' \prod SKS1 \setminus \{qc\downarrow\} \\ A' \\ \prod \{qc\downarrow\} \\ A \end{array}$$

such that  $|\phi'| = \exp^3(O(|\phi| \ln |\phi|))$ . Furthermore, if  $\phi$  is cut-free then  $\phi'$  may be chosen to be cut-free.

*Proof.* We permute all occurrences of  $qc\downarrow$  down the proof using the rewriting system defined as follows. At each inductive step, we select a lowermost occurrence of  $qc\downarrow$  in the proof which occurs above some occurrence of an inference rule other than  $qc\downarrow$ . We then permute this occurrence of  $qc\downarrow$  down through a rule occurrence  $P$  of the form  $\rho \frac{B}{C}$  which occurs immediately below it in the proof in the following manner:

If  $P$  occurs inside the context of  $qc\downarrow$ , we apply the following rewrite, replacing the subderivation in the proof using sequential composition:

$$\begin{array}{c} \frac{\exists x K \{B\} \vee \exists x K \{B\}}{qc\downarrow \frac{\exists x K \left\{ \rho \frac{B}{C} \right\}}{\exists x K \{C\}}} \end{array} \rightarrow \frac{\boxed{\exists x K \left\{ \rho \frac{B}{C} \right\}} \vee \boxed{\exists x K \left\{ \rho \frac{B}{C} \right\}}}{qc\downarrow \frac{\boxed{\exists x K \{C\}}}{\exists x K \{C\}}}$$

Otherwise, if  $P$  occurs outside the context of the selected occurrence of  $qc\downarrow$ , in a subderivation of the form

$$\frac{\rho \frac{K \left\{ qc\downarrow \frac{\exists x D \vee \exists x D}{\exists x D} \right\}}{C}}{C}$$

for some formula context  $K\{\}$  such that  $K\{\exists x D\}$  is the formula  $B$ , we replace the above subderivation in the proof with the derivation given by Lemma 5.2 using sequential composition.

The procedure terminates once every occurrence of  $qc\downarrow$  in the proof occurs above only other occurrences of  $qc\downarrow$ . Each step of the rewriting system replaces a subderivation in the proof which consists of two rule occurrences, one of which is an occurrence of  $qc\downarrow$  which is also an uppermost rule occurrence in the subderivation, with a subderivation in which all occurrences of  $qc\downarrow$  occur as lowermost rule occurrences in the subderivation. Since this

operation can be performed on a proof only finitely many times before all occurrences of  $\text{qc}\downarrow$  occur at the bottom of the proof, termination is guaranteed.

The proof resulting from these permutations is of the form

$$\begin{array}{c} \phi'' \parallel \\ (SKS1 \setminus \{\text{qc}\downarrow, \text{qc}\uparrow\}) \cup \{\text{c}\uparrow\} \\ A' \\ \parallel \\ \{\text{qc}\downarrow\} \\ A \end{array}$$

To obtain a proof in the desired form, we replace all occurrences of  $\text{c}\uparrow$  in  $\phi''$  with derivations in  $\{\text{ac}\uparrow, \text{m}, \text{qc}\uparrow, \exists, =_{\mathbf{P}}, =_{\exists}\}$  using sequential composition and Lemma 3.13. The resultant proof is of the form

$$\begin{array}{c} \phi' \parallel \\ SKS1 \setminus \{\text{qc}\downarrow\} \\ A' \\ \parallel \\ \{\text{qc}\downarrow\} \\ A \end{array}$$

as required.

By Lemma 5.2, if  $\phi$  is cut-free then the rewrites presented do not introduce any further occurrences of  $\text{ai}\uparrow$  into the proof. Therefore, if  $\phi$  is cut-free then  $\phi'$  is cut-free. Similarly, by Lemma 5.2, since  $\phi$  is epsilon-free the rewrites presented do not introduce any  $\varepsilon$ -terms into the proof. Therefore  $\phi'$  is epsilon-free.

### Complexity

Each rewrite for permuting an occurrence of  $\text{qc}\downarrow$  down through a rule occurrence immediately below it in the proof replaces a subderivation  $\chi$  of the proof with a derivation of size at most  $k|\chi|^2$  for some constant  $k$ , by the rewrites described above and Lemma 5.2. Therefore

$$|\phi''| \leq k^{(2^N - 1)} |\phi|^{(2^N)} \quad (5)$$

where  $N$  is the number of occurrences of  $\text{qc}\downarrow$  permuted down the proof. As such, I compute an upper bound for the number of occurrences of  $\text{qc}\downarrow$  in the proof during the procedure.

The rewrites for permuting occurrences of  $\text{qc}\downarrow$  down through a rule occurrence immediately below them are described above and in the proof of Lemma 5.2. The reductions which potentially increase the number of  $\text{qc}\downarrow$  occurrences in the proof correspond to when an occurrence of  $\text{qc}\downarrow$  is permuted down through an occurrence of a vacuous  $=_{\exists}$  rule or an occurrence of the  $\text{c}\uparrow$  rule (cases (2) and (5) in the proof of Lemma 5.2):

$$\begin{array}{c} \text{qc}\downarrow \frac{\exists xB \vee \exists xB}{= \frac{\exists xB}{B}} \end{array} \rightarrow \begin{array}{c} \boxed{\boxed{= \frac{\exists xB}{B}} \vee \boxed{= \frac{\exists xB}{B}}} \\ \parallel \\ B \\ \{\text{ac}\downarrow, \text{m}, \text{qc}\downarrow, \forall, =_{\mathbf{P}}, =_{\forall}\} \end{array}$$

where the subderivation in  $\{\text{ac}\downarrow, \text{m}, \text{qc}\downarrow, \forall, =_{\mathbf{P}}, =_{\forall}\}$  in the derivation above results from decomposing an instance of the contraction rule  $\text{c}\downarrow$  on the formula  $B$  by Lemma 3.13 and hence contains at most  $|B|$  occurrences of  $\text{qc}\downarrow$ .

$$\begin{array}{c}
\boxed{\boxed{K \left\{ \frac{\boxed{\text{qc}\downarrow \frac{\exists xB \vee \exists xB}{\exists xB}}}{\exists xB} \right\}}} \\
\text{c}\uparrow \frac{}{K\{\exists xB\} \wedge K\{\exists xB\}} \\
\downarrow \\
\boxed{\boxed{K\{\exists xB \vee \exists xB\}}} \\
\text{c}\uparrow \frac{}{\boxed{K \left\{ \frac{\boxed{\text{qc}\downarrow \frac{\exists xB \vee \exists xB}{\exists xB}}}{\exists xB} \right\} \wedge K \left\{ \frac{\boxed{\text{qc}\downarrow \frac{\exists xB \vee \exists xB}{\exists xB}}}{\exists xB} \right\}}}
\end{array} \tag{6}$$

As such, I will consider which reductions potentially increase the number of occurrences of  $\text{c}\uparrow$  or vacuous  $=\exists$  rules in the proof and use this to provide a bound for the number of occurrences of  $\text{qc}\downarrow$  in the proof during the procedure. By the reductions described in the proof above and in the proof of Lemma 5.2, most of the subderivations resulting from permuting an occurrence of  $\text{qc}\downarrow$  down through a rule occurrence  $P$  immediately below it in the proof only contain occurrences of  $\text{c}\uparrow$  or a vacuous  $=\exists$  rule if  $P$  is respectively an occurrence of  $\text{c}\uparrow$  or a vacuous  $=\exists$  rule, and in which case the occurrences of these rules in the resultant subderivation occur in parallel in the subderivation so that the reduction does not increase the number of occurrences of  $\text{c}\uparrow$  or a vacuous  $=\exists$  which a single occurrence of  $\text{qc}\downarrow$  may be permuted down through during the procedure. The only exception to this is a particular reduction for permuting an occurrence of  $\text{qc}\downarrow$  down through an occurrence of  $\text{r1}\uparrow$  (case (1) in the proof of Lemma 5.2), which introduces a fresh occurrence of  $\text{c}\uparrow$  into the proof:

$$\begin{array}{c}
\boxed{\boxed{\text{qc}\downarrow \frac{\exists xB \vee \exists xB}{\exists xB} \wedge C}} \\
\text{r1}\uparrow \frac{}{\exists x(B \wedge C)}
\end{array}
\rightarrow
\begin{array}{c}
\boxed{\boxed{(\exists xB \vee \exists xB) \wedge \frac{\text{c}\uparrow \frac{C}{C \wedge C}}{C}}} \\
\text{s} \frac{}{(\exists xB \vee (\exists xB \wedge C)) \wedge C} \\
\text{s} \frac{}{\boxed{\boxed{\text{r1}\uparrow \frac{\exists xB \wedge C}{\exists x(B \wedge C)} \vee \text{r1}\uparrow \frac{\exists xB \wedge C}{\exists x(B \wedge C)}}}} \\
\text{qc}\downarrow \frac{}{\exists x(B \wedge C)}
\end{array}$$

As such, permuting multiple occurrences of  $\text{qc}\downarrow$  down through an occurrence of  $\text{r1}\uparrow$  in this manner results in a subderivation of the form

$$\begin{array}{c}
\boxed{\boxed{(\exists xB \vee \dots \vee \exists xB) \wedge \frac{C}{\frac{\text{c}\uparrow \frac{C}{C \wedge \dots \wedge C}}{C \wedge \dots \wedge C}}}} \\
\parallel \{s\} \\
\boxed{\boxed{\text{r1}\uparrow \frac{\exists xB \wedge C}{\exists x(B \wedge C)} \vee \dots \vee \text{r1}\uparrow \frac{\exists xB \wedge C}{\exists x(B \wedge C)}}} \\
\parallel \{\text{qc}\downarrow\} \\
\exists x(B \wedge C)
\end{array} \tag{7}$$

where the subderivation in  $\{\text{c}\uparrow\}$  in the derivation above contains  $n$  occurrences of  $\text{c}\uparrow$ , where  $n$  is the number of occurrences of  $\text{qc}\downarrow$  permuted down through the occurrence of  $\text{r1}\uparrow$ . By reduction (6) above, each occurrence of  $\text{c}\uparrow$  may introduce a further occurrence of  $\text{qc}\downarrow$  into the proof when an occurrence of  $\text{qc}\downarrow$  is permuted down through it. As such, an occurrence

of  $r1\uparrow$  in the proof which has  $n$  occurrences of  $qc\downarrow$  permuted down through it introduces at most  $n$  occurrences of  $c\uparrow$  into the proof, each of which potentially introduces a further occurrence of  $qc\downarrow$  into the proof for each occurrence of  $qc\downarrow$  permuted down, resulting in at most  $n^2$  occurrences of  $qc\downarrow$ .

It follows that for a given occurrence of an inference rule in the proof which has  $n$  occurrences of  $qc\downarrow$  permuted down through it, the greatest possible number of occurrences of  $qc\downarrow$  resulting from these permutations is  $n^2$ , when the inference rule is  $r1\uparrow$ .  $\phi$  contains at most  $|\phi|$  occurrences of  $qc\downarrow$  and  $|\phi|$  total occurrences of inference rules, yielding an upper bound of  $|\phi|^{(2|\phi|)}$  occurrences of  $qc\downarrow$  in the proof during the procedure. By (5) above, this yields the bound

$$|\phi''| = \exp^3(O(|\phi| \ln |\phi|))$$

Finally, since  $\phi'$  is obtained from  $\phi''$  by replacing all occurrences of  $c\uparrow$  in  $\phi''$  with derivations in  $\{ac\uparrow, m, qc\uparrow, \exists, =_P, =_\exists\}$  using sequential composition and Lemma 3.13, it follows that  $|\phi'| = O(|\phi''|^2)$  and hence

$$|\phi'| = \exp^3(O(|\phi| \ln |\phi|))$$

as required. □

### 5.3 Phase 2

In this subsection I will formalise the second phase of the proof of Theorem 4.13 the Falsifier Decomposition Theorem, in which occurrences of the existential witnessing rule  $\exists$  are permuted down to the bottom of the proof.

The following lemma provides reduction rules for permuting occurrences of the existential witnessing rule  $\exists$  down through occurrences of other inference rules, which will be used during the second phase of the procedure (see Lemma 5.5 below). Prior to performing these permutations, occurrences of the quantifier-shift  $r1\downarrow$  in the proof are replaced with equivalent occurrences of the falsifier rule  $\varepsilon$  to ensure that the permutations are possible. When an occurrence of  $\exists$  is permuted down through an occurrence of  $\varepsilon$ , an  $\varepsilon$ -term may be introduced into the proof (see case (3) in the proof below). The lemma does not provide a reduction for permuting occurrences of  $\exists$  down through occurrences of vacuous  $=_\exists$  rules, since these rule occurrences will be eliminated separately in the proof of Lemma 5.5.

**Lemma 5.4.** *For every inference rule  $\rho \in (SKS1 \setminus \{qc\downarrow, r1\downarrow, \exists\}) \cup \{\varepsilon\}$  which is not a vacuous  $=_\exists$  rule and every derivation  $\phi$  of the form*

$$\frac{\rho \left( K \left\{ \frac{\exists \frac{A(t)}{\exists x A(x)}}{\right\} \right)}{B}$$

where  $K\{\}$  is a formula context, there exists a derivation of the form

$$\frac{\frac{K\{A(t)\}}{\phi' \parallel \{\rho\}}}{B'} \parallel \{\exists\} B_\varepsilon$$

such that  $|\phi'| \leq |t||\phi|$ ,  $\phi'$  contains at most one occurrence of  $\rho$  and  $B_\varepsilon$  is a formula obtained by possibly replacing some  $\varepsilon$ -terms of the form  $\varepsilon_z(J\{\exists x C(x)\}[z/y])$  which occur in  $B$  with  $\varepsilon_z(J\{C(t)\}[z/y])$  for some formula context  $J\{\}$  and variables  $y, z$ .

*Proof.* I present transformations for each possible inference rule  $\rho$ .

1.

$$\begin{array}{c}
\boxed{\forall y K \left\{ \frac{\boxed{A(t)}}{\exists x A(x)} \right\}} \\
\text{qc}\uparrow \frac{}{\forall y K \{ \exists x A(x) \} \wedge \forall y K \{ \exists x A(x) \}} \\
\downarrow \\
\text{qc}\uparrow \frac{\forall y K \{ A(t) \}}{\boxed{\forall y K \left\{ \frac{\boxed{A(t)}}{\exists x A(x)} \right\} \wedge \forall y K \left\{ \frac{\boxed{A(t)}}{\exists x A(x)} \right\}}}
\end{array}$$

2.

$$\boxed{\forall y J \left\{ \frac{\boxed{A(t)}}{\exists x A(x)} \right\}} \quad \rightarrow \quad \forall \frac{\forall y J \{ A(t) \}}{\boxed{J \left\{ \frac{\boxed{A(t)}}{\exists x A(x)} \right\} [s/y]}}$$

3.

$$\begin{array}{c}
\boxed{\forall y \left( C(y) \vee J \left\{ \frac{\boxed{A(t)}}{\exists x A(x)} \right\} \right)} \\
\varepsilon \frac{}{\forall y C(y) \vee J \{ \exists x A(x) \} [\varepsilon_z \overline{C}(z)/y]} \\
\downarrow \\
\varepsilon \frac{\forall y (C(y) \vee J \{ A(t) \})}{\boxed{\forall y C(y) \vee J \left\{ \frac{\boxed{A(t)}}{\exists x A(x)} \right\} [\varepsilon_z \overline{C}(z)/y]}}
\end{array}$$

4.

$$\begin{array}{c}
\boxed{\forall y \left( J \left\{ \frac{\boxed{A(t)}}{\exists x A(x)} \right\} \vee C(y) \right)} \\
\varepsilon \frac{}{\forall y J \{ \exists x A(x) \} \vee C(\varepsilon_z \overline{J} \{ \exists x A(x) \} [z/y])} \\
\downarrow \\
\varepsilon \frac{\forall y (J \{ A(t) \} \vee C(y))}{\boxed{\forall y J \left\{ \frac{\boxed{A(t)}}{\exists x A(x)} \right\} \vee C(\varepsilon_z \overline{J} \{ A(t) \} [z/y])}}
\end{array}$$

5.

$$\boxed{\boxed{\frac{\boxed{A(t)}}{\exists x A(x)}} \wedge C} \quad \rightarrow \quad \frac{\boxed{A(t)} \wedge C}{\exists x (A(x) \wedge C)}$$



6.

$$\frac{\frac{\exists \frac{A(t)}{\exists x A(x)}}{= \exists y A(y)}}{\exists \frac{A(t)}{\exists y A(y)}} \rightarrow \frac{\exists \frac{A(t)}{\exists y A(y)}}{\exists y A(y)}$$

7.

$$\frac{s \left[ J \left\{ \frac{\exists \frac{A(t)}{\exists x A(x)}}{\exists x A(x)} \right\} \wedge (C \vee D) \right]}{(J\{\exists x A(x)\} \wedge C) \vee D} \rightarrow \frac{s \left[ J\{A(t)\} \wedge (C \vee D) \right]}{\left[ J \left\{ \frac{\exists \frac{A(t)}{\exists x A(x)}}{\exists x A(x)} \right\} \wedge C \right] \vee D}$$

The remaining cases for  $\mathbf{s}$ ,  $\mathbf{m}$ ,  $\mathbf{r1}\uparrow$  and the remaining  $=$  rules are similar to case (7) above in that they are immediately verified.  $\square$

The following lemma corresponds to the second phase of the procedure, which permutes occurrences of the existential witnessing rule  $\exists$  down to the bottom of the proof resulting from Phase 1 (Lemma 5.3), using Lemma 5.4 to locally rewrite subderivations in the proof when performing the permutations. During this phase of the procedure, all occurrences of existential rules other than  $\exists$  and  $\mathbf{qc}\downarrow$ , namely the quantifier-shift  $\mathbf{r1}\uparrow$  and existential equality rule  $=_{\exists}$ , are eliminated from the proof by the permutations so that the resultant proof is comprised of an upper segment consisting only of weakly existential-free formulae and a lower segment in  $\{\exists, \mathbf{qc}\downarrow\}$ . Eliminating occurrences of the vacuous  $=_{\exists}$  rule given by

$$\frac{=_{\exists} \frac{\exists x A}{A}}{A}$$

where  $x$  does not occur free in  $A$ , is delicate since existential quantifiers in  $A$  may be witnessed by terms which depend on  $x$ . Such occurrences of  $=_{\exists}$  are eliminated from the proof by appropriately ordering the occurrences of  $\exists$  being permuted down the proof during the procedure such that all occurrences of  $\exists$  which witness an existential quantifier in  $A$  occur inside the context of the occurrence of  $\exists$  which instantiates the existential quantifier in the occurrence of  $=_{\exists}$  above. The occurrence of  $=_{\exists}$  is then eliminated by the following transformation:

$$\frac{\frac{\exists \frac{A'[t/x]}{\exists x \left[ \frac{A'}{\chi \parallel \{\exists\}} \right] A}}{= A}}{\frac{A'[t/x]}{\chi[t/x] \parallel \{\exists\}} \frac{A}{A}} \rightarrow \frac{A'[t/x]}{\chi[t/x] \parallel \{\exists\}} \frac{A}{A}$$

**Lemma 5.5** (Phase 2, Existential Witness Extraction). *For every epsilon-free proof of the form*

$$\frac{\phi \parallel \text{SKS1} \setminus \{\mathbf{qc}\downarrow\}}{A' \parallel \{\mathbf{qc}\downarrow\}} \frac{A'}{A}$$

*there exists an epsilon-closed proof of the form*

$$\frac{\phi' \parallel (\text{SKS1} \setminus \{\mathbf{qc}\downarrow, \mathbf{r1}\downarrow, \mathbf{r1}\uparrow, \exists, =_{\exists}\}) \cup \{\varepsilon\}}{A'' \parallel \{\exists\}} \frac{A'}{A \parallel \{\mathbf{qc}\downarrow\}} \frac{A'}{A}$$

such that  $|\phi'| = \exp^2(O(|\phi| \ln |\phi|))$  and no two occurrences of  $\varepsilon$  in  $\phi'$  share the same variable name. Furthermore, if  $\phi$  is cut-free then  $\phi'$  may be chosen to be cut-free.

*Proof.* To begin, to ensure that occurrences of the  $\exists$  rule can be permuted down the proof, we replace every occurrence of  $r1\downarrow$  in the proof with an occurrence of the falsifier rule  $\varepsilon$ , using the following transformation:

$$r1\downarrow \frac{\forall x(B(x) \vee C)}{\forall x B(x) \vee C} \rightarrow \varepsilon \frac{\forall x(B(x) \vee C)}{\forall x B(x) \vee C}$$

To ensure that the desired conditions of the final proof are met, we assume that the variable names of all occurrences of the  $\varepsilon$  rule in the proof use distinct variables which are not used anywhere else in the proof.

We now permute occurrences of the  $\exists$  rule down the proof using the rewriting system defined as follows. At each inductive step, we select an uppermost occurrence of  $\exists$  in the proof which occurs above some occurrence of a rule other than  $\exists$  or  $qc\downarrow$ . We then permute this occurrence of  $\exists$  down through an occurrence  $P$  of an inference rule  $\rho$  immediately below it in the proof in the following manner:

If  $P$  is an occurrence of  $\exists$ , we order the rule occurrences so that the innermost occurrence occurs below the outermost occurrence, applying the following rewrite if necessary, replacing the subderivation in the proof using sequential composition:

$$\exists \frac{K \left\{ \exists \frac{B(t)}{\exists x B(x)} \right\} [s/y]}{\exists y K \{ \exists x B(x) \}} \rightarrow \exists \frac{K \{ B(t) \} [s/y]}{\exists y K \left\{ \exists \frac{B(t)}{\exists x B(x)} \right\}}$$

We then permute the lower occurrence of  $\exists$  down before permuting the upper occurrence down likewise. This guarantees that when an occurrence of  $\exists$  is being permuted down, it occurs in a subderivation of the form

$$\begin{array}{c} D \\ \chi \parallel \{ \exists \} \\ \boxed{J \left\{ \exists \frac{E(s)}{\exists y E(y)} \right\}} \end{array} \quad (8)$$

in which the displayed occurrence of  $\exists$  is the occurrence being permuted and all occurrences of  $\exists$  which occur above the permuted occurrence in the proof occur in  $\chi$  such that they act outside the context of the displayed occurrence of  $\exists$ , meaning their corresponding existential quantifiers occur outside the hole of  $J\{\}$ .

Otherwise, if  $P$  occurs inside the context of the selected occurrence of  $\exists$ , we apply the following rewrite, replacing the subderivation in the proof using sequential composition:

$$\exists \frac{K \{ B \} [t/x]}{\exists x K \left\{ \left[ \rho \frac{B}{C} \right] \right\}} \rightarrow \exists \frac{K \left\{ \left[ \rho \frac{B}{C} \right] \right\} [t/x]}{\exists x K \{ C \}} \quad (9)$$

Otherwise,  $P$  must occur outside the context of the occurrence of  $\exists$ , in a subderivation of the form

$$\rho \frac{K \left\{ \exists \frac{B(t)}{\exists x B(x)} \right\}}{C} \quad (10)$$

for some formula context  $K\{\}$ . If  $\rho$  is a vacuous  $=_{\exists}$  rule, we eliminate it in the following manner. Since, as described above, all occurrences of  $\exists$  which occur above the selected occurrence of  $\exists$  in the proof occur in a subderivation as given by (8) above, we apply the following rewrite, replacing the subderivation in the proof using sequential composition:

$$\frac{\exists \frac{B(t)}{\frac{\exists x \frac{B(x)}{\chi \parallel \{\exists\}} \frac{C}{=}}}{C}}{\rightarrow \frac{B(t)}{\chi[t/x] \parallel \{\exists\}} \frac{C}{=}}$$

Otherwise, if  $\rho$  is an inference rule other than a vacuous  $=_{\exists}$  rule, we replace the subderivation (10) in the proof with the derivation given by Lemma 5.4 using sequential composition. To maintain correctness of the proof, if an  $\varepsilon$ -term  $\varepsilon_z(\overline{J\{\exists x B(x)\}}[z/y])$  is locally renamed to  $\varepsilon_z(\overline{J\{B(t)\}}[z/y])$  by this reduction, we replace every occurrence of  $\varepsilon_z(\overline{J\{\exists x B(x)\}}[z/y])$  in the proof with  $\varepsilon_z(\overline{J\{B(t)\}}[z/y])$ . When performing this replacement, in the case of nested  $\varepsilon$ -terms, we replace innermost occurrences of the term  $\varepsilon_z(\overline{J\{\exists x B(x)\}}[z/y])$  before outermost occurrences.

The procedure terminates once every occurrence of  $\exists$  in the proof occurs above only other occurrences of  $\exists$  and  $\text{qc}\downarrow$ . Each step of the rewriting system replaces a subderivation in the proof which consists of two rule occurrences, one of which is an occurrence of  $\exists$  which is also an uppermost rule occurrence in the subderivation, with a subderivation in which all occurrences of  $\exists$  occur as lowermost rule occurrences in the subderivation. Since this operation can be performed on a proof only finitely many times before all occurrences of  $\exists$  occur at the bottom of the proof, termination is guaranteed.

The resultant proof is of the form

$$\begin{array}{c} \phi' \parallel (\text{SKS1} \setminus \{\text{qc}\downarrow, \text{r1}\downarrow, \text{r1}\uparrow, \exists, =_{\exists}\}) \cup \{\varepsilon\} \\ A'' \\ \parallel \{\exists\} \\ A' \\ \parallel \{\text{qc}\downarrow\} \\ A \end{array}$$

as required.

By Lemma 5.4, if  $\phi$  is cut-free then the rewrites presented do not introduce any further occurrences of  $\text{ai}\uparrow$  into the proof. Therefore, if  $\phi$  is cut-free then  $\phi'$  is cut-free.

When replacing occurrences of  $\text{r1}\downarrow$  in the proof with occurrences of  $\varepsilon$ , we assumed that the variable names of all such occurrences of  $\varepsilon$  used distinct variable names. It is easily verified that the rewrites presented do not alter the variable names of any occurrences of  $\varepsilon$  in the proof or introduce any new occurrences of  $\varepsilon$  into the proof which may use a new variable name. It follows that no two occurrences of  $\varepsilon$  in  $\phi'$  share the same variable name, as required.

### Epsilon-closure

I will now demonstrate that  $\phi'$  is epsilon-closed. During the procedure, each step of the rewriting system permutes uppermost occurrences of  $\exists$  in the proof down through rule occurrences which occur below them in the proof. As such, the proof after each step of the procedure is of the form

$$\begin{array}{c} \psi \parallel \\ B \\ \psi' \parallel \\ B' \\ \parallel \{\text{qc}\downarrow\} \\ A \end{array} \tag{11}$$

where  $\psi'$  is epsilon-free and  $\psi$  is a derivation which contains no existential quantifiers outside the scope of epsilon symbols, which we assume to be maximal in the sense that every uppermost occurrence of an inference rule in  $\psi'$  is an occurrence of  $\exists$ . I will show by induction on the procedure that the corresponding  $\psi$  after each step of the rewriting system is epsilon-closed, thereby demonstrating that  $\phi'$  is epsilon-closed, as desired.

Since the proof prior to the procedure is epsilon-free, it follows that when it is expressed in the form (11) above, the corresponding  $\psi$  is epsilon-free and therefore epsilon-closed.

Now suppose that the proof prior to a step of the rewriting system may be expressed in the form (11) such that the corresponding  $\psi$  is epsilon-closed and, for the sake of exposition, denote this  $\psi$  by  $\psi_0$ . Prior to the step of the rewriting system, the occurrence  $\chi$  of an inference rule  $\rho$  which occurrences of the  $\exists$  rule are permuted down through occurs in the proof in a subderivation of the form

$$\begin{array}{c} \psi_0 \parallel \\ C \\ \omega \parallel \{\exists\} \\ C' \\ \chi \parallel \{\rho\} \\ C'' \end{array}$$

where  $\omega$  is a derivation of the form (8) representing the occurrences of  $\exists$  to be permuted down through  $\chi$  and  $\chi$  and  $\omega$  are epsilon-free. After the permutations have been performed, by the reductions described above and Lemma 5.4, the subderivation above is replaced in the proof with a derivation of the form

$$\begin{array}{c} \psi_0 \parallel \\ C \\ \chi' \parallel \{\rho\} \\ C''' \\ \omega' \parallel \{\exists\} \\ C'''' \end{array}$$

where  $\chi'$  is either a formula or a single occurrence of  $\rho$ .  $\psi_0$  is not altered by the permutations even if  $\varepsilon$ -terms are renamed due to the assumption that all occurrences of  $\varepsilon$  in the proof use distinct variable names. As such, if the proof after the step of the rewriting system is expressed in the form (11), the corresponding  $\psi$  is given by  $\psi_0; \chi'$ . There are three possible cases: either (1)  $\rho$  is  $\forall$ , (2)  $\rho$  is  $\varepsilon$ , or (3)  $\rho$  is another inference rule.

In case (1), that  $\rho$  is  $\forall$ , the occurrences of  $\exists$  in  $\omega$  are permuted down through  $\chi$  by appropriately using reduction (9) in the proof above:

$$\exists \frac{K\{\forall y D(y)\}[t/x]}{\exists x K \left\{ \frac{\forall y D(y)}{D(s)} \right\}} \rightarrow \exists \frac{K \left\{ \frac{\forall y D(y)}{\frac{D(s)}{D(s)}} \right\} [t/x]}{\exists x K \{D(s)\}}$$

and reduction (2) in the proof of Lemma 5.4:

$$\forall \frac{\forall y K \left\{ \frac{D(t)}{\exists x D(x)} \right\}}{K \{ \exists x D(x) \} [s/y]} \rightarrow \forall \frac{\forall y K \{D(t)\}}{K \left\{ \frac{D(t)}{\exists x D(x)} \right\} [s/y]}$$

As such,  $\chi'$  is an occurrence of  $\forall$  of the form

$$\forall \frac{\forall x D(x)}{D(t[s_1/y_1] \dots [s_n/y_n])}$$

where  $t$  is the term that instantiates  $\chi$  and  $s_1, \dots, s_n$  are some of the terms that witness occurrences of  $\exists$  in  $\omega$ . Since  $\forall x D(x)$  is a subformula of the conclusion of  $\psi_0$  and  $\psi_0$  is epsilon-closed, every  $\varepsilon$ -term which occurs in  $\forall x D(x)$  must be constructed by some occurrence of  $\varepsilon$  or  $\forall$  in  $\psi_0$ . Since  $t$  is the term that instantiates  $\chi$ ,  $t$  must be epsilon-free and hence all  $\varepsilon$ -terms which occur in  $t[s_1/y_1] \dots [s_n/y_n]$  occur in a term of the form  $y_i[s_1/y_1] \dots [s_n/y_n]$  for some  $i \in \{1, \dots, n\}$ . By the two reductions presented above, all such terms occur in the conclusion of  $\psi_0$ . Since  $\psi_0$  is epsilon-closed, it follows that every  $\varepsilon$ -term which occurs in  $t[s_1/y_1] \dots [s_n/y_n]$  is constructed by some occurrence of  $\varepsilon$  or  $\forall$  in  $\psi_0$ . Now let  $\preceq_\varepsilon$  be an epsilon closure ordering for  $\psi_0$  and let  $\preceq_\varepsilon^+$  be the total order on occurrences of  $\varepsilon$  and  $\forall$  in  $\psi_0; \chi'$  given by extending  $\preceq_\varepsilon$  such that  $\gamma \prec_\varepsilon \chi'$  for all occurrences  $\gamma$  of  $\varepsilon$  and  $\forall$  in  $\psi_0$ .  $\preceq_\varepsilon^+$  meets the conditions of an epsilon closure ordering for  $\psi_0; \chi'$  and therefore  $\psi_0; \chi'$  is epsilon-closed, as desired.

In case (2), that  $\rho$  is  $\varepsilon$ , the occurrences of  $\exists$  in  $\omega$  are permuted down through  $\chi$  by appropriately using reduction (9) in the proof above:

$$\frac{\exists \frac{K\{\forall y(D(y) \vee E(y))\}[t/x]}{\exists x K \left\{ \frac{\forall y(D(y) \vee E(y))}{\forall y D(y) \vee E(\varepsilon_z \bar{D}(z))} \right\}}}{\exists x K \left\{ \frac{\forall y(D(y) \vee E(y))}{\forall y D(y) \vee E(\varepsilon_z \bar{D}(z))} \right\}}} \rightarrow \frac{\exists \frac{K \left\{ \frac{\forall y(D(y) \vee E(y))}{\forall y D(y) \vee E(\varepsilon_z \bar{D}(z))} \right\}}{K \left\{ \frac{\forall y(D(y) \vee E(y))}{\forall y D(y) \vee E(\varepsilon_z \bar{D}(z))} \right\}} [t/x]}{\exists x K \{\forall y D(y) \vee E(\varepsilon_z \bar{D}(z))\}}$$

reduction (3) in the proof of Lemma 5.4:

$$\frac{\forall y \left( E(y) \vee K \left\{ \frac{\exists \frac{D(t)}{\exists x D(x)}}{\exists x D(x)} \right\} \right)}{\forall y E(y) \vee K \{\exists x D(x)\}[\varepsilon_z \bar{E}(z)/y]} \xrightarrow{\varepsilon} \frac{\forall y (E(y) \vee K \{D(t)\})}{\forall y E(y) \vee K \left\{ \frac{\exists \frac{D(t)}{\exists x D(x)}}{\exists x D(x)} \right\} [\varepsilon_z \bar{E}(z)/y]} \xrightarrow{\varepsilon}$$

and reduction (4) in the proof of Lemma 5.4:

$$\frac{\forall y \left( K \left\{ \frac{\exists \frac{D(t)}{\exists x D(x)}}{\exists x D(x)} \right\} \vee E(y) \right)}{\forall y K \{\exists x D(x)\} \vee E(\varepsilon_z K \{\exists x D(x)\}[z/y])} \xrightarrow{\varepsilon} \frac{\forall y (K \{D(t)\} \vee E(y))}{\forall y K \left\{ \frac{\exists \frac{D(t)}{\exists x D(x)}}{\exists x D(x)} \right\} \vee E(\varepsilon_z K \{D(t)\}[z/y])} \xrightarrow{\varepsilon}$$

As such,  $\chi'$  is an occurrence of  $\varepsilon$  of the form

$$\frac{\forall x(D(x) \vee E(x))}{\forall x D(x) \vee E(\varepsilon_y D(y))} \xrightarrow{\varepsilon}$$

for some formulae  $D(x)$  and  $E(x)$  and variables  $x$  and  $y$ . Since  $\forall x(D(x) \vee E(x))$  is a subformula of the conclusion of  $\psi_0$  and  $\psi_0$  is epsilon-closed, every  $\varepsilon$ -term which occurs in  $\forall x(D(x) \vee E(x))$  must be constructed by some occurrence of  $\varepsilon$  or  $\forall$  in  $\psi_0$ . Let  $\preceq_\varepsilon$  be an epsilon closure ordering for  $\psi_0$  and let  $\preceq_\varepsilon^+$  be the total order on occurrences of  $\varepsilon$  and  $\forall$  in  $\psi_0; \chi'$  given by extending  $\preceq_\varepsilon$  such that  $\gamma \prec_\varepsilon \chi'$  for all occurrences  $\gamma$  of  $\varepsilon$  and  $\forall$  in  $\psi_0$ .  $\preceq_\varepsilon^+$  meets the conditions of an epsilon closure ordering for  $\psi_0; \chi'$  and therefore  $\psi_0; \chi'$  is epsilon-closed, as desired.

In case (3), that  $\rho$  is an inference rule other than  $\varepsilon$  or  $\forall$ , by the reductions described in the proof above and in the proof of Lemma 5.4, no new  $\varepsilon$ -terms are introduced into the proof so that every  $\varepsilon$ -term which occurs in  $\chi'$  is constructed by some occurrence of  $\varepsilon$  or  $\forall$  in  $\psi_0$ . It follows that  $\psi_0; \chi'$  is epsilon-closed since any epsilon closure ordering  $\preceq_\varepsilon$  for  $\psi_0$  is also an epsilon closure ordering for  $\psi_0; \chi'$ .

It follows that when the proof after each step of the rewriting procedure is expressed in the form (11), the corresponding  $\psi$  is epsilon-closed. Therefore  $\phi'$  is epsilon-closed.

### Complexity

Each rewrite for permuting an occurrence of the  $\exists$  rule which is witnessed by a term  $t$  down through a rule occurrence which occurs immediately below it in the proof replaces a subderivation  $\chi$  of the proof with a derivation of size at most  $|t||\chi|$ , by the rewrites described above and Lemma 5.4, since the existential quantifier in the conclusion of the rule occurrence is replaced by  $t$ . Therefore, since there are at most  $|\phi|$  existential quantifiers in  $\phi$ ,

$$|\phi'| \leq |t|^{|\phi|} |\phi| \quad (12)$$

where  $t$  is the largest term which witnesses an occurrence of  $\exists$  during the procedure. I therefore compute an upper bound for  $|t|$ .

There is only one possible type of reduction which increases the size of terms which instantiate occurrences of  $\exists$  during the procedure, when an occurrence of  $\exists$  is permuted down through an occurrence of  $\forall$  (case (2) in the proof of Lemma 5.4):

$$\frac{\forall x K \left\{ \frac{\exists \frac{B(s)}{\exists y B(y)}}{\exists y B(y)} \right\}}{K \{ \exists y B(y) \} [r/x]} \rightarrow \frac{\forall x K \{ B(s) \}}{K \left\{ \frac{\exists \frac{B(s)}{\exists y B(y)}}{\exists y B(y)} \right\} [r/x]} \quad (13)$$

Observe that the occurrence of  $\exists$  in the subderivation prior to the reduction is witnessed by  $s$  and that the occurrence of  $\exists$  in the subderivation after the reduction is witnessed by  $s[r/x]$ .

There is also one type of reduction which increases the size of terms which instantiate occurrences of  $\forall$  during the procedure, when an occurrence of  $\exists$  is permuted down an occurrence of  $\forall$  which occurs inside its context (see reduction (9) above):

$$\frac{\exists \frac{K \{ \forall x B(s) \}}{\exists y K \left\{ \frac{\forall x B(y)}{B(y) [r(y)/x]} \right\}}}{\exists y K \{ B(y) [r(y)/x] \}} \rightarrow \frac{K \left\{ \frac{\forall x B(s)}{B(s) [r(s)/x]} \right\}}{\exists y K \{ B(y) [r(y)/x] \}} \quad (14)$$

Observe that the occurrence of  $\forall$  in the subderivation prior to the reduction is instantiated by  $r(y)$  and that the occurrence of  $\forall$  in the subderivation after the reduction is instantiated by  $r(s)$ .

Consider a single occurrence of  $\forall$  in  $\phi$  which is instantiated by a term  $r$  and has  $M$  occurrences of  $\exists$  permuted down through it during the procedure, either by reduction (13) or (14), which are witnessed by terms  $s_1, \dots, s_M$ . After the occurrences of  $\exists$  have been permuted down, the occurrence of  $\forall$  will be instantiated by a term of size at most  $|r[s_1/x_1] \dots [s_M/x_M]| \leq |r||s|^M$ , where  $s$  is the largest of the terms  $s_i$ , by reduction (14).

Therefore each of the  $M$  occurrences of  $\exists$  will be witnessed by terms of size at most  $|r||s|^{M+1}$  after being permuted down through the occurrence of  $\forall$ , by reduction (13). If  $M$  occurrences of  $\exists$  are permuted down through  $L$  occurrences of  $\forall$ , the resultant occurrences of  $\exists$  are thus witnessed by terms of size at most  $|r|^{O(M^{L-1})}|s|^{O(M^L)}$ , where  $r$  is the largest term which instantiates one of the  $\forall$  occurrences.

Since  $r$  is a term which instantiates an occurrence of  $\forall$  in  $\phi$  and  $s$  is a term which witnesses an occurrence of  $\exists$  in  $\phi$ , we have  $|r|, |s| \leq |\phi|$ . Since  $\phi$  contains at most  $|\phi|$  occurrences of  $\forall$ , we have  $L \leq |\phi|$  and since it contains at most  $|\phi|$  existential quantifiers, we have  $M \leq |\phi|$ . Therefore the largest term which witnesses an occurrence of  $\exists$  during the procedure is of size at most  $|\phi|^{O(|\phi|^{|\phi|})}$ . By (12), this yields the bound

$$|\phi'| = \exp^2(O(|\phi| \ln |\phi|))$$

as required.  $\square$

### 5.4 Phase 3

In this subsection I will formalise the third and final phase of the proof of Theorem 4.13 the Falsifier Decomposition Theorem, in which occurrences of the universal cocontraction rule  $\text{qc}\uparrow$  are permuted up the proof until they are eliminated.

The following lemma provides reduction rules for permuting occurrences of the universal cocontraction rule  $\text{qc}\uparrow$  up through occurrences of most other inference rules, which will be used during the third phase of the procedure (see Lemma 5.8 below).

**Lemma 5.6.** *For every inference rule  $\rho \in \text{SKS1} \setminus \{\text{qc}\downarrow, \text{qc}\uparrow, \text{r1}\downarrow, \text{r1}\uparrow, \exists, =\exists\}$  and every derivation  $\phi$  of the form*

$$\rho \frac{B}{K \left\{ \frac{\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}}{} \right\}}$$

where  $K\{\}$  is a formula context and  $B$  is weakly existential-free, there exists a derivation of the form

$$\begin{array}{c} B \\ \parallel \{\text{qc}\uparrow\} \\ B' \\ \phi' \parallel \text{SKS1} \setminus \{\text{qc}\downarrow, \text{qc}\uparrow, \text{ai}\uparrow, \text{r1}\downarrow, \text{r1}\uparrow, \exists, =\exists\} \\ K\{\forall x A \wedge \forall x A\} \end{array}$$

such that  $|\phi'| \leq k|\phi|^2$  for some constant  $k$ .

*Proof.* I present transformations for each possible inference rule  $\rho$ .

1.

$$\frac{\forall y K\{\forall x A\}}{K \left\{ \frac{\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}}{} \right\} [t/y]} \rightarrow \frac{\forall y K \left\{ \frac{\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}}{} \right\}}{K\{\forall x A \wedge \forall x A\} [t/y]}$$

2.

$$\frac{= \frac{A}{\forall x A}}{\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}} \rightarrow \frac{\text{c}\uparrow \frac{A}{\frac{= \frac{A}{\forall x A} \wedge \frac{= \frac{A}{\forall x A}}}}{= \frac{A}{\forall x A} \wedge \frac{= \frac{A}{\forall x A}}}}$$

where the occurrence of  $\text{c}\uparrow$  in the derivation above is replaced with a derivation in  $\{\text{ac}\uparrow, \text{m}, \text{qc}\uparrow, =_{\text{P}}\}$  using Lemma 3.13 and sequential composition.

3.

$$\frac{\frac{\forall y(A[y/x])}{\forall x A}}{\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}} \rightarrow \frac{\forall y(A[y/x])}{\text{qc}\uparrow \left[ \frac{\forall y(A[y/x])}{\forall x A} \wedge \frac{\forall y(A[y/x])}{\forall x A} \right]}$$

4.

$$\frac{s \left[ J \left\{ \frac{\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}}{\forall x A} \right\} \wedge C \vee D \right]}{J \left\{ \frac{\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}}{\forall x A} \right\} \wedge (C \vee D)} \rightarrow \frac{s \left[ J \left\{ \frac{\text{qc}\uparrow \frac{\forall x A}{\forall x A \wedge \forall x A}}{\forall x A} \right\} \wedge (C \vee D) \right]}{(J \{ \forall x A \wedge \forall x A \} \wedge C) \vee D}$$

The remaining cases for **s**, **m** and the remaining  $=$  rules are similar to case (4) above in that they are immediately verified.  $\square$

During the third phase of the procedure (see Lemma 5.8 below), when occurrences of the universal cocontraction rule  $\text{qc}\uparrow$  are permuted up the proof, the greatest source of complexity and most troublesome case is when occurrences of  $\text{qc}\uparrow$  are permuted up through occurrences of the falsifier rule  $\varepsilon$ . The following lemma provides reduction rules for permuting occurrences of  $\text{qc}\uparrow$  up through a certain construction which is invariant under the permutation, and which is a generalisation of a single occurrence of the falsifier rule  $\varepsilon$ . Observe that the construction given in this lemma resembles the resultant derivation from transformation (3) in Subsection 4.2, reflecting that the falsifier rule  $\varepsilon$  and more expressive syntax of the epsilon-calculus are able to express that certain terms in the derivation are equal so that occurrences of the regular contraction rule  $\text{c}\downarrow$  may be used instead of occurrences of the existential contraction rule  $\text{qc}\downarrow$  which would be present in the epsilon-free syntax of predicate logic, representing superfluous case analyses.

**Lemma 5.7.** *For all variables  $x$  and  $y$  and all weakly existential-free formulae  $A(x)$  and  $B(x)$ , let  $D(A(x), B(x), x, y, n)$  denote the derivation*

$$\frac{\left[ \frac{\varepsilon \frac{\forall x(A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))}}{\wedge \dots \wedge} \frac{\varepsilon \frac{\forall x(A(x) \vee B(x))}{\forall x A(x) \vee B(\varepsilon_y \bar{A}(y))}}{\right]} \parallel_{\{s\}} \left[ (\forall x A(x) \wedge \dots \wedge \forall x A(x)) \vee \frac{B(\varepsilon_y \bar{A}(y)) \vee \dots \vee B(\varepsilon_y \bar{A}(y))}{\parallel_{\{c\downarrow\}} B(\varepsilon_y \bar{A}(y))} \right]$$

where the premise of the derivation is a conjunction of  $n$  copies of the formula  $\forall x(A(x) \vee B(x))$ .

For all derivations of the form

$$\frac{\forall x(A(x) \vee B(x)) \wedge \dots \wedge \forall x(A(x) \vee B(x))}{D(A(x), B(x), x, y, n)} \parallel \left[ K \left\{ \frac{\text{qc}\uparrow \frac{\forall z C}{\forall z C \wedge \forall z C}}{\right\} \right]$$



where  $K\{\}$  is a formula context, there exists a derivation of the form

$$\begin{array}{c}
\forall x(A(x) \vee B(x)) \wedge \cdots \wedge \forall x(A(x) \vee B(x)) \\
\parallel \{\mathbf{qc}\uparrow\} \\
\forall x(A'(x) \vee B'(x)) \wedge \cdots \wedge \forall x(A'(x) \vee B'(x)) \\
D(A'(x), B'(x), x, y, n') \parallel \\
(\forall x A'(x) \wedge \cdots \wedge \forall x A'(x)) \vee B'(\varepsilon_y \overline{A'}(y)) \\
\parallel \{\mathbf{aw}\uparrow, \forall, =_P\} \\
(\forall x A(x) \wedge \cdots \wedge \forall x A'(x) \wedge \cdots \wedge \forall x A(x)) \vee B'(\varepsilon_y \overline{A'}(y))
\end{array}$$

of size  $O((n+1)^2(|A(x)| + |B(x)|))$ , where  $A'(x)$ ,  $B'(x)$  and  $n'$  are given by one of the following

1.  $A'(x)$  is obtained by replacing a subformula  $\forall zC$  of  $A(x)$  with  $\forall zC \wedge \forall zC$ ,  $B'(x)$  is the formula  $B(x)$  and  $n' = n$
2.  $A'(x)$  is the formula  $A(x)$ ,  $B'(x)$  is obtained by replacing a subformula  $\forall zC$  of  $B(x)$  with  $\forall zC \wedge \forall zC$  and  $n' = n$
3.  $A'(x)$  is the formula  $A(x)$ ,  $B'(x)$  is the formula  $B(x)$  and  $n' = n + 1$

*Proof.* There are three possible cases: either (1) the occurrence of  $\mathbf{qc}\uparrow$  acts on a subformula  $\forall zC$  of  $A(x)$ , (2) the occurrence of  $\mathbf{qc}\uparrow$  acts on a subformula  $\forall zC$  of  $B(\varepsilon_y \overline{A'}(y))$  or (3) the occurrence of  $\mathbf{qc}\uparrow$  acts on the formula  $\forall x A(x)$ . The transformations to be applied in each case are respectively shown in Figures 4, 5 and 6.

Observe that  $|D(A(x), B(x), x, y, n)| = O(n^2(|A(x)| + |B(x)|))$  for all weakly existential-free formulae  $A(x)$ ,  $B(x)$ , all variables  $x$ ,  $y$  and  $n \in \mathbb{N}$ . It follows from Lemma 3.14 that the subderivation in  $\{\mathbf{aw}\uparrow, \forall, =_P\}$  in case (1) is of size  $O(n|A(x)|)$ . Hence the resultant derivation in each case is of size  $O((n+1)^2(|A(x)| + |B(x)|))$ .  $\square$

With these lemmas in place, I can now state the lemma that corresponds to the third phase of the procedure, which permutes occurrences of the universal cocontraction rule  $\mathbf{qc}\uparrow$  up the proof resulting from Phase 2 (Lemma 5.5) until they are eliminated, using Lemmas 5.6 and 5.7 to locally rewrite subderivations in the proof when performing the permutations. The proof resulting from this phase of the procedure is in falsifier normal form, and this lemma includes further information about the structure of the proof than was given in the statement of Theorem 4.13 the Falsifier Decomposition Theorem.

**Lemma 5.8** (Phase 3, Universal Cocontraction Elimination). *For every epsilon-closed proof of the form*

$$\begin{array}{c}
\phi \parallel (\text{SKS1} \setminus \{\mathbf{qc}\downarrow, \mathbf{r1}\downarrow, \mathbf{r1}\uparrow, \exists, =_\exists\}) \cup \{\varepsilon\} \\
A' \\
\parallel \{\exists\} \\
A'' \\
\parallel \{\mathbf{qc}\downarrow\} \\
A
\end{array}$$

such that no two occurrences of  $\varepsilon$  in  $\phi$  share the same variable name, there exists an epsilon-closed proof of the form

$$\begin{array}{c}
\phi' \parallel \text{SKSg}\varepsilon \\
A''' \\
\parallel \{\exists\} \\
A'' \\
\parallel \{\mathbf{qc}\downarrow\} \\
A
\end{array}$$

$$\begin{array}{c}
\forall x(A(x) \vee B(x)) \wedge \dots \wedge \forall x(A(x) \vee B(x)) \\
\parallel \\
D(A(x), B(x), x, y, n) \\
\boxed{\forall x A(x) \wedge \dots \wedge \forall x J \left\{ \frac{\forall z C}{\forall z C \wedge \forall z C} \right\} \wedge \dots \wedge \forall x A(x) \vee B(\varepsilon_y \overline{A}(y))} \\
\downarrow \\
\boxed{\forall x \left\{ J \left\{ \frac{\forall z C}{\forall z C \wedge \forall z C} \right\} \vee B(x) \right\} \wedge \dots \wedge \forall x \left\{ J \left\{ \frac{\forall z C}{\forall z C \wedge \forall z C} \right\} \vee B(x) \right\}} \\
\parallel \\
D(J\{\forall z C \wedge \forall z C\}, B(x), x, y, n) \\
\boxed{\forall x J \left\{ \forall z C \wedge \frac{\forall z C}{\mathbf{t}} \right\} \wedge \dots \wedge \forall x J \{\forall z C \wedge \forall z C\} \wedge \dots \wedge \forall x J \left\{ \forall z C \wedge \frac{\forall z C}{\mathbf{t}} \right\} \vee B(\varepsilon_y \overline{J\{\forall z C \wedge \forall z C\}})} \\
\parallel \\
(\forall x A(x) \wedge \dots \wedge \forall x J \{\forall z C \wedge \forall z C\} \wedge \dots \wedge \forall x A(x)) \vee B(\varepsilon_y \overline{J\{\forall z C \wedge \forall z C\}})
\end{array}$$

where  $J\{\forall z C\}$  is the formula  $A(x)$  and each occurrence of the  $\mathbf{w}\uparrow$  rule in the derivation above is replaced with a derivation in  $\{\mathbf{aw}\uparrow, \forall, =\mathbf{p}\}$  using Lemma 3.14 and sequential composition.

**Figure 4:** Transformation for case (1) of Lemma 5.7

$$\begin{array}{c}
\forall x(A(x) \vee B(x)) \wedge \dots \wedge \forall x(A(x) \vee B(x)) \\
D(A(x), B(x), x, y, n) \parallel \\
\boxed{(\forall x A(x) \wedge \dots \wedge \forall x A(x)) \vee J \left\{ \boxed{\frac{\forall z C}{\forall z C \wedge \forall z C}} \right\} [\varepsilon_y \bar{A}(y)/x]} \\
\downarrow \\
\boxed{\forall x \left[ A(x) \vee J \left\{ \boxed{\frac{\forall z C}{\forall z C \wedge \forall z C}} \right\} \right] \wedge \dots \wedge \forall x \left[ A(x) \vee J \left\{ \boxed{\frac{\forall z C}{\forall z C \wedge \forall z C}} \right\} \right]} \\
D(A(x), J\{\forall z C \wedge \forall z C\}, x, y, n) \parallel \\
(\forall x A(x) \wedge \dots \wedge \forall x A(x)) \vee J\{\forall z C \wedge \forall z C\}[\varepsilon_y \bar{A}(y)/x]
\end{array}$$

where  $J\{\forall z C\}$  is the formula  $B(x)$ .

**Figure 5:** Transformation for case (2) of Lemma 5.7

$$\begin{array}{c}
\forall x(A(x) \vee B(x)) \wedge \dots \wedge \forall x(A(x) \vee B(x)) \\
D(A(x), B(x), x, y, n) \parallel \\
\boxed{\forall x A(x) \wedge \dots \wedge \boxed{\frac{\forall x A(x)}{\forall x A(x) \wedge \forall x A(x)}} \wedge \dots \wedge \forall x A(x) \vee B(\varepsilon_y \bar{A}(y))} \\
\downarrow \\
\boxed{\forall x(A(x) \vee B(x)) \wedge \dots \wedge \boxed{\frac{\forall x(A(x) \vee B(x))}{\forall x(A(x) \vee B(x)) \wedge \forall x(A(x) \vee B(x))}} \wedge \dots \wedge \forall x(A(x) \vee B(x))} \\
D(A(x), B(x), x, y, n+1) \parallel \\
(\forall x A(x) \wedge \dots \wedge \forall x A(x)) \vee B(\varepsilon_y \bar{A}(y))
\end{array}$$

**Figure 6:** Transformation for case (3) of Lemma 5.7

such that  $|\phi'| = \exp^3(O(|\phi| \ln |\phi|))$ ,  $|A'''| = |A'|$  and every occurrence of  $\varepsilon$  in  $\phi'$  occurs within a subderivation of the form

$$\begin{array}{c}
\boxed{\frac{\varepsilon \frac{\forall x(B(x) \vee C(x))}{\forall x B(x) \vee C(\varepsilon_y \overline{B}(y))}}{\forall x B(x) \vee C(\varepsilon_y \overline{B}(y))}} \wedge \dots \wedge \boxed{\frac{\varepsilon \frac{\forall x(B(x) \vee C(x))}{\forall x B(x) \vee C(\varepsilon_y \overline{B}(y))}}{\forall x B(x) \vee C(\varepsilon_y \overline{B}(y))}} \\
\parallel_{\{s\}} \\
\boxed{(\forall x B(x) \wedge \dots \wedge \forall x B(x)) \vee \frac{C(\varepsilon_y \overline{B}(y)) \vee \dots \vee C(\varepsilon_y \overline{B}(y))}{\parallel_{\{c, \downarrow, m, \forall, =_P, =_V\}} C(\varepsilon_y \overline{B}(y))}}
\end{array}$$

where every occurrence of  $\varepsilon$  in  $\phi'$  with variable name  $y$  is contained in the subderivation above. Furthermore, if  $\phi$  is cut-free then  $\phi'$  may be chosen to be cut-free.

*Proof.* We permute occurrences of  $\mathbf{qc}\uparrow$  up the proof using the rewriting system defined as follows. At each inductive step, we select an uppermost occurrence of  $\mathbf{qc}\uparrow$  in the proof and permute it up through an occurrence  $P$  of an inference rule  $\rho$  which occurs immediately above it in the proof in the following manner:

If  $P$  occurs inside the context of the selected occurrence of  $\mathbf{qc}\uparrow$ , we apply the following rewrite, replacing the subderivation in the proof using sequential composition:

$$\frac{\mathbf{qc}\uparrow \boxed{\forall x K \left\{ \frac{\rho \frac{C}{B}}{B} \right\}}}{\forall x K \{B\} \wedge \forall x K \{B\}} \rightarrow \frac{\forall x K \{C\}}{\mathbf{qc}\uparrow \boxed{\boxed{\forall x K \left\{ \frac{\rho \frac{C}{B}}{B} \right\}} \wedge \boxed{\forall x K \left\{ \frac{\rho \frac{C}{B}}{B} \right\}}}}$$

Otherwise, if the selected occurrence of  $\mathbf{qc}\uparrow$  occurs in the proof in a subderivation of the form

$$\frac{\forall x(B(x) \vee C(x)) \wedge \dots \wedge \forall x(B(x) \vee C(x))}{D(B(x), C(x), x, y, n)} \parallel \boxed{K \left\{ \frac{\mathbf{qc}\uparrow \frac{\forall z E}{\forall z E \wedge \forall z E}}{\forall z E \wedge \forall z E} \right\}}$$

where  $D(B(x), C(x), x, y, n)$  is a derivation of the form described in Lemma 5.7, we replace the whole subderivation with the appropriate of the three derivations described in Lemma 5.7, using sequential composition. To maintain correctness of the proof, if an  $\varepsilon$ -term is locally renamed by this reduction (case (1) of Lemma 5.7), we replace every occurrence of the  $\varepsilon$ -term in the proof with the renamed  $\varepsilon$ -term. When performing this replacement, in the case of nested  $\varepsilon$ -terms, we replace innermost occurrences of the  $\varepsilon$ -term before outermost occurrences.

Otherwise, if  $P$  occurs outside the context of the selected occurrence of  $\mathbf{qc}\uparrow$ , in a subderivation of the form

$$\frac{C}{\rho \boxed{K \left\{ \frac{\mathbf{qc}\uparrow \frac{\forall x B}{\forall x B \wedge \forall x B}}{\forall x B \wedge \forall x B} \right\}}}$$

for some formula context  $K\{\}$ , we replace the above subderivation in the proof with the derivation given by Lemma 5.6 using sequential composition. Note that if  $P$  is an occurrence of  $\varepsilon$ , it is of the form  $D(E(x), F(x), x, y, 1)$  as described in Lemma 5.7 and hence is handled by the reduction described in the previous paragraph.

The procedure terminates once the proof contains no occurrences of  $\mathbf{qc}\uparrow$ . Every universal quantifier in the proof must be introduced by an occurrence of a vacuous  $=_{\forall}$  rule and when an occurrence of  $\mathbf{qc}\uparrow$  is permuted up through such a rule occurrence it is eliminated, introducing one further occurrence of  $\mathbf{qc}\uparrow$  into the proof for each universal quantifier in the premise of the occurrence of  $=_{\forall}$ , by reduction (2) in the proof of Lemma 5.6 and Lemma 3.13. It follows that termination of the procedure is guaranteed, since every occurrence of  $\mathbf{qc}\uparrow$  in the proof must eventually be eliminated by permuted up through an occurrence of a vacuous  $=_{\forall}$  rule which contains no universal quantifiers in its premise.

After termination, we replace every occurrence of  $\mathbf{c}\downarrow$  in the proof (resulting from the constructions of Lemma 5.7) with derivations in  $\{\mathbf{ac}\downarrow, \mathbf{m}, \forall, =_{\mathbf{P}}, =_{\forall}\}$  using Lemma 3.13 and sequential composition. We then replace all occurrences of atomic rules in  $\mathbf{SKS1}$  in the proof with equivalent occurrences of their regular variants in  $\mathbf{SKSg}_{\mathbf{P}}$  (occurrences of the atomic contraction rule  $\mathbf{ac}\downarrow$  are replaced with equivalent occurrences of the contraction rule  $\mathbf{c}\downarrow$ , etc.) so that they are all occurrences of rules in  $\mathbf{SKSg}_{\varepsilon}$ . The resultant proof is of the form

$$\begin{array}{c} \phi' \\ \parallel \mathbf{SKSg}_{\varepsilon} \\ A''' \\ \parallel \{\exists\} \\ A'' \\ \parallel \{\mathbf{qc}\downarrow\} \\ A \end{array}$$

as required.

By Lemmas 5.6 and 5.7, if  $\phi'$  is cut-free then the rewrites presented do not introduce any further occurrences of  $\mathbf{ai}\uparrow$  into the proof. Therefore, if  $\phi$  is cut-free then  $\phi'$  is cut-free.

It is easily verified that the proofs obtained after each step of the rewriting system are epsilon-closed since none of the rewrites described in the proof above or in the proofs of Lemmas 5.6 and 5.7 introduce new  $\varepsilon$ -terms into the proof, and instead only rename  $\varepsilon$ -terms in a manner that preserves epsilon closure orderings. If an occurrence  $\chi$  of  $\varepsilon$  or  $\forall$  in the proof is duplicated into two occurrences  $\chi'$  and  $\chi''$  by a step of the rewriting system, an epsilon closure ordering  $\preceq_{\varepsilon}^{+}$  may be obtained for the resultant proof by modifying an epsilon closure ordering  $\preceq_{\varepsilon}$  for the proof prior to the rewrite such that  $\omega \preceq_{\varepsilon}^{+} \chi' \preceq_{\varepsilon}^{+} \chi'' \preceq_{\varepsilon}^{+} \omega'$  for all occurrences  $\omega$  and  $\omega'$  of  $\varepsilon$  or  $\forall$  in the proof such that  $\omega \preceq_{\varepsilon} \chi \preceq_{\varepsilon} \omega'$  and  $\pi \preceq_{\varepsilon}^{+} \pi'$  if and only if  $\pi \preceq_{\varepsilon} \pi'$  for all occurrences of  $\varepsilon$  or  $\forall$  in the proof other than  $\chi$ . It follows that  $\phi'$  is epsilon-closed.

Furthermore, since no two occurrences of  $\varepsilon$  in  $\phi$  share the same variable name, it follows from the rewrites described in the proof above and in the proof of Lemma 5.7 that every occurrence of  $\varepsilon$  in  $\phi'$  occurs within a subderivation of the form

$$\begin{array}{c} \boxed{\frac{\varepsilon \quad \forall x(B(x) \vee C(x))}{\forall x B(x) \vee C(\varepsilon_y \overline{B}(y))} \wedge \dots \wedge \frac{\varepsilon \quad \forall x(B(x) \vee C(x))}{\forall x B(x) \vee C(\varepsilon_y \overline{B}(y))}} \\ \parallel \{\mathbf{s}\} \\ \boxed{(\forall x B(x) \wedge \dots \wedge \forall x B(x)) \vee \begin{array}{c} C(\varepsilon_y \overline{B}(y)) \vee \dots \vee C(\varepsilon_y \overline{B}(y)) \\ \parallel \{\mathbf{c}\downarrow, \mathbf{m}, \forall, =_{\mathbf{P}}, =_{\forall}\} \\ C(\varepsilon_y \overline{B}(y)) \end{array}} \end{array}$$

where every occurrence of  $\varepsilon$  in  $\phi'$  with variable name  $y$  is contained in the subderivation above.

### Complexity

The complexity increase from the procedure is analogous to that of Phase 1 (Lemma 5.3), since each rewrite for permuting an occurrence of  $\mathbf{qc}\uparrow$  up through a rule occurrence

immediately above it replaces a subderivation  $\chi$  of the proof with a derivation of size at most  $k|\chi|^2$  for some constant  $k$  and the constructions of Lemma 5.7 duplicate  $\mathbf{qc}\uparrow$  occurrences in the same manner that  $\mathbf{r1}\uparrow$  occurrences duplicate  $\mathbf{qc}\downarrow$  occurrences in the proof of Lemma 5.3. Therefore

$$|\phi'| = \exp^3(O(|\phi| \ln |\phi|))$$

Finally, since  $A'''$  is obtained by renaming  $\varepsilon$ -terms in  $A'$  during the procedure,  $|A'''| = |A'|$ .  $\square$

## 5.5 Final Proof

Theorem 4.13 the Falsifier Decomposition Theorem may now be proved by successively applying the above lemmas corresponding to the three phases of the procedure to an epsilon-free proof in SKSg1, as follows.

*Proof of Theorem 4.13.* Let  $\phi$  be an epsilon-free proof with conclusion  $A$  in SKSg1. By Lemma 5.1, there exists an epsilon-free proof  $\psi$  with conclusion  $A$  in  $(\text{SKS1} \setminus \{\mathbf{qc}\uparrow\}) \cup \{\mathbf{c}\uparrow\}$  such that  $|\psi| = O(|\phi|^2)$  and if  $\phi$  is cut-free then  $\psi$  is cut-free.

It follows from Lemma 5.3 that there exists an epsilon-free proof of the form

$$\begin{array}{c} \psi_1 \prod \text{SKS1} \setminus \{\mathbf{qc}\downarrow\} \\ A_1 \\ \prod \{\mathbf{qc}\downarrow\} \\ A \end{array}$$

such that

$$|\psi_1| = \exp^3(O(|\psi| \ln |\psi|)) = \exp^3(O(|\phi|^2 \ln |\phi|))$$

and if  $\phi$  is cut-free then  $\psi_1$  is cut-free.

It follows from Lemma 5.5 that there exists an epsilon-closed proof of the form

$$\begin{array}{c} \psi_2 \prod (\text{SKS1} \setminus \{\mathbf{qc}\downarrow, \mathbf{r1}\downarrow, \mathbf{r1}\uparrow, \exists, =\exists\}) \cup \{\varepsilon\} \\ A_2 \\ \prod \{\exists\} \\ A_1 \\ \prod \{\mathbf{qc}\downarrow\} \\ A \end{array}$$

such that

$$|\psi_2| = \exp^2(O(|\psi_1| \ln |\psi_1|)) = \exp^5(O(|\phi|^2 \ln |\phi|)),$$

no two occurrences of  $\varepsilon$  in  $\psi_2$  share the same variable name and if  $\phi$  is cut-free then  $\psi_2$  is cut-free.

It follows from Lemma 5.8 that there exists an epsilon-closed proof of the form

$$\begin{array}{c} \psi_3 \prod \text{SKSg}\varepsilon \\ A_3 \\ \prod \{\exists\} \\ A_1 \\ \prod \{\mathbf{qc}\downarrow\} \\ A \end{array}$$

such that

$$|\psi_3| = \exp^3(O(|\psi_2| \ln |\psi_2|)) = \exp^8(O(|\phi|^2 \ln |\phi|)),$$

$$|A_3| = |A_2| \leq |\psi_2| = \exp^5(O(|\phi|^2 \ln |\phi|))$$

and if  $\phi$  is cut-free then  $\psi_3$  is cut-free. By Lemma 4.12,

$$|\psi_3|_\varepsilon = O(\exp(\exp |\psi_3| \ln |\psi_3|)) = \exp^{10}(O(|\phi|^2 \ln |\phi|))$$

and therefore

$$|A_3|_\varepsilon \leq |\psi_3|_\varepsilon = \exp^{10}(O(|\phi|^2 \ln |\phi|))$$

Thus  $\phi' = \psi_3$ ,  $A' = A_3$  and  $A'' = A_1$  are as desired.  $\square$

## 5.6 Optimisations

The procedure presented above for transforming an epsilon-free proof in **SKSg1** into falsifier normal form is not optimal, but demonstrates elementary complexity of the size and  $\varepsilon$ -size of the proof resulting from the procedure with respect to the size of the original proof. I will now present some optimisations which can be made during the procedure. Although these optimisations are only of polynomial complexity, I believe them to be of interest for the sake of retaining the structure of the initial proof prior to the procedure.

When decomposing occurrences of inference rules in the proof using Lemma 5.1, it is not necessary to decompose all occurrences of the quantifier-shifts  $r2\downarrow$  and  $r2\uparrow$  in the proof into derivations in **SKS1**. This has the benefit of introducing fewer occurrences of the existential contraction rule  $qc\downarrow$  and universal cocontraction rule  $qc\uparrow$  into the proof, representing fewer superfluous case analyses. Instead of decomposing occurrences of  $r2\downarrow$  and  $r2\uparrow$  in the proof, we may handle them in the following manner.

During Phase 1 (Lemma 5.3), when permuting occurrences of  $qc\downarrow$  down the proof, they trivially permute through occurrences of  $r2\downarrow$  by the following reductions:

$$\begin{array}{ccc} \frac{\forall x \left( K \left\{ \frac{qc\downarrow \frac{\exists y A \vee \exists y A}{\exists y A}}{\exists y A} \right\} \wedge B \right)}{r2\downarrow \frac{\forall x K \{ \exists y A \} \wedge B}{\forall x K \{ \exists y A \} \wedge B}} & \rightarrow & \frac{\forall x (K \{ \exists y A \vee \exists y A \} \wedge B)}{r2\downarrow \frac{\forall x K \left\{ \frac{qc\downarrow \frac{\exists y A \vee \exists y A}{\exists y A}}{\exists y A} \right\} \wedge B}{\forall x K \{ \exists y A \} \wedge B}} \\[10pt] \frac{\forall x \left( A(x) \wedge K \left\{ \frac{qc\downarrow \frac{\exists y B \vee \exists y B}{\exists y B}}{\exists y B} \right\} \right)}{r2\downarrow \frac{\forall x A(x) \wedge K \{ \exists y B \}}{\forall x A(x) \wedge K \{ \exists y B \}}} & \rightarrow & \frac{\forall x (A(x) \wedge K \{ \exists y B \vee \exists y B \})}{r2\downarrow \frac{\forall x A(x) \wedge K \left\{ \frac{qc\downarrow \frac{\exists y B \vee \exists y B}{\exists y B}}{\exists y B} \right\}}{\forall x A(x) \wedge K \{ \exists y B \}}} \end{array}$$

Similarly, the reductions for permuting occurrences of  $qc\downarrow$  down through occurrences of  $r2\uparrow$  are also immediately verified, except in the case that the occurrence of  $qc\downarrow$  acts on the existential quantifier shifted by the occurrence of  $r2\uparrow$ . In this case, we apply the following reduction:

$$\frac{\frac{qc\downarrow \frac{\exists x A(x) \vee \exists x A(x)}{\exists x A(x)} \vee B}{r2\uparrow \frac{\exists x (A(x) \vee B)}}{\exists x (A(x) \vee B)} \rightarrow \frac{\frac{\frac{A(x)}{= \frac{A(x)}{A(x) \vee \frac{w\downarrow \frac{f}{B}}{B}}} \vee \frac{r2\uparrow \frac{\exists x A(x) \vee B}{\exists x (A(x) \vee B)}}}{qc\downarrow \frac{\exists x (A(x) \vee B)}}{\exists x (A(x) \vee B)}$$

where the occurrence of  $w\downarrow$  in the derivation above is replaced with a derivation in **SKS1** using Lemma 3.14 and sequential composition.

During Phase 2 (Lemma 5.5), when permuting occurrences of the existential witnessing rule  $\exists$  down the proof, they permute through occurrences of  $r2\downarrow$  by the following reductions:

$$\frac{\forall x \left( K \left\{ \frac{\exists \frac{A(t)}{\exists y A(y)}}{\exists y A(y)} \right\} \wedge B \right)}{r2\downarrow \frac{\forall x K \{ \exists y A \} \wedge B}} \rightarrow \frac{\forall x (K \{ A(t) \} \wedge B)}{r2\downarrow \frac{\forall x K \left\{ \frac{\exists \frac{A(t)}{\exists y A(y)}}{\exists y A(y)} \right\} \wedge B}{\forall x K \{ \exists y A \} \wedge B}}$$

$$\frac{\forall x \frac{A(x) \wedge K \left\{ \frac{\exists \frac{B(t)}{\exists y B(y)}}{\exists y B(y)} \right\}}{\forall x A(x) \wedge K \{ \exists y B \}}}{\text{r2}\downarrow} \rightarrow \frac{\forall x(A(x) \wedge K \{B(t)\})}{\text{r2}\downarrow} \frac{\forall x A(x) \wedge K \left\{ \frac{\exists \frac{B(t)}{\exists y B(y)}}{\exists y B(y)} \right\}}{\text{r2}\downarrow}$$

if  $x$  does not occur free in  $t$ , and

$$\begin{array}{c} \frac{\forall x \frac{A(x) \wedge K \left\{ \frac{\exists \frac{B(t)}{\exists y B(y)}}{\exists y B(y)} \right\}}{\forall x A(x) \wedge K \{ \exists y B(y) \}}}{\text{r2}\downarrow} \\ \downarrow \\ \frac{\forall x(A(x) \wedge K \{B(t)\})}{\text{qc}\uparrow} \frac{\forall x \frac{A(x) \wedge \frac{K \{B(t)\}}{t}}{t} \wedge \frac{\forall x \frac{A(x)}{t} \wedge K \{B(t)\}}{\forall \frac{t \wedge K \{B(t)[c/x]\}}{t \wedge K \{B(t)[c/x]\}}} \\ = \frac{\forall x A(x) \wedge K \left\{ \frac{\exists \frac{B(t)[c/x]}{\exists y B(y)}}{\exists y B(y)} \right\}}{\text{r2}\downarrow} \end{array}$$

if  $x$  occurs free in  $t$ , where the occurrences of  $w\uparrow$  in the derivation above are replaced with derivations in  $\text{SKS1}$  using Lemma 3.14 and sequential composition. Occurrences of  $\exists$  trivially permute down through occurrences of  $\text{r2}\uparrow$ , except in the case that the occurrence of  $\exists$  instantiates the existential quantifier which is shifted by the occurrence of  $\text{r2}\uparrow$ . In this case, we eliminate the occurrence of  $\text{r2}\uparrow$  by the following reduction:

$$\frac{\frac{\frac{\exists \frac{A(t)}{\exists x A(x)}}{\exists x A(x)} \vee B}{\text{r2}\uparrow} \rightarrow \frac{\exists \frac{A(t) \vee B}{\exists x(A(x) \vee B)}}{\text{r2}\uparrow}$$

In this manner, all occurrences of  $\text{r2}\uparrow$  are eliminated from the proof during the second phase of the procedure.

During Phase 3 (Lemma 5.8), when permuting occurrences of  $\text{qc}\uparrow$  up the proof, it is possible to eliminate every occurrence of  $\text{qc}\uparrow$  in the proof without decomposing every occurrence of  $\text{r2}\downarrow$ , so that a proof is obtained with an upper segment in  $\text{SKSg}\varepsilon \cup \{\text{r2}\downarrow\}$  instead of  $\text{SKSg}\varepsilon$ . Indeed, occurrences of  $\text{qc}\uparrow$  trivially permute up through occurrences of  $\text{r2}\downarrow$  except in the case that the occurrence of  $\text{qc}\uparrow$  acts on the universal quantifier that is shifted by the occurrence of  $\text{r2}\downarrow$ , in a subderivation of the form

$$\frac{\forall x(A(x) \wedge B)}{\text{r2}\downarrow} \frac{\frac{\forall x A(x)}{\text{qc}\uparrow} \wedge B}{\forall x A(x) \wedge \forall x A(x)}$$

in which case, the occurrence of  $\text{r2}\downarrow$  should be replaced with the relevant derivation given in the proof of Proposition 3.16 using sequential composition before permuting the newly introduced occurrence of  $\text{qc}\uparrow$  up the proof.



## 6 The Restricted Falsifier Calculus

Science is voiceless; it is the scientists who talk.

*S. Weil*

Section 4 introduced the falsifier calculus  $\text{SKSg}\varepsilon$  and used it to provide a decomposition theorem for proofs in first-order predicate logic, Theorem 4.13 the Falsifier Decomposition Theorem. The falsifier calculus  $\text{SKSg}\varepsilon$  includes the universal instantiation rule  $\forall$ , given by

$$\forall \frac{\forall x A(x)}{A(t)}$$

where  $t$  is free for  $x$  in  $A(x)$ , in which the term  $t$  that instantiates the rule may contain  $\varepsilon$ -terms. In this section, I will introduce a subsystem of the falsifier calculus  $\text{SKSg}\varepsilon$ , called the *restricted falsifier calculus*  $\text{SKSg}\varepsilon'$ , which restricts the universal instantiation rule  $\forall$  in the falsifier calculus  $\text{SKSg}\varepsilon$  so that  $t$  must be a variable. The restricted falsifier calculus  $\text{SKSg}\varepsilon'$  admits the same normalisation and complexity properties as the full falsifier calculus  $\text{SKSg}\varepsilon$ , demonstrating that only such trivial instances of the  $\forall$  rule are necessary to attain the desirable properties of proofs in falsifier normal form.

In Subsection 6.1 I explore the connection between the universal instantiation rule  $\forall$  and the cut rule  $i\uparrow$  and illustrate how the three-phase procedure given in Section 5 for transforming an epsilon-free proof in  $\text{SKSg}1$  into falsifier normal form acts on occurrences of  $i\uparrow$  in a proof to introduce occurrences of  $\forall$  into the proof. In Subsection 6.2 I introduce the restricted falsifier calculus  $\text{SKSg}\varepsilon'$  and use it to prove a restricted form of Theorem 4.13 the Falsifier Decomposition Theorem.

### 6.1 Universal Instantiation Rules

The falsifier calculus  $\text{SKSg}\varepsilon$  includes the universal instantiation rule  $\forall$ , given by

$$\forall \frac{\forall x A(x)}{A(t)}$$

where  $t$  is free for  $x$  in  $A(x)$ , in which the term  $t$  that instantiates the rule may contain  $\varepsilon$ -terms, raising the question of what role such non-trivial instances of the  $\forall$  rule that are instantiated by terms which contain  $\varepsilon$ -terms play in the complexity and normalisation properties of the falsifier calculus.

As demonstrated by Lemma 3.12 and discussed in Subsection 3.3, instances of the universal instantiation rule  $\forall$  play a fundamental role in decomposing instances of the cut rule  $i\uparrow$  into instances of smaller rules. In particular, the part of an instance of the cut rule  $i\uparrow$  which acts on quantifiers may be decomposed into an instance of the universal instantiation rule  $\forall$  and an instance of the quantifier-shift  $r1\uparrow$  such that the instance of  $\forall$  is instantiated by the variable which corresponds to the outermost existential quantifier in the instance of  $i\uparrow$ , as follows.

$$i\uparrow \frac{\forall x A(x) \wedge \exists x \bar{A}(x)}{f} \rightarrow \begin{array}{c} \boxed{\forall x A(x)} \\ = \frac{\boxed{\forall y A(y)}}{\forall y A(y)} \wedge \exists x \bar{A}(x) \\ \hline r1\uparrow \\ \boxed{\exists x \left( \frac{\boxed{\forall y A(y)}}{\forall y A(y)} \wedge \bar{A}(x) \right)} \\ \boxed{i\uparrow \frac{\boxed{\forall y A(y)}}{\forall y A(y)} \wedge \bar{A}(x)} \\ \boxed{f} \\ = \frac{\boxed{\boxed{\boxed{\boxed{\forall y A(y)}} \wedge \bar{A}(x)}}}{f} \end{array}$$

This is relevant to Theorem 4.13 the Falsifier Decomposition Theorem since the proof presented in Section 5 for transforming an epsilon-free proof in SKSg1 into falsifier normal form first decomposes occurrences of the cut rule  $i\uparrow$  in the proof using Lemma 5.1 (which in turn uses Lemma 3.12) so that occurrences of  $\forall$  in a proof in falsifier normal form may correspond to parts of decomposed occurrences of the cut rule  $i\uparrow$ .

I will now illustrate how the three-phase procedure for transforming an epsilon-free proof in SKSg1 into falsifier normal form presented in Section 5 affects the outermost quantifiers in an occurrence of  $i\uparrow$  on a quantified formula in the initial proof. Consider the following occurrence of  $i\uparrow$  contained in the proof prior to the procedure.

$$i\uparrow \frac{\forall x A(x) \wedge \exists x \bar{A}(x)}{f}$$

The occurrence is first decomposed using Lemma 5.1 so that the part of the occurrence which acts on the quantifiers is decomposed as follows.

$$\begin{array}{c} \boxed{\frac{\forall x A(x)}{\forall y A(y)} \wedge \exists x \bar{A}(x)} \\ r1\uparrow \frac{}{\boxed{\exists x \left( \boxed{\frac{\forall y A(y)}{A(x)} \wedge \bar{A}(x)} \right)}} \\ i\uparrow \frac{}{f} \\ = \frac{}{f} \end{array}$$

Observe that the highlighted occurrence of  $\forall$  in the derivation above is instantiated by the variable  $x$ , which corresponds to the outermost existential quantifier  $\exists x$  in the occurrence of the cut rule  $i\uparrow$ . An explicit disjunction of term witnesses  $t_1, \dots, t_n$  may be obtained for the existential quantifier  $\exists x$  by permuting occurrences of the existential contraction rule  $qc\downarrow$  and existential witnessing rule  $\exists$  down the proof (see Lemmas 5.3 and 5.5), with  $\varepsilon$ -terms generated as witnesses when permuting occurrences of  $\exists$  down through occurrences of the falsifier rule  $\varepsilon$  which replace occurrences of the quantifier-shift  $r1\downarrow$ , yielding the following derivation.

$$\begin{array}{c} \boxed{\frac{\forall x A(x)}{\forall y A(y)} \wedge \left( \begin{array}{c} \bar{A}(t_1) \vee \dots \vee \bar{A}(t_n) \\ \parallel \{\exists, qc\downarrow\} \\ \exists x \bar{A}(x) \end{array} \right)} \\ r1\uparrow \frac{}{\boxed{\exists x \left( \boxed{\frac{\forall y A(y)}{A(x)} \wedge \bar{A}(x)} \right)}} \\ i\uparrow \frac{}{f} \\ = \frac{}{f} \end{array}$$

By permuting the occurrences of  $qc\downarrow$  down the derivation above (as in the procedure given

by Lemma 5.3), the following derivation is obtained.

$$\begin{array}{c}
\boxed{\begin{array}{c} \boxed{\frac{\forall x A(x)}{\forall y A(y)}} \\ \parallel \{\text{qc}\uparrow\} \\ \forall y A(y) \wedge \cdots \wedge \forall y A(y) \end{array}} \wedge \boxed{\begin{array}{c} \overline{A}(t_1) \vee \cdots \vee \overline{A}(t_n) \\ \parallel \{\exists\} \\ \exists x \overline{A}(x) \vee \cdots \vee \exists x \overline{A}(x) \end{array}} \\
\parallel \{\text{s}\} \\
\boxed{\begin{array}{c} \text{r1}\uparrow \frac{\forall y A(y) \wedge \exists x \overline{A}(x)}{\vdots} \\ \vdots \\ \exists x \boxed{\begin{array}{c} \boxed{\frac{\forall y A(y)}{\forall A(x)}} \wedge \overline{A}(x) \\ \text{i}\uparrow \\ \text{f} \end{array}} \vee \cdots \vee \boxed{\begin{array}{c} \boxed{\frac{\forall y A(y)}{\forall A(x)}} \wedge \overline{A}(x) \\ \text{i}\uparrow \\ \text{f} \end{array}} \\
= \text{f} \end{array}}
\end{array}$$

Observe that the resultant derivation contains  $n$  occurrences of the universal instantiation rule  $\forall$ , which are instantiated by the variables corresponding to  $n$  copies of the existential quantifier  $\exists x$ . By permuting the occurrences of  $\exists$  down the derivation above (as in the procedure given by Lemma 5.5), an explicit witness is assigned to each existential quantifier, resulting in the following derivation.

$$\begin{array}{c}
\boxed{\begin{array}{c} \boxed{\frac{\forall x A(x)}{\forall y A(y)}} \\ \parallel \{\text{qc}\uparrow\} \\ \forall y A(y) \wedge \cdots \wedge \forall y A(y) \end{array}} \wedge (\overline{A}(t_1) \vee \cdots \vee \overline{A}(t_n)) \\
\parallel \{\text{s}\} \\
\boxed{\begin{array}{c} \boxed{\frac{\forall y A(y)}{\forall A(t_1)}} \wedge \overline{A}(t_1) \\ \text{i}\uparrow \\ \text{f} \end{array}} \vee \cdots \vee \boxed{\begin{array}{c} \boxed{\frac{\forall y A(y)}{\forall A(t_n)}} \wedge \overline{A}(t_n) \\ \text{i}\uparrow \\ \text{f} \end{array}}
\end{array}$$

Observe that the occurrences of the universal instantiation rule  $\forall$  in the resultant derivation are instantiated by the terms  $t_1, \dots, t_n$ , which may contain  $\varepsilon$ -terms. Finally, by permuting the occurrences of  $\text{qc}\uparrow$  up the proof (as in the procedure given by Lemma 5.8), the following derivation is obtained.

$$\begin{array}{c}
\boxed{\frac{\forall x A(x)}{\forall y A(y)}} \wedge \cdots \wedge \boxed{\frac{\forall x A(x)}{\forall y A(y)}} \wedge (\overline{A}(t_1) \vee \cdots \vee \overline{A}(t_n)) \\
\parallel \{\text{s}\} \\
\boxed{\begin{array}{c} \boxed{\frac{\forall y A(y)}{\forall A(t_1)}} \wedge \overline{A}(t_1) \\ \text{i}\uparrow \\ \text{f} \end{array}} \vee \cdots \vee \boxed{\begin{array}{c} \boxed{\frac{\forall y A(y)}{\forall A(t_n)}} \wedge \overline{A}(t_n) \\ \text{i}\uparrow \\ \text{f} \end{array}}
\end{array}$$

The derivation above is therefore the resultant subderivation yielded by applying the three-phase procedure corresponding to the Falsifier Decomposition Theorem to the outermost quantifiers in an instance of the cut rule  $i\uparrow$  on a quantified formula  $\forall xA(x)$ . Since the derivation contains occurrences of the universal instantiation rule  $\forall$  instantiated by terms  $t_1, \dots, t_n$ , it follows that occurrences of  $\forall$  in proofs in falsifier normal form may correspond to residual parts of decomposed occurrences of the cut rule  $i\uparrow$ . As such, the universal instantiation rule  $\forall$  is intimately connected to the cut rule  $i\uparrow$  and cut elimination in the falsifier calculus and in proofs in falsifier normal form. In particular, occurrences of  $\forall$  which are instantiated by terms which contain  $\varepsilon$ -terms in a proof resulting from the three-phase procedure corresponding to the proof of Theorem 4.13 the Falsifier Decomposition Theorem must have resulted from an occurrence of  $\forall$  which was contained in an existential quantifier that was witnessed by an  $\varepsilon$ -term, as is the case in decomposing certain occurrences of  $i\uparrow$ .

## 6.2 A Restricted Form of The Falsifier Decomposition Theorem

The above discussion demonstrates that non-trivial occurrences of the universal instantiation rule  $\forall$  that are instantiated by terms which contain  $\varepsilon$ -terms are intimately connected to the cut rule  $i\uparrow$  in proofs in falsifier normal form. Since the cut rule is admissible in proof systems for first-order predicate logic, it is natural then to consider whether such non-trivial instances of  $\forall$  are necessary to attain completeness of the open-deduction proof system with rules  $\text{SKSg}\varepsilon \cup \{\exists, \text{qc}\downarrow\}$  given by Corollary 4.15. I will now demonstrate that, in fact, only trivial instances of  $\forall$  that are instantiated by variables are necessary to attain this completeness and that an analogue of Theorem 4.13 the Falsifier Decomposition Theorem holds using only trivial instances of  $\forall$ , demonstrating that only trivial instances of  $\forall$  are required for the desirable properties of the falsifier calculus.

The *trivial universal instantiation rule*  $\forall_0$  is given as follows.

**Definition 6.1.** The *trivial universal instantiation rule*  $\forall_0$  is given by

$$\forall_0 \frac{\forall xA(x)}{A(y)}$$

where  $y$  is a variable that is free for  $x$  in  $A(x)$ .

I define the *restricted falsifier calculus*  $\text{SKSg}\varepsilon'$  to be the subsystem of the full falsifier calculus  $\text{SKSg}\varepsilon$  which contains the rule  $\forall_0$  in place of  $\forall$ , as follows.

**Definition 6.2.** The *restricted falsifier calculus* is given by  $\text{SKSg}\varepsilon' = (\text{SKSg}\varepsilon \setminus \{\forall\}) \cup \{\forall_0\}$ .

I now state an analogue of Theorem 4.13 the Falsifier Decomposition Theorem for the restricted falsifier calculus  $\text{SKSg}\varepsilon'$ , which decomposes an epsilon-free proof in  $\text{SKSg}1 \setminus \{\forall, i\uparrow\}$  into an upper segment in the cut-free restricted falsifier calculus  $\text{SKSg}\varepsilon' \setminus \{i\uparrow\}$  and a lower segment in  $\{\exists, \text{qc}\downarrow\}$ . The completeness of the open-deduction proof system with rules  $\text{SKSg}1 \setminus \{\forall, i\uparrow\}$  follows from Lemma 3.8 so that this decomposition theorem provides a factorisation of proofs for all valid formulae in epsilon-free first-order predicate logic.

**Theorem 6.3** (The Falsifier Decomposition Theorem for  $\text{SKSg}\varepsilon'$ ). *For every epsilon-free proof  $\phi$  with conclusion  $A$  in  $\text{SKSg}1 \setminus \{\forall, i\uparrow\}$ , there exists an epsilon-closed proof of the form*

$$\begin{array}{c} \phi' \parallel \text{SKSg}\varepsilon' \setminus \{i\uparrow\} \\ A' \\ \parallel \{\exists\} \\ A'' \\ \parallel \{\text{qc}\downarrow\} \\ A \end{array} \quad (15)$$

such that the following elementary bounds hold

$$\begin{aligned} |\phi'| &= \exp^8(O(|\phi|^2 \ln |\phi|)) \\ |A'| &= \exp^5(O(|\phi|^2 \ln |\phi|)) \\ |\phi'|_\varepsilon &= \exp^{10}(O(|\phi|^2 \ln |\phi|)) \\ |A'|_\varepsilon &= \exp^{10}(O(|\phi|^2 \ln |\phi|)) \end{aligned}$$

Furthermore, every occurrence of  $\varepsilon$  in  $\phi'$  occurs within a subderivation of the form

$$\boxed{\begin{array}{c} \frac{\varepsilon \quad \forall x(B(x) \vee C(x))}{\forall x B(x) \vee C(\varepsilon_y \overline{B}(y))} \wedge \dots \wedge \frac{\varepsilon \quad \forall x(B(x) \vee C(x))}{\forall x B(x) \vee C(\varepsilon_y \overline{B}(y))} \\ \parallel \{s\} \\ \boxed{(\forall x B(x) \wedge \dots \wedge \forall x B(x)) \vee \begin{array}{c} C(\varepsilon_y \overline{B}(y)) \vee \dots \vee C(\varepsilon_y \overline{B}(y)) \\ \parallel \{c\downarrow, m, \forall_0, =_P, =_V\} \\ C(\varepsilon_y \overline{B}(y)) \end{array}} \end{array}}$$

where every occurrence of  $\varepsilon$  in  $\phi'$  with variable name  $y$  is contained in the subderivation above.

*Proof.* This may be proved analogously to Theorem 4.13, by applying the three-phase procedure presented in Section 5 to  $\phi$  and observing that the relevant lemmas and propositions only introduce occurrences of the  $\forall$  rule of the form  $\forall_0$  into the proof, by the constructions presented in their proofs.  $\square$

Observe that the statement and proof of the theorem given above require that the proof prior to performing the procedure be cut-free, since a proof which contains occurrences of the cut rule  $i\uparrow$  may introduce non-trivial occurrences of  $\forall$  into the proof during the procedure. As such, it is possible to transform proofs which are cut-free and contain no occurrences of  $\forall$  into the restricted falsifier calculus with elementary complexity, but not proofs which contain occurrences of the cut rule  $i\uparrow$ .

I define the notions of *restricted falsifier normal form* and *restricted falsifier disjunctions*, which are the respective analogues to falsifier normal form and falsifier disjunctions induced by Theorem 6.3.

**Definition 6.4.** The normal form for proofs given by (15) is called *restricted falsifier normal form* and the formula  $A'$  is called a *restricted falsifier disjunction* for  $A$ .

I now note some of the consequences of Theorem 6.3. The following corollaries are analogues to Corollaries 4.15 and 4.16 of Theorem 4.13 the Falsifier Decomposition Theorem and may be proved from Theorem 6.3 by the same means.

It follows from Theorem 6.3 that the open-deduction proof system with rules  $(\text{SKSg}\varepsilon' \setminus \{i\uparrow\}) \cup \{\exists, \text{qc}\downarrow\}$  forms a complete proof system for epsilon-free first-order predicate logic and admits non-elementarily smaller cut-free proofs than the sequent calculus **LK** for certain classes of theorems, as follows.

**Corollary 6.5.** Every valid epsilon-free formula has a proof in  $(\text{SKSg}\varepsilon' \setminus \{i\uparrow\}) \cup \{\exists, \text{qc}\downarrow\}$  and there is no elementary function bounding the size of the smallest cut-free **LK** proof of a formula in terms of the size or  $\varepsilon$ -size of its smallest proof in  $(\text{SKSg}\varepsilon' \setminus \{i\uparrow\}) \cup \{\exists, \text{qc}\downarrow\}$ .

It also follows from Theorem 6.3 that there exist classes of first-order theorems which admit non-elementarily smaller restricted falsifier disjunctions than Herbrand disjunctions, as follows.

**Corollary 6.6.** *There is no elementary function bounding the size of the smallest Herbrand disjunction of a valid epsilon-free formula in terms of the size or  $\varepsilon$ -size of its smallest restricted falsifier disjunction.*

It follows that the restricted falsifier calculus  $\text{SKSg}\varepsilon'$  shares the desirable properties of the full falsifier calculus  $\text{SKSg}\varepsilon$ , namely in its normalisation properties and speedups over traditional proof systems. There is a sense in which the cut-free restricted falsifier calculus  $\text{SKSg}\varepsilon' \setminus \{\text{i}\uparrow\}$  can be understood as the “true” cut-free fragment of the falsifier calculus  $\text{SKSg}\varepsilon$ , due to the relationship between non-trivial instances of the universal instantiation rule  $\forall$  and the cut rule  $\text{i}\uparrow$  discussed in the previous subsection, Subsection 6.1. As was discussed prior to the definition of a cut-free derivation, Definition 3.7, there are several ways that cut-freeness may be defined in deep inference due to the greater flexibility of deep-inference formalisms over conventional proof formalisms. Corollaries 4.15 and 6.5 demonstrate that the falsifier calculus  $\text{SKSg}\varepsilon$  admits non-elementarily smaller cut-free proofs of certain classes of first-order theorems than the sequent calculus for two different notions of “cut-free”, showing that this compression is attained regardless of the chosen definition of “cut-free”.

The restricted falsifier calculus  $\text{SKSg}\varepsilon'$  also has certain normalisation benefits over the falsifier calculus  $\text{SKSg}\varepsilon$ . In an epsilon-closed proof in  $\text{SKSg}\varepsilon'$ , such as those in the upper segment of a proof in restricted falsifier normal form resulting from the three-phase procedure presented, every  $\varepsilon$ -term which occurs in the proof is constructed by some occurrence of the falsifier rule  $\varepsilon$  in the proof. It follows that every  $\varepsilon$ -term which occurs in the proof is associated with a particular universal quantifier, unlike in epsilon-closed proofs in the falsifier calculus  $\text{SKSg}\varepsilon$ , in which non-trivial occurrences of the universal instantiation rule  $\forall$  can result in  $\varepsilon$ -terms being associated with multiple universal quantifiers in this manner. This property of epsilon-closed proofs in restricted falsifier normal form makes proofs more amenable to further normalisation and it is expected that this will provide a simpler means of deriving Herbrand disjunctions than for proofs in falsifier normal form.

Like falsifier normal form, restricted falsifier normal form provides a normal form for proofs of formulae in epsilon-free first-order predicate logic using the falsifier rule  $\varepsilon$ , for which the proofs of certain classes of theorems are non-elementarily smaller than the corresponding smallest proofs in Herbrand normal form. This is due to the presence of occurrences of the falsifier rule  $\varepsilon$  which replace occurrences of the quantifier-shift  $\text{r1}\downarrow$  in the original proof in the procedures corresponding to both normal forms, so that the non-elementary proof compression yielded by quantifier-shifts is observed in both normal forms. However, this does not provide insight as to whether there are further differences in complexity between falsifier normal form and restricted falsifier normal form, resulting from the presence of non-trivial occurrences of the universal instantiation rule  $\forall$  which may be present in proofs in falsifier normal form but not in proofs in restricted falsifier normal form. The following example demonstrates that such occurrences of  $\forall$  are indeed a source of complexity difference between the two normal forms for some formulae.

**Example 6.7.** The following is a proof of the Skolemised form of the drinker’s paradox in falsifier normal form.

$$\begin{array}{c}
 = \frac{t}{\forall y \left[ \frac{t}{P(y) \vee \overline{P}(y)} \right]} \\
 \varepsilon \frac{\boxed{\frac{\forall y P(y)}{P(f(\varepsilon_z \overline{P}(z)))}} \vee \overline{P}(\varepsilon_z \overline{P}(z))}{\exists x (P(f(x)) \vee \overline{P}(x))}
 \end{array}$$

In this example, the falsifier disjunction for the formula  $\exists x (P(f(x)) \vee \overline{P}(x))$  is  $P(f(\varepsilon_z \overline{P}(z))) \vee \overline{P}(\varepsilon_z \overline{P}(z))$ . Observe that the highlighted occurrence of  $\forall$  is instantiated by the term

$f(\varepsilon_z \bar{P}(z))$ , which contains an  $\varepsilon$ -term, and hence is not an occurrence of  $\forall_0$ . The above proof is therefore in falsifier normal form, but not in restricted falsifier normal form.

Indeed, the smallest proof in restricted falsifier normal form for this formula is the following.

$$\begin{array}{c}
\boxed{\text{t}} \\
\text{i}\downarrow \frac{\quad}{P(f(c)) \vee \bar{P}(f(c))} \\
= \\
\boxed{\boxed{\boxed{P(f(c)) \vee \boxed{\text{f}} \downarrow \bar{P}(c)}} \vee \boxed{\boxed{\text{f}} \downarrow \boxed{P(f(f(c)))} \vee \bar{P}(f(c))}} \\
\exists \downarrow \frac{\quad}{\exists x(P(f(x)) \vee \bar{P}(x))} \quad \vee \quad \exists \downarrow \frac{\quad}{\exists x(P(f(x)) \vee \bar{P}(x))} \\
\text{qc}\downarrow \frac{\quad}{\exists x(P(f(x)) \vee \bar{P}(x))}
\end{array}$$

The restricted falsifier disjunction for  $\exists x(P(f(x)) \vee \bar{P}(x))$  given by this proof is  $(P(f(c)) \vee \bar{P}(c)) \vee (P(f(f(c))) \vee \bar{P}(f(c)))$ , which is also the smallest Herbrand disjunction for the formula. Indeed, the above proof is also the smallest proof in Herbrand normal form for the formula.

It follows that there is a difference in complexity between proofs in falsifier normal form and proofs in restricted falsifier normal form. Given the relationship between non-trivial instances of the universal instantiation rule  $\forall$  and the cut rule  $\text{i}\uparrow$ , since in general there is a non-elementary difference in complexity between proof systems with and without the cut rule, it is possible that the difference in complexity between the two normal forms is significant. A complexity comparison between the two normal forms and their corresponding proof systems, the falsifier calculus  $\text{SKSg}\varepsilon$  and the restricted falsifier calculus  $\text{SKSg}\varepsilon'$ , has fallen outside the scope of this thesis. This does however provide an interesting research question which could perhaps also deepen our understanding of the role of non-trivial occurrences of the universal instantiation rule  $\forall$  in proofs in falsifier normal form and in first-order proof theory more broadly.

As discussed, deep-inference proof systems provide a means of extracting proofs in Herbrand normal form from arbitrary first-order proofs without eliminating all occurrences of the cut rule from the proof [21]. Procedures for extracting Herbrand proofs from first-order sequent calculus proofs usually require all cut rules to be eliminated from the proof, meaning that deep inference reveals an independence between Herbrand's Theorem and cut elimination which traditional proof formalisms cannot detect. Since non-trivial occurrences of  $\forall$  are closely related to occurrences of the cut rule on quantified formulae, developing a procedure for extracting proofs in restricted falsifier normal form from proofs in falsifier normal form could also provide a new perspective on the relationship between Herbrand's Theorem and cut elimination, suggesting that an analysis of the relationship between the falsifier calculus and the restricted falsifier calculus may be a fruitful direction for future research.

## 7 Conclusion

On and on we go, for the mental consciousness labours  
under the illusion that there is somewhere to go to, a goal  
to consciousness. Whereas of course there is no goal.  
Consciousness is an end in itself. We torture ourselves  
getting somewhere, and when we get there it is nowhere,  
for there is nowhere to get to.

---

*D. H. Lawrence*

This thesis introduced the falsifier calculus, a deep-inference proof system for first-order predicate logic in the language of Hilbert's epsilon-calculus, and used it to provide a new decomposition theorem for first-order proofs, Theorem 4.13 the Falsifier Decomposition Theorem. This result provides a new perspective on the structure and complexity of proofs in first-order predicate logic as well as a novel insight into the phenomenon of the non-elementary compression of cut-free proofs. The notion of falsifier disjunctions that is induced by the Falsifier Decomposition Theorem provides a new perspective on the structure and complexity of Herbrand disjunctions, giving us a new way to look at one of the central theorems of classical proof theory.

The falsifier calculus is related to several core concepts in first-order proof theory and other areas of research, such as Herbrand's Theorem, the non-elementary compression of cut-free proofs, quantifier-shifts, the epsilon-calculus and other proof interpretations in the literature, such as the no-counterexample interpretation. These topics are all seemingly related to the same fundamental phenomena in first-order proof theory, but the precise connections between them can be unclear. Providing a syntactic formalism connected to these areas can serve as a bridge between them, revealing new insights into the relationships between them and into the nature of first-order proofs more generally.

The most obvious next step for this research is to develop a procedure for extracting proofs in Herbrand normal form from proofs in falsifier normal form and restricted falsifier normal form, providing a novel proof of Herbrand's Theorem using the falsifier calculus. Such a procedure would eliminate the  $\varepsilon$ -terms from proofs in the falsifier calculus, resembling the procedure of epsilon substitution used for the traditional epsilon-calculus. It is expected that the disjunctive composition mechanism of open deduction will provide a simpler means of expressing the case distinctions implicitly contained in  $\varepsilon$ -terms than is available in traditional proof formalisms, providing a simpler and more semantically natural means of eliminating  $\varepsilon$ -terms from proofs. Traditional epsilon substitution is notoriously convoluted in Hilbert-Frege systems and the sequent calculus, suggesting that the falsifier calculus may be able to provide further simplification in this area.

Furthermore, it is possible that the falsifier calculus and the falsifier rule may be able to inform the design of a more compact syntax for the epsilon-calculus. Some of the results in this thesis suggest that not all of the syntactic information contained in  $\varepsilon$ -terms is required to attain completeness of the falsifier calculus or its desirable complexity and normalisation properties. In particular, in proofs in restricted falsifier norm, every  $\varepsilon$ -term which occurs in a proof is associated with a particular universal quantifier in the proof, suggesting that it is this information which is required for completeness. Several alternate syntaxes for the epsilon-calculus have already been developed in the literature [14, 36, 38], suggesting that the falsifier calculus may indeed be able to provide one as well.

The falsifier calculus and the falsifier rule may also be able to inform our understanding of how to extend non-classical logics by the epsilon operator. The problem of incorporating the epsilon operator into intuitionistic logic has long been of interest, with the first conservative extension of intuitionistic predicate logic by the epsilon operator and a corresponding Kripke semantics being proposed in Meyer Viol's 1995 thesis [70]. Extensions of various intermediate logics by the epsilon operator and epsilon theorems for these logics have recently been studied by Baaz and Zach [18] and the role of the quantifier-shifts in intermediate logics has also



recently been studied by Baaz, Gamsakhurdia, Iemhoff and Jalali [12]. Existing studies into extensions of intuitionistic logic by the epsilon operator usually base the extension around encoding existential quantifiers by  $\varepsilon$ -terms or modelling critical axioms in intuitionistic logic. Perhaps the falsifier rule may be able to provide another means of defining such extensions and it would be interesting to study the logic obtained by extending intuitionistic predicate logic by the falsifier rule.

## References

- [1] B. Afshari, S. Hetzl, and G. E. Leigh. Herbrand’s theorem as higher order recursion. *Annals of Pure and Applied Logic*, 171(6):102792, 2020.
- [2] J. P. Aguilera and M. Baaz. Unsound Inferences Make Proofs Shorter. *The Journal of Symbolic Logic*, 84(1):102–122, 2019.
- [3] A. Alcolei, P. Clairambault, M. Hyland, and G. Winskel. The true concurrency of herbrand’s theorem. In D. R. Ghica and A. Jung, editors, *27th EACSL Annual Conference on Computer Science Logic (CSL 2018)*, Leibniz International Proceedings in Informatics, LIPIcs, pages 5:1–5:22. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, GBR, August 2018.
- [4] A. Aler Tubella. *A Study of Normalisation Through Subatomic Logic*. PhD thesis, University of Bath, 2017.
- [5] A. Aler Tubella and A. Guglielmi. Subatomic Proof Systems: Splittable Systems. *ACM Transactions on Computational Logic*, 19(1):5:1–33, 2017.
- [6] C. Allett. Non-elementary compression of first-order proofs in deep inference using epsilon-terms. In *Proceedings of the 39th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS ’24*. Association for Computing Machinery, 2024.
- [7] G. A. Antonelli. Numerical abstraction via the frege quantifier. *Notre Dame Journal of Formal Logic*, 51(2):161–179, 2010.
- [8] T. Arai. Epsilon substitution method for id1. *Annals of Pure and Applied Logic*, 121(2-3):163–208, 2003.
- [9] F. Aschieri, S. Hetzl, and D. Weller. Expansion trees with cut. *Mathematical Structures in Computer Science*, 29(8):1009–1029, 2019.
- [10] A. Avron. The method of hypersequents in the proof theory of propositional non-classical logics. In *Logic: from Foundations to Applications: European logic colloquium*. Oxford University Press, 06 1996.
- [11] M. Baaz, A. Ciabattoni, and C. G. Fermüller. Hypersequent calculi for gödel logics — a survey. *Journal of Logic and Computation*, 13(6):835–861, 12 2003.
- [12] M. Baaz, M. Gamsakhurdia, R. Iemhoff, and R. Jalali. Skolemization in intermediate logics, 2025.
- [13] M. Baaz and A. Leitsch. On Skolemization And Proof Complexity. *Fundam. Inf.*, 20(4):353–379, dec 1994.
- [14] M. Baaz, A. Leitsch, and A. Lolic. A Sequent-Calculus Based Formulation of the Extended First Epsilon Theorem. In S. Artemov and A. Nerode, editors, *Logical Foundations of Computer Science*, pages 55–71, Cham, 2018. Springer International Publishing.
- [15] M. Baaz and A. Lolic. A Globally Sound Analytic Calculus for Henkin Quantifiers. *Logical Foundations of Computer Science: International Symposium, LFCS 2020*, 2020.
- [16] M. Baaz and A. Lolic. Towards a proof theory for henkin quantifiers. *Journal of Logic and Computation*, 31(1):40–66, 01 2021.
- [17] M. Baaz and A. Lolic. Epsilon calculus provides shorter cut-free proofs, 2024.
- [18] M. Baaz and R. Zach. Epsilon theorems in intermediate logics. *Journal of Symbolic Logic*, 87(2):682–720, 2022.

- [19] C. Barrett and A. Guglielmi. A subatomic proof system for decision trees. *ACM Trans. Comput. Logic*, 23(4), 2022.
- [20] K. Brünnler. *Deep Inference and Symmetry in Classical Proofs*. Logos Verlag, Berlin, 2004.
- [21] K. Brünnler. Cut Elimination inside a Deep Inference System for Classical Predicate Logic. *Studia Logica*, 82(1):51–71, 2006.
- [22] K. Brünnler and A. Guglielmi. A first order system with finite choice of premises. In V. Hendricks, F. Neuhaus, S. A. Pedersen, U. Scheffler, and H. Wansing, editors, *First-Order Logic Revisited*, pages 59–74. Logos Verlag, 2004.
- [23] P. Bruscoli and A. Guglielmi. On analyticity in deep inference, 2009.
- [24] P. Bruscoli and A. Guglielmi. On the proof complexity of deep inference. *ACM Transactions on Computational Logic*, 10(2):14:1–34, 2009.
- [25] P. Bruscoli, A. Guglielmi, T. Gundersen, and M. Parigot. A Quasipolynomial Cut-Elimination Procedure in Deep Inference Via Atomic Flows and Threshold Formulae. In E. M. Clarke and A. Voronkov, editors, *Logic for Programming, Artificial Intelligence, and Reasoning (LPAR-16)*, volume 6355 of *Lecture Notes in Computer Science*, pages 136–153. Springer, 2010.
- [26] K. Brünnler. Deep inference and its normal form of derivations. In A. Beckmann, U. Berger, B. Löwe, and J. V. Tucker, editors, *Logical Approaches to Computational Barriers—2nd Conference on Computability in Europe*, volume 3988 of *Lecture Notes in Computer Science*, pages 65–74. Springer, 2006.
- [27] K. Brünnler. Deep sequent systems for modal logic. *Archive for Mathematical Logic*, 48(6):551–577, 2009.
- [28] K. Brünnler and A. Guglielmi. A first order system with finite choice of premises. In V. Hendricks, F. Neuhaus, S. A. Pedersen, U. Scheffler, and H. Wansing, editors, *First-Order Logic Revisited*, Logische Philosophie, pages 59–74. Logos Verlag, 2004.
- [29] K. Brünnler and A. F. Tiu. A local system for classical logic. In R. Nieuwenhuis and A. Voronkov, editors, *Logic for Programming, Artificial Intelligence, and Reasoning (LPAR)*, volume 2250 of *Lecture Notes in Computer Science*, pages 347–361. Springer, 2001.
- [30] S. R. Buss. On herbrand’s theorem. In D. Leivant, editor, *Logic and Computational Complexity*, pages 195–209, Berlin, Heidelberg, 1995. Springer Berlin Heidelberg.
- [31] S. R. Buss. An Introduction to Proof Theory. *Bulletin of Symbolic Logic*, 6(4):464–465, 2000.
- [32] S. Cook and R. Reckhow. On the lengths of proofs in the propositional calculus (preliminary version). In *Proceedings of the Sixth Annual ACM Symposium on Theory of Computing*, STOC ’74, page 135–148, New York, NY, USA, 1974. Association for Computing Machinery.
- [33] T. Coquand. A semantics of evidence for classical arithmetic. *Journal of Symbolic Logic*, 60(1):325–337, 1995.
- [34] A. Das. *The Complexity of Propositional Proofs in Deep Inference*. PhD thesis, University of Bath, 2014.

- [35] A. Das. On the pigeonhole and related principles in deep inference and monotone systems. In *Proceedings of the Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, CSL-LICS '14, New York, NY, USA, 2014. Association for Computing Machinery.
- [36] M. Davis and R. Fechter. A free variable version of the first-order predicate calculus. *Journal of Logic and Computation*, 1(4):431–451, 09 1991.
- [37] B. Dreben and P. Andrews. False lemmas in herbrand. *Journal of Symbolic Logic*, 31(4):657–659, 1966.
- [38] R. Fechter. *Skolem functions, Hilbert's epsilon-symbol, and the lambda calculus*. PhD thesis, USA, 1992. UMI Order No. GAX92-13227.
- [39] G. Frege. Begriffsschrift. In J. V. Heijenoort, editor, *From Frege to Gödel*, pages 1–83. Harvard University Press, 1967.
- [40] G. Gentzen. Die widerspruchsfreiheit der reinen zahlentheorie. *Journal of Symbolic Logic*, 1(2):75–75, 1936.
- [41] G. Gentzen. Investigations into Logical Deduction. *American Philosophical Quarterly*, 1(4):288–306, 1964.
- [42] P. Gerhardy and U. Kohlenbach. Extracting herbrand disjunctions by functional interpretation. *Arch. Math. Log.*, 44(5):633–644, 2005.
- [43] J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50(1):1–101, 1987.
- [44] K. Gödel. *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*. Basic Books, New York, NY, USA, 1931.
- [45] A. Guglielmi. A System of Interaction and Structure. *ACM Transactions on Computational Logic*, 8(1):1:1–64, 2007.
- [46] A. Guglielmi and T. Gundersen. Normalisation control in deep inference via atomic flows. *Logical Methods in Computer Science*, 4(1):9:1–36, 2008.
- [47] A. Guglielmi, T. Gundersen, and M. Parigot. A Proof Calculus Which Reduces Syntactic Bureaucracy. In C. Lynch, editor, *21st International Conference on Rewriting Techniques and Applications (RTA)*, volume 6 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 135–150. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2010.
- [48] T. Gundersen. *A General View of Normalisation Through Atomic Flows*. PhD thesis, University of Bath, 2009.
- [49] K. Gödel, S. Feferman, J. Dawson, John W, S. C. Kleene, G. H. Moore, R. M. Solovay, and J. v. Heijenoort. Über die vollständigkeit des logikkalküls (1929). In *Collected Works: Volume I Publications 1929-1936*. Oxford University Press, 1990.
- [50] K. Gödel, S. Feferman, J. Dawson, John W, S. C. Kleene, G. H. Moore, R. M. Solovay, and J. v. Heijenoort. Über eine bisher noch nicht benützte erweiterung des finiten standpunktes (1958). In *Collected Works: Volume II Publications 1938-1974*. Oxford University Press, 06 2001.
- [51] W. Heijltjes. Classical proof forestry. *Annals of Pure and Applied Logic*, 161(11):1346–1366, 2010. Special Issue: Classical Logic and Computation (2008).

- [52] W. Heijltjes. *Graphical Representation of Canonical Proof: Two Case Studies*. PhD thesis, University of Edinburgh, 2012.
- [53] L. Henkin. Some Remarks on Infinitely Long Formulas. In *Journal of Symbolic Logic*, pages 167–183. Pergamon Press, 1961.
- [54] J. Herbrand. *Recherches sur la théorie de la démonstration*. PhD thesis, University of Paris, 1930.
- [55] H. Herre, M. Krynicki, A. Pinus, and J. Vä”anänen. The härting quantifier: A survey. *Journal of Symbolic Logic*, 56(4):1153–1183, 1991.
- [56] S. Hetzl. A simplified proof of the epsilon theorems. *The Review of Symbolic Logic*, 17(4):1248–1263, 2024.
- [57] D. Hilbert. *The Foundations of Mathematics*. 1927.
- [58] J. Hintikka. *The Principles of Mathematics Revisited*. Cambridge University Press, 1996.
- [59] Z. Hóu, A. Tiu, and R. Goré. A labelled sequent calculus for bbi: Proof theory and proof search. In D. Galmiche and D. Larchey-Wendling, editors, *Automated Reasoning with Analytic Tableaux and Related Methods*, pages 172–187, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- [60] D. J. D. Hughes. Proofs without syntax. *Annals of Mathematics*, 164(3):1065–1076, 2006.
- [61] D. J. D. Hughes. First-order proofs without syntax, 2019.
- [62] D. J. D. Hughes, L. Straßburger, and J.-H. Wu. Combinatorial Proofs and Decomposition Theorems for First-order Logic. *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–13, 2021.
- [63] E. Jeřábek. Proof complexity of the cut-free calculus of structures. *Journal of Logic and Computation*, 19(2):323–339, 2009.
- [64] U. Kohlenbach. On the no-counterexample interpretation. *The Journal of Symbolic Logic*, 64(4):1491–1511, 1999.
- [65] G. Kreisel. On the interpretation of non-finitist proofs—part i. *The Journal of Symbolic Logic*, 16(4):241–267, 1951.
- [66] G. Kreisel. On the interpretation of non-finitist proofs—part ii. *Journal of Symbolic Logic*, 17(1):43–58, 1952.
- [67] T. Lyon, C. Ittner, T. Eckhardt, and N. Gratzl. The basics of display calculi. *Kriterion (Austria)*, 31:55–100, 09 2022.
- [68] R. McKinley. A sequent calculus demonstration of herbrand’s theorem, 2010.
- [69] R. McKinley. Proof nets for herbrand’s theorem. *ACM Trans. Comput. Logic*, 14(1), feb 2013.
- [70] W. Meyer-Viol. *Instantial Logic*. PhD thesis, 1995.
- [71] D. A. Miller. A compact representation of proofs. *Studia Logica: An International Journal for Symbolic Logic*, 46(4):347–370, 1987.
- [72] G. Mints. Epsilon substitution for first- and second-order predicate logic. *Annals of Pure and Applied Logic*, 164(6):733–739, 2013. Classical Logic and Computation 2010(CLAC 2010).

- [73] K. Miyamoto and G. Moser. Herbrand complexity and the epsilon calculus with equality. *Archive for Mathematical Logic*, 2023.
- [74] G. Moser and R. Zach. The Epsilon Calculus. In *Computer Science Logic*, pages 455–455, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.
- [75] G. Moser and R. Zach. The epsilon calculus and herbrand complexity. *Studia Logica*, 82(1):133–155, 2006.
- [76] D. Prawitz. *Natural Deduction: A Proof-Theoretical Study*. Dover Publications, Mineola, N.Y., 1965.
- [77] B. Ralph. Decomposing First-Order Proofs using Deep Inference. *Proof, Computation, Complexity 2016*, 2016.
- [78] B. Ralph. *Modular Normalisation of Classical Proofs*. PhD thesis, University of Bath, 2019.
- [79] B. Ralph. Herbrand Proofs and Expansion Proofs as Decomposed Proofs. *Journal of Logic and Computation*, 30(8):1711–1742, 10 2020.
- [80] J. R. Shoenfield. *Mathematical Logic*. A K Peters/CRC Press, 1967.
- [81] R. Statman. Lower Bounds on Herbrand’s Theorem. *Proceedings of the American Mathematical Society*, 75(1):104–107, 1979.
- [82] C. Stewart and P. Stouppa. A systematic proof theory for several modal logics. In R. Schmidt, I. Pratt-Hartmann, M. Reynolds, and H. Wansing, editors, *Advances in Modal Logic (AiML)*, volume 5, pages 309–333. King’s College Publications, 2005.
- [83] L. Straßburger. A local system for linear logic. In M. Baaz and A. Voronkov, editors, *Logic for Programming, Artificial Intelligence, and Reasoning (LPAR)*, volume 2514 of *Lecture Notes in Computer Science*, pages 388–402. Springer, 2002.
- [84] A. Tarski. The concept of truth in formalized languages. In A. Tarski, editor, *Logic, semantics, metamathematics*, pages 152–278. Clarendon Press, 1956.
- [85] A. Tiu. A local system for intuitionistic logic. In M. Hermann and A. Voronkov, editors, *Logic for Programming, Artificial Intelligence, and Reasoning (LPAR)*, volume 4246 of *Lecture Notes in Computer Science*, pages 242–256. Springer, 2006.
- [86] A. Tiu. A system of interaction and structure ii: The need for deep inference. *Logical Methods in Computer Science*, Volume 2, Issue 2, Apr. 2006.
- [87] R. Zach. *Hilbert’s Finitism: Historical, Philosophical, and Metamathematical Perspectives*. PhD thesis, University of California, Berkeley, 2001.
- [88] R. Zach. The practice of finitism: Epsilon calculus and consistency proofs in hilbert’s program. *Synthese*, 137(1):211–259, 2003.
- [89] R. Zach. Semantics and proof theory of the epsilon calculus. In S. Ghosh and S. Prasad, editors, *Logic and Its Applications*, pages 27–47, Berlin, Heidelberg, 2017. Springer Berlin Heidelberg.

## A Appendix

### A.1 Tangential Remarks

The intelligence has nothing to discover, it has only to clear the ground. It is only good for servile tasks.

---

*S. Weil*

For those who relinquish the intellect, the world is anything but subtle or complex: it is immediately clear; it appears subtle only to the intellect that struggles laboriously and sees no end to its struggle.

---

*L. E. J. Brouwer*

Critics in their secret hearts love continuities, but he who lives with continuity alone cannot be a poet. The God of poets is not Apollo, who lives in rhythm of recurrence, but the bald gnome Error, who lives at the back of a cave; and skulks forth only at irregular intervals, to feast upon the mighty dead, in the dark of the moon.

---

*H. Bloom*

My father was killed in the Boer War. He went off and joined the Irish Brigade and fought for the Boers.

---

*G. B. Edwards*

In all the works on pedagogy that ever I read — and they have been many, big, and heavy — I don't remember that any one has advocated a system of teaching by practical jokes, mostly cruel. That, however, describes the method of our great teacher, Experience.

---

*C. S. Peirce*

Passion has little to do with euphoria and everything to do with patience. It is not about feeling good. It is about endurance. Like patience, passion comes from the same Latin root: pati. It does not mean to flow with exuberance. It means to suffer.

---

*M. Z. Danielewski*

Courage, not cleverness; not even inspiration, is the grain of mustard that grows up to be a great tree.

---

*L. Wittgenstein*