ONE-AND-A-HALFTH-ORDER LOGIC

MURDOCH J. GABBAY AND AAD MATHIJSSEN

ABSTRACT. The practice of first-order logic is replete with meta-level concepts. Most notably there are meta-variables ranging over formulae, variables, and terms, and properties of syntax such as alpha-equivalence, capture-avoiding substitution and assumptions about freshness of variables with respect to meta-variables. We present one-and-a-halfth-order logic, in which these concepts are made explicit. We exhibit both sequent and algebraic specifications of one-and-a-halfth-order logic derivability, show them equivalent, show that the derivations satisfy cut-elimination, and prove correctness of an interpretation of first-order logic within it.

We discuss the technicalities in a wider context as a case-study for nominal algebra, as a logic in its own right, as an algebraisation of logic, as an example of how other systems might be treated, and also as a theoretical foundation for future implementation.

Partially supported by EPSRC grant number EP/C013573/1.

Contents

1. Introduction

Consider the following valid sequents in first-order predicate logic with equality (FOL) [3, 5], written in standard notation:

- $\vdash \phi \supset (\psi \supset \phi)$,
- if $a \notin fn(\phi)$ then $\phi \vdash \phi[\![a \mapsto t]\!]$,
- if $b \notin fn(\phi)$ then $\forall a.\phi \vdash \forall b.(\phi[\![a \mapsto b]\!])$,
- if $a \notin fn(\phi)$ then $\phi \vdash \forall a.\phi$,
- if $a \notin fn(\phi)$ then $\phi, \psi \vdash \forall a.\phi$,
- if $a \notin fn(\phi)$ then $\phi, \forall a.(\phi \supset \psi) \vdash \forall a.\psi$,
- $\forall b.\forall a.\phi \vdash \forall a.(\phi[\![b \mapsto a]\!])$.

These sequents cannot be derived in FOL, since derivations involve FOL syntax only, while the syntax of the sequents just given contains meta-variables ϕ, ψ, a, b and t. These are *not* FOL syntax, they *vary over* it. Also we refer to properties of syntax when we write '$a \notin fn(\phi)$' and '$\phi[\![a \mapsto t]\!]$', but FOL syntax cannot represent these explicitly.

Of course to us humans this is all obvious. One reason is that the derivations fall into a limited number of schema. For example the 'derivation' below on the left

$$\dfrac{\dfrac{\dfrac{}{\psi, \phi \vdash \phi}\,(\mathbf{Ax})}{\phi \vdash \psi \supset \phi}\,(\supset\mathbf{R})}{\vdash \phi \supset (\psi \supset \phi)}\,(\supset\mathbf{R}) \qquad\qquad \dfrac{\dfrac{\dfrac{}{\bot, \bot \vdash \bot}\,(\mathbf{Ax})}{\bot \vdash \bot \supset \bot}\,(\supset\mathbf{R})}{\vdash \bot \supset (\bot \supset \bot)}\,(\supset\mathbf{R})$$

is not a derivation, but it obviously represents a schema of derivations of which the (real) derivation on the right is an instance setting ϕ and ψ to $\bot$. But is there a logic in which the beast on the left is a derivation too?

It is not a new observation that meta-variables varying over syntax are not syntax, and that schematic derivations are not derivations [27, page 7] (Hodges calls them 'argument schema'). Many authors do leave meta-variables at the meta-level. Some suggest that this is where they belong.

Yet logic teaches us that reasoning can and should be formalised, not only its conclusions. So if we use meta-variables in reasoning, we can and should ask 'what is the mathematics of this reasoning'?

This paper presents **one-and-a-halfth-order logic**. This logic generalises first-order logic by adding explicit meta-variables (**unknowns**) P and Q which in the logic represent the ϕ and ψ of the derivations above.

If we think about how we might do this we immediately encounter serious technical barriers, due to the interaction of unknowns with quantifiers.

- $\forall a.\phi$ and $\forall b.\phi$ need not be *α-convertible* if ϕ mentions a and b free. So if we permit terms with unknowns such as $\forall a.P$ and $\forall a.Q$, what is the suitable generalisation of α-equivalence? It is not acceptable to write 'a does not occur in the syntax of P and Q' (even though this is quite true) — because P and Q are syntax representing *unknown predicates* and should also represent the intuition of ϕ and ψ that they might be instantiated to predicates mentioning a.

  Yet we do need to generalise α-equivalence, for example so that we can represent the $\forall$ right quantifier introduction rule ($\forall\mathbf{R}$):

  $$\frac{\phi \vdash \psi}{\phi \vdash \forall a.\psi}\,(\forall\mathbf{R}) \quad (a \notin \mathit{fn}(\phi))$$

  Here as is well-known we expect to be able to α-rename a to guarantee the side-condition. We would still like to be able to do so in a derivation with unknowns, but it is absolutely not clear how to do so if we are just given a sequent $P \vdash \forall a.Q$, because P and Q are just variable symbols (representing unknown predicates) and it is quite unclear how to rename a in Q to avoid capture in P.
- In the presence of meta-variables, *substitution* becomes nontrivial. For example $\phi[\![a \mapsto t]\!]$ where $[\![a \mapsto t]\!]$ means 'replace a by t' has some meaning if we can rely on ϕ and t being meta-variables which are bound to 'real' syntax. What is a correct representation of substitution which permits us to write $P[\![a \mapsto T]\!]$ and thus represent the $\forall$ left introduction rule:

  $$\frac{\phi[\![a \mapsto t]\!] \vdash \psi}{\forall a.\phi \vdash \psi}\,(\forall\mathbf{L})$$

So in this paper we are seeking to construct a logic that enriches the syntax of first-order logic with explicit meta-variables, and which *still* permits generalised forms of α-equivalence, substitution, capture-avoidance, and quantifier introduction rules which are only ϵ (a very small distance) away from the 'schematic derivations' which we see all the time in informal practice, such as:

$$\dfrac{\dfrac{\begin{matrix}\vdots\\ \phi, (\phi \supset \psi)[\![a \mapsto a]\!] \vdash \psi\end{matrix}}{\phi, \forall a.(\phi \supset \psi) \vdash \psi}\,(\forall\mathbf{L})}{\phi, \forall a.(\phi \supset \psi) \vdash \forall a.\psi}\,(\forall\mathbf{R}) \quad (a \notin \mathit{fn}(\phi))$$

We shall demonstrate not only that such a logic exists, but that it satisfies many of the good properties of first-order logic including cut-elimination. Even though one-and-a-halfth-order logic is a completely formal logic with clearly-specified derivation rules, we shall see that reasoning in one-and-a-halfth-order logic is remarkably close to reasoning in informal practice using schematic derivations. One of the obvious future applications of this logic is as a basis for a theorem-prover with some of the generality of second-order logic, but with the flavour of familiar pencil-and-paper schematic derivations in first-order logic.

Map of the paper. In Section 2 we introduce the syntax of one-and-a-halfth-order logic. In Sections 3 to 5 we develop a sequent calculus and establish proof-theoretical properties including cut-elimination. In Section 6 we give an equational axiomatisation of one-and-a-halfth-order logic and show it equivalent to the sequent calculus. In Section 7 we show that a subset of one-and-a-halfth-order logic is equivalent to first-order logic. We discuss related and future work in the Conclusions.

2. Nominal terms

We need a syntax in which expressions with meta-variables may be represented. Examples we have already mentioned include $\forall a.\phi$, $a \notin \mathit{fn}(\phi)$, and $\phi[\![a \mapsto t]\!]$.

We use Nominal Terms [42] because they offer built-in support for meta-variables, abstraction, and freshness in a way that is close to informal practice.

2.1. **Sorts and terms.**

Our sorts will be very tailored to the needs of this paper; see elsewhere for a more general exposition [18, 19].

Fix **base sorts**:

- $\mathbb{F}$ of **formulae**,
- $\mathbb{T}$ of **terms**.

Fix a countably infinite collection $\mathbb{A}$ of **atoms**. Then **sorts** τ are inductively defined by:

$$\tau \quad ::= \quad \mathbb{F} \mid \mathbb{T} \mid [\mathbb{A}]\mathbb{F} \mid [\mathbb{A}]\mathbb{T}.$$

The intuition of $[\mathbb{A}]\tau$ is 'elements of τ with an atom abstracted'.

$[\mathbb{A}]\tau$ has no intuitive functional denotation; we should *not* think of $[\mathbb{A}]\tau$ as a space of functions (in a standard notation: as $\mathbb{A} \to \tau$). So for example $[\tau']\tau$ is not a valid sort. We should think of $[\mathbb{A}]\tau$ as the set of α-equivalence classes of elements of τ with a distinguished bound atom.

So much for the sorts. Now we construct the terms.

Let atoms $a, b, c, \ldots$ be the elements of $\mathbb{A}$. For each sort $\tau \in \{\mathbb{F}, \mathbb{T}\}$ fix a countably infinite collection $X_\tau, Y_\tau, Z_\tau, \ldots$ of **unknowns of sort** τ. Atoms and sets of unknowns are all distinct.

Atoms represent object-level variable symbols, for examples see a, b in the Introduction. Unknowns represent meta-level variables, for examples see ϕ, ψ, t in the Introduction. We may drop the sorting subscript and write X_τ as X. We may also write $X : \tau$ as shorthand for 'X, which has sort τ'. We tend to give unknowns of sort $\mathbb{F}$ names P, Q, R and unknowns of sort $\mathbb{T}$ names T, U.

A **permutation** π of atoms is a total bijection $\mathbb{A} \to \mathbb{A}$ with **finite support**, meaning that for some finite set of atoms (which may be empty) $\pi(a) \neq a$, *but* for all atoms not in that set, $\pi(a) = a$.

This is a mathematical notion of 'most': π is a bijection on atoms such that $\pi(a) = a$ for *most* a.

As usual, we write **Id** for the **identity** permutation, $\pi^{\text{-}1}$ for the **inverse** of π, and $\pi \circ \pi'$ for the **composition** of π and π', i.e. $(\pi \circ \pi')(a) = \pi(\pi'(a))$. **Id** is also the identity of composition, i.e. $\mathbf{Id} \circ \pi = \pi$ and $\pi \circ \mathbf{Id} = \pi$.

Terms t, u, v are inductively defined by:

$$\begin{array}{rcl} t & ::= & a \mid \pi \cdot X \mid [a]t \mid \bot \mid t \supset t \mid \forall t \mid t \approx t \mid \mathsf{sub}(t, t) \\ & \mid & \mathsf{f}_1(t_1, \ldots, t_{n_1}) \mid \mathsf{f}_2(t_1, \ldots, t_{n_2}) \mid \ldots \end{array}$$

Call $\pi \cdot X$ a **moderated unknown**. We may abbreviate $\mathbf{Id} \cdot X$ to X.

We call the symbols

$$\bot, \quad \supset, \quad \forall, \quad \approx, \quad \mathsf{sub}, \quad \mathsf{f}_1, \quad \mathsf{f}_2, \quad \ldots$$

term-formers. We shall usually let f vary over term-formers.

Amongst the f_i the reader should understand symbols such as 0, S, $+$, $\mathsf{issocrates}$ and greek, but also λ, Σ, $\mathsf{fixpoint}$, $\exists!$ (standard notation for the 'there exists exactly one' quantifier), and so on.

We define **sorting assertions** $t : \tau$ (t has sort τ) inductively by:

$$\frac{}{a : \mathbb{T}} \qquad \frac{}{\pi \cdot X_\tau : \tau} \ (\tau \in \{\mathbb{F}, \mathbb{T}\}) \qquad \frac{t : \tau}{[a]t : [\mathbb{A}]\tau} \ (\tau \in \{\mathbb{F}, \mathbb{T}\})$$

$$\frac{}{\bot : \mathbb{F}} \qquad \frac{t : \mathbb{F} \quad u : \mathbb{F}}{t \supset u : \mathbb{F}} \qquad \frac{t : [\mathbb{A}]\mathbb{F}}{\forall t : \mathbb{F}}$$

$$\frac{t : \mathbb{T} \quad u : \mathbb{T}}{t \approx u : \mathbb{F}} \qquad \frac{t : [\mathbb{A}]\tau \quad u : \mathbb{T}}{\mathsf{sub}(t, u) : \tau} \ (\tau \in \{\mathbb{F}, \mathbb{T}\})$$

There may be additional sorting rules for other term formers, such as:

$$\frac{}{0 : \mathbb{T}} \qquad \frac{t : \mathbb{T}}{\mathsf{S}(t) : \mathbb{T}} \qquad \frac{t : \mathbb{T} \quad u : \mathbb{T}}{t + u : \mathbb{T}} \qquad \frac{t : \mathbb{T}}{\mathsf{issocrates}(t) : \mathbb{F}} \qquad \frac{t : \mathbb{T}}{\mathsf{greek}(t) : \mathbb{F}}$$

$$\frac{t : [\mathbb{A}]\mathbb{T}}{\lambda t : \mathbb{T}} \qquad \frac{t : [\mathbb{A}]\mathbb{T}}{\Sigma t : \mathbb{T}} \qquad \frac{t : [\mathbb{A}]\mathbb{T}}{\mathsf{fixpoint}(t) : \mathbb{T}} \qquad \frac{t : [\mathbb{A}]\mathbb{F}}{\exists! t : \mathbb{F}}$$

We write $t : \tau$ as a shorthand for 't of sort τ'. We may call terms $\phi : \mathbb{F}$ **formulae**.

Let a be any atom, $t, u : \mathbb{T}$ be any terms, and ϕ, ψ be any formulae. We discuss some intuitions and introduce some sugar:

- $\bot$ represents false.
- $\phi \supset \psi$ is a logical implication.
- We write $\forall([a]\phi)$ as $\forall[a]\phi$; this represents a universal quantification (which takes an *abstraction* of a formula and yields a formula).
- We write $\mathsf{sub}([a]\phi, t)$ and $\mathsf{sub}([a]u, t)$ as $\phi[a \mapsto u]$ and $t[a \mapsto u]$, representing capture-avoiding substitution in the object-language. We may also refer to these terms as **substitutions**.
- $t \approx u$ is equality in the object-language.
- a is a term of sort $\mathbb{T}$, representing an object-level variable symbol of sort $\mathbb{T}$.
- We may use some other standard notation for term-formers with suggestive names; e.g. we may write $t + u$ as shorthand for $+(t, u)$.
- We use standard **classical logic sugar**:

$$\begin{array}{rl} \neg\phi \text{ is } \phi \supset \bot & \top \text{ is } \neg\bot \\ \phi \wedge \psi \text{ is } \neg(\phi \supset \neg\psi) & \phi \vee \psi \text{ is } \neg\phi \supset \psi \\ \phi \Leftrightarrow \psi \text{ is } (\phi \supset \psi) \wedge (\psi \supset \phi) & \exists[a]\phi \text{ is } \neg\forall[a]\neg\phi \end{array}$$

  Note that these are *abbreviations*, not term-formers.

  To save on (unnecessary) parentheses, take $[a]\_$, $\_[\_ \mapsto \_]$, $\approx$, $\{\neg, \forall, \exists\}$, $\{\wedge, \vee\}$, $\supset$, $\Leftrightarrow$ as the descending order of precedence. So for example $P \wedge Q \supset R \vee S$ is $(P \wedge Q) \supset (R \vee S)$ and $\forall[a]P \wedge Q$ is $(\forall[a]P) \wedge Q$. Also let $\wedge$, $\vee$, $\supset$ and $\Leftrightarrow$ associate to the right; for example $P \supset Q \supset R$ is $P \supset (Q \supset R)$.

If a term-former forms terms of sort $\mathbb{T}$ call it an **object-level term-former**. If a term-former forms terms of sort $\mathbb{F}$ call it a **predicate term-former**. 0, S, $+$, λ, Σ and $\mathsf{fixpoint}$ are object-level term-formers and $\mathsf{issocrates}$, greek and $\exists!$ are predicate term-formers. These extra term-formers would cause no difficulties for the results which follow — aside from some extra cases.

We will not want any term-former to *make* terms of sort $[\mathbb{A}]\mathbb{T}$ or $[\mathbb{A}]\mathbb{F}$ directly; the correct way to form an abstraction is to use $[a]$.

Write **syntactic identity** of terms t, u as $t \equiv u$. This emphasises the difference from provable equality $t = u$, which is a logical assertion defined later in Section 4.2, and object-level equality $t \approx u$, which is a term.

Note that the sorting system is such that a well-sorted term of the form $\forall t$ *must* be of the form $\forall[a]t'$ (so $t \equiv [a]t'$), and a well-sorted term of the form $\mathsf{sub}(t_1, t_2)$ *must* be of the form $t[a \mapsto t_2]$.

We write $a \in t$ (or $X \in t$) for 'a (or X) **occurs in (the syntax of)** t'. Occurrence is literal, e.g. $a \in [a]a$ and $a \in \pi \cdot X$ if $\pi(a) \neq a$. We omit inductive definitions. Similarly we may write $a \notin t$ and $X \notin t$ for 'does not occur in the syntax of t'.

Call t **closed** when t mentions no unknowns — t may still mention atoms, e.g. the terms a and $[a]b$ are closed and the terms X and $[a]X$ are not.

2.2. **Freshnesses.** A **freshness (assertion)** is a pair $a\#t$ of an atom and a term. Intuitively we should read $a\#t$ as meaning '$a \notin fn(t)$' or in words 'a is fresh for t'.

Call the assertion $a\#X$ (so $t \equiv X$) **primitive**. Write Δ for a (possibly infinite) set of *primitive* freshnesses and call it a **freshness context**. We may drop set brackets in freshnesses, e.g. writing $a\#t, b\#u$ for $\{a\#t, b\#u\}$. Also, we may write $a\#t, u$ for $a\#t, a\#u$.

$$\frac{}{a\#b}\,(\#\mathbf{ab}) \qquad \frac{\pi^{\text{-}1}(a)\#X}{a\#\pi\cdot X}\,(\#\mathbf{X}) \quad (\pi \neq \mathbf{Id})$$

$$\frac{}{a\#[a]t}\,(\#[]\mathbf{a}) \qquad \frac{a\#t}{a\#[b]t}\,(\#[]\mathbf{b}) \qquad \frac{a\#t_1 \ \cdots \ a\#t_n}{a\#\mathsf{f}(t_1,\dots,t_n)}\,(\#\mathbf{f})$$

FIGURE 1. Derivation rules for freshness

A reason this notion is quite subtle in nominal techniques is the unknowns X; $a\#X$ is *not* necessarily true even though $a \notin X$ *is* a fact of the syntax. X represents an unknown term in the syntax; $a\#X$ has the quality of a *promise* or *assertion* about what term that can be, or put another way, about what we can instantiate X to.

Freshness enjoys a notion of derivation which will be very useful:

Definition 2.1. *Write $\Delta \vdash a\#t$ when a derivation of a freshness assertion $a\#t$ exists using the elements of Δ as assumptions, according to the rules in Figure 1. Say that Δ **entails** $a\#t$ or $a\#t$ **is derivable from** Δ.*

In Figure 1...

- a and b *permutatively* range over atoms, i.e. a and b represent any two *distinct* atoms;
- π ranges over permutations.[1]
- X ranges over unknowns;
- t and $t_1,\dots,t_n$ range over terms;
- f ranges over term-formers.

We use similar conventions henceforth.

Examples of derivable freshness assertions are:

$$\vdash a\#\forall[a]P \qquad a\#T \vdash a\#(a\approx a)[a\mapsto T] \qquad a\#X \vdash b\#(b\ a)\cdot X.$$

Examples of *non*-derivable freshness assertions are:

$$\vdash a\#a \qquad \vdash a\#\forall[b]P \qquad \vdash a\#(a\approx a)[a\mapsto T] \qquad a\#X \vdash a\#(b\ a)\cdot X.$$

Note that freshness is decidable; we obtain an algorithm by reading the rules in Figure 1 bottom-up.

3. DERIVATIONS OF ONE-AND-A-HALFTH-ORDER LOGIC

Recall that by our terminology a formula ϕ is a term of sort $\mathbb{F}$.

Let **(formula) contexts** Φ,Ψ be finite (possibly empty) sets of formulae. A **sequent** is a triple $\Phi \vdash_{\vartriangle} \Psi$ where Δ is a freshness context and Φ and Ψ are formula contexts; when a context appears to the right of $\vdash$ we may call it a **co-context**.

We may write ϕ for $\{\phi\}$, ϕ,Φ for $\{\phi\}\cup\Phi$, and Φ,Φ' for $\Phi\cup\Phi'$, and we may omit empty formula contexts, e.g. writing $\vdash_{\vartriangle}$ for $\emptyset\vdash_{\vartriangle}\emptyset$.

[1] The $(\#\mathbf{X})$ rule excludes the empty permutation $\mathbf{Id}$. While there is no mathematical reason for this, there is a nice *computational* one: the algorithm obtained by reading rules bottom-up, must terminate.

Extend the notions of occurrence, closedness, permutation actions and substitution action to formula contexts elementwise; for example $a \in \Phi$ if $a \in \phi$ for some $\phi \in \Phi$.

Definition 3.1. *Let the **valid sequents** of one-and-a-halfth-order logic be inductively specified by the rules in Figure 2.*

We may call the set of valid sequents an **entailment relation**.

Our rules resemble those of Gentzen's sequent calculus for classical first-order logic with equality [8, 23, 39], but with the following distinctive features:

- *Unknown formulae* and *unknown terms* are represented explicitly as unknowns of sort $\mathbb{F}$ and $\mathbb{T}$ respectively.
- We can make *freshness* assumptions about unknowns (using (**Fr**)), and these affect derivability, for example in ($\forall$**R**).
- A theory of equality of terms up to α-equivalence and capture-avoiding substitution is represented by an equality $\Delta \vdash_{\mathsf{SUB}} t = u$ in a theory SUB (formally defined in the next section).
- Side-conditions on substitution, freshness and atoms not occurring in terms, are all *decidable*.

Call (**StructL**) and (**StructR**) **structural rules**. (**Cut**) can emulate them, but we would lose cut-elimination. Note that the side-conditions of these rules refer to equational derivability in SUB. (**StructL**) and (**StructR**) exist to help us manage substitutions, see the example derivations below.

(**Fr**) helps us to introduce new atoms into a derivation. The difficulty is that one-and-a-halfth-order logic has explicit meta-variables, so we need (**Fr**) so that we can enrich the freshness context with information that the atom is fresh for those meta-variables.

Example derivations in one-and-a-halfth-order logic are in Figure 3. We give some intuitions:

- $\vdash P \supset (Q \supset P)$ represents a family of tautologies of propositional logic $\phi \supset (\psi \supset \phi)$. The only difference is that here we are using the provision of explicit meta-variables to represent this family directly as a single sequent.
- $P \vdash_{\mathsf{a\#P}} P[a \mapsto T]$ represents a family of tautologies of predicate logic. The condition $a\#P$ intuitively guarantees that whatever formula P represents, it is not one that mentions a free in its syntax. It corresponds to writing $a \notin \mathit{fn}(\phi)$.

  A side-condition verifies formally that if $a \notin \mathit{fn}(\phi)$ then ϕ and $\phi[a \mapsto t]$ are the same. Since unknowns represent meta-variables ranging over formulae and terms, such an equality is not a syntactic fact; it becomes something worthy of proof. We develop the theory of this equality next.
- $\forall[a]P \vdash_{\mathsf{b\#P}} \forall[b](P[a \mapsto b])$ is α-equivalence... in the presence of unknowns! The derivation exploits the power to prove equalities in SUB.
- $P \vdash_{\mathsf{a\#P}} \forall[a]P$ expresses that if a does not appear free in ϕ and ϕ holds, then so does $\forall a.\phi$.
- $P, Q \vdash_{\mathsf{a\#P}} \forall[a]P$ expresses the same as the previous example, except that we have an additional assumption Q. The derivation becomes significantly more complex: we cannot use ($\forall$**R**) on $\forall[a]P$ because we do not know $a\#Q$.

$$\frac{}{\phi,\, \Phi \vdash_{\Delta} \Psi,\, \phi}\,(\mathbf{Ax}) \qquad \frac{}{\bot,\, \Phi \vdash_{\Delta} \Psi}\,(\bot\mathbf{L})$$

$$\frac{\Phi \vdash_{\Delta} \Psi,\, \phi \quad \psi,\, \Phi \vdash_{\Delta} \Psi}{\phi \supset \psi,\, \Phi \vdash_{\Delta} \Psi}\,(\supset\mathbf{L}) \qquad \frac{\phi,\, \Phi \vdash_{\Delta} \Psi,\, \psi}{\Phi \vdash_{\Delta} \Psi,\, \phi \supset \psi}\,(\supset\mathbf{R})$$

$$\frac{\phi[a \mapsto t],\, \Phi \vdash_{\Delta} \Psi}{\forall[a]\phi,\, \Phi \vdash_{\Delta} \Psi}\,(\forall\mathbf{L}) \qquad \frac{\Phi \vdash_{\Delta} \Psi,\, \psi}{\Phi \vdash_{\Delta} \Psi,\, \forall[a]\psi}\,(\forall\mathbf{R}) \quad (\Delta \vdash a\#\Phi, \Psi)$$

$$\frac{\phi[a \mapsto t'],\ \Phi \vdash_{\Delta} \Psi}{t' \approx t,\ \phi[a \mapsto t],\ \Phi \vdash_{\Delta} \Psi}\,(\approx\mathbf{L}) \qquad \frac{}{\Phi \vdash_{\Delta} \Psi,\ t \approx t}\,(\approx\mathbf{R})$$

$$\frac{\phi',\, \Phi \vdash_{\Delta} \Psi}{\phi,\, \Phi \vdash_{\Delta} \Psi}\,(\mathbf{StructL}) \quad (\Delta \vdash_{\mathsf{SUB}} \phi' = \phi) \qquad \frac{\Phi \vdash_{\Delta} \Psi,\, \psi'}{\Phi \vdash_{\Delta} \Psi,\, \psi}\,(\mathbf{StructR}) \quad (\Delta \vdash_{\mathsf{SUB}} \psi' = \psi)$$

$$\frac{\Phi \vdash_{\Delta, a\#X_1, \ldots, a\#X_n} \Psi}{\Phi \vdash_{\Delta} \Psi}\,(\mathbf{Fr}) \quad (n \geq 1,\ a \notin \Phi, \Psi, \Delta)$$

$$\frac{\Phi \vdash_{\Delta} \Psi,\, \phi \qquad \phi',\, \Phi \vdash_{\Delta} \Psi}{\Phi \vdash_{\Delta} \Psi}\,(\mathbf{Cut}) \quad (\Delta \vdash_{\mathsf{SUB}} \phi = \phi')$$

FIGURE 2. Sequent calculus for one-and-a-halfth-order logic

The solution is to use (**Fr**) to generate $b\#P, Q$, use structural rules to α-rename, and *then* use ($\forall$**R**). This use of (**Fr**) is essential: we need this mechanism to introduce a fresh atom into the derivation.

- $P, \forall[a](P \supset Q) \vdash_{\mathsf{a\#P}} \forall[a]Q$ represents another family of tautologies of predicate logic [12, page 4, axiom (2a)]. For the instance of ($\forall$**R**) to be valid we must show $a\#\forall[a](P \supset Q)$. We have made no assumptions about what is fresh for Q, but the abstraction by a guarantees this property anyway.
- $\forall[b]\forall[a]P \vdash \forall[a](P[b \mapsto a])$ is a relatively non-trivial tautology which might be written in semi-formal notation as '$\forall a.\forall b.\phi(a, b)$ implies $\forall a.\phi(a, a)$'.

4. Nominal algebra

We now develop the theory of equality of nominal terms up to substitution and formally construct the judgement $\Delta \vdash_{\mathsf{SUB}} t = u$ which we used in Figure 2. This framework is called nominal algebra [18, 19]. In this section we omit proofs of most lemmas and theorems. Full proofs are available elsewhere [19].

4.1. **Permutations and substitutions.** We write $\pi \cdot t$ for the **action** of a permutation on a term, defined inductively on syntax by:

$$\pi \cdot a \ \equiv\ \pi(a) \qquad \pi \cdot (\pi' \cdot X) \ \equiv\ (\pi \circ \pi') \cdot X \qquad \pi \cdot [a]t \ \equiv\ [\pi(a)](\pi \cdot t)$$
$$\pi \cdot \mathsf{f}(t_1, \ldots, t_n) \ \equiv\ \mathsf{f}(\pi \cdot t_1, \ldots, \pi \cdot t_n)$$

Lemma 4.1. $\pi \cdot (\pi' \cdot t) \equiv (\pi \circ \pi') \cdot t$ *and* $\mathbf{Id} \cdot t \equiv t$.

A **substitution** σ is a finitely supported sort-respecting function from unknowns to terms. Here, **finitely supported** means that for some finite set of unknowns $\sigma(X) \not\equiv \mathbf{Id} \cdot X$, but for all other unknowns $\sigma(X) \equiv \mathbf{Id} \cdot X$. **Sort-respecting**

$$\dfrac{\dfrac{\dfrac{}{Q, P \vdash_{\emptyset} P}\,(\mathbf{Ax})}{P \vdash_{\emptyset} Q \supset P}\,(\supset\mathbf{R})}{\vdash_{\emptyset} P \supset (Q \supset P)}\,(\supset\mathbf{R})
\qquad
\dfrac{\dfrac{}{P \vdash_{a\#P} P}\,(\mathbf{Ax})}{P \vdash_{a\#P} P[a \mapsto T]}\,(\mathbf{StructR}) \quad (a\#P \vdash_{\mathsf{SUB}} P = P[a \mapsto T])$$

$$\dfrac{\dfrac{}{\forall[a]P \vdash_{b\#P} \forall[a]P}\,(\mathbf{Ax})}{\forall[a]P \vdash_{b\#P} \forall[b](P[a \mapsto b])}\,(\mathbf{StructR}) \quad (b\#P \vdash_{\mathsf{SUB}} \forall[a]P = \forall[b](P[a \mapsto b]))$$

$$\dfrac{\dfrac{}{P \vdash_{a\#P} P}\,(\mathbf{Ax})}{P \vdash_{a\#P} \forall[a]P}\,(\forall\mathbf{R}) \quad (a\#P \vdash a\#P)$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{P, Q \vdash_{a\#P, b\#P, Q} P}\,(\mathbf{Ax})}{P, Q \vdash_{a\#P, b\#P, Q} (a\ b) \cdot P}\,(\mathbf{StructR}) \quad (a\#P, b\#P, Q \vdash_{\mathsf{SUB}} P = (a\ b) \cdot P)}{P, Q \vdash_{a\#P, b\#P, Q} \forall[b](a\ b) \cdot P}\,(\forall\mathbf{R}) \quad (a\#P, b\#P, Q \vdash b\#P, Q)}{P, Q \vdash_{a\#P, b\#P, Q} \forall[a]P}\,(\mathbf{StructR}) \quad (a\#P, b\#P, Q \vdash_{\mathsf{SUB}} \forall[b](a\ b) \cdot P = \forall[a]P)}{P, Q \vdash_{a\#P} \forall[a]P}\,(\mathbf{Fr}) \quad (b \notin P,\ Q,\ \forall[a]P,\ a\#P)$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{P \vdash_{a\#P} Q, P}\,(\mathbf{Ax}) \qquad \dfrac{}{Q, P \vdash_{a\#P} Q}\,(\mathbf{Ax})}{P, P \supset Q \vdash_{a\#P} Q}\,(\supset\mathbf{L})}{P, (P \supset Q)[a \mapsto a] \vdash_{a\#P} Q}\,(\mathbf{StructL}) \quad (a\#P \vdash_{\mathsf{SUB}} P \supset Q = (P \supset Q)[a \mapsto a])}{P, \forall[a](P \supset Q) \vdash_{a\#P} Q}\,(\forall\mathbf{L})}{P, \forall[a](P \supset Q) \vdash_{a\#P} \forall[a]Q}\,(\forall\mathbf{R}) \quad (a\#P \vdash a\#P,\ \forall[a](P \supset Q))$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{P[b \mapsto c][a \mapsto c] \vdash_{c\#P} P[b \mapsto c][a \mapsto c]}\,(\mathbf{Ax})}{\forall[a](P[b \mapsto c]) \vdash_{c\#P} P[b \mapsto c][a \mapsto c]}\,(\forall\mathbf{L})}{(\forall[a]P)[b \mapsto c] \vdash_{c\#P} P[b \mapsto c][a \mapsto c]}\,(\mathbf{StructL}) \quad (c\#P \vdash_{\mathsf{SUB}} \forall[a](P[b \mapsto c]) = (\forall[a]P)[b \mapsto c])}{\forall[b]\forall[a]P \vdash_{c\#P} P[b \mapsto c][a \mapsto c]}\,(\forall\mathbf{L})}{\forall[b]\forall[a]P \vdash_{c\#P} \forall[c](P[b \mapsto c][a \mapsto c])}\,(\forall\mathbf{R}) \quad (c\#P \vdash c\#\forall[b]\forall[a]P)}{\forall[b]\forall[a]P \vdash_{c\#P} \forall[a](P[b \mapsto a])}\,(\mathbf{StructR}) \quad (c\#P \vdash_{\mathsf{SUB}} \forall[c](P[b \mapsto c][a \mapsto c]) = \forall[a](P[b \mapsto a]))}{\forall[b]\forall[a]P \vdash_{\emptyset} \forall[a](P[b \mapsto a])}\,(\mathbf{Fr}) \quad (c \notin \forall[b]\forall[a]P,\ \forall[a](P[b \mapsto a]))$$

FIGURE 3. Example derivations in one-and-a-halfth-order logic

means that for each X the term $\sigma(X)$ should have the same sort as X. Write $[t_1/X_1, \dots, t_n/X_n]$ for the substitution σ such that $\sigma(X_i) \equiv t_i$ and $\sigma(Y) \equiv \mathbf{Id} \cdot Y$, for all $Y \not\equiv X_i$, $1 \leq i \leq n$. Write $[]$ for the **empty** substitution, which maps each X to $\mathbf{Id} \cdot X$.

Write $a \in \sigma$ if there exists an X such that $a \in \sigma(X)$, and similarly write $a \notin \sigma$ if there is no such X. For example $a \in [a/X]$ and $a \notin []$.

A substitution σ has a natural **action** on terms t, inductively defined by:

$$a\sigma \;\equiv\; a \qquad (\pi \cdot X)\sigma \;\equiv\; \pi \cdot \sigma(X) \qquad ([a]t)\sigma \;\equiv\; [a](t\sigma)$$
$$\mathsf{f}(t_1, \dots, t_n)\sigma \;\equiv\; \mathsf{f}(t_1\sigma, \dots, t_n\sigma)$$

Give substitution and permutation actions higher precedence than abstraction and any of the sugared term-formers, and put substitution before permutation.

Note how substitution interacts with permutation in the case of an unknown, for example $((a\ b) \cdot X)[b/X] \equiv (a\ b) \cdot b \equiv a$. So π in X is 'waiting for a substitution to arrive', as also made formal in the following property:

Lemma 4.2. $\pi \cdot t\sigma \equiv (\pi \cdot t)\sigma$.

Another permutation action will be useful. Write t^π for the **meta-level action** of π on t, which is defined by:

$$a^\pi \;\equiv\; \pi(a) \qquad (\pi' \cdot X)^\pi \;\equiv\; (\pi \circ \pi' \circ \pi^{-1}) \cdot X \qquad ([a]t)^\pi \;\equiv\; [\pi(a)](t^\pi)$$
$$\mathsf{f}(t_1, \dots, t_n)^\pi \;\equiv\; \mathsf{f}({t_1}^\pi, \dots, {t_n}^\pi)$$

Lemma 4.3. *Fix t and π, and let σ map $X \in t$ to $\pi \cdot X$, and σ' map $X \in t$ to $\pi^{-1} \cdot X$. Then $\pi \cdot t \equiv t^\pi\sigma$ and $t^\pi \equiv (\pi \cdot t)\sigma'$.*

So the two permutation actions are interdefinable in the presence of substitution σ; however, sometimes one is more natural than the other, we shall point out how, later.

We extend notation for t^π, $\pi \cdot t$ and $t\sigma$ to freshness contexts Δ as follows:

$$\begin{aligned} \Delta^\pi &= \{\pi(a)\#X \mid a\#X \in \Delta\} \\ \pi \cdot \Delta &= \{\pi(a)\#\pi \cdot X \mid a\#X \in \Delta\} \\ \Delta\sigma &= \{a\#\sigma(X) \mid a\#X \in \Delta\} \end{aligned}$$

Note that Δ^π is a freshness context, but $\pi \cdot \Delta$ and $\Delta\sigma$ need not be.

4.2. **Assertions, axioms and derivations.** An **equality assertion** is a pair of terms $t = u$ of the same sort. Call a pair $\Delta \vdash t = u$ of a finite freshness context Δ and an equality assertion $t = u$ an **axiom**. If $\Delta = \emptyset$, we may write just $t = u$.

Definition 4.4. *Call a (possibly infinite) set of axioms T a* ***theory****.*

Write $\Delta \vdash_{\mathsf{T}} t = u$ when a derivation of $t = u$ exists using the rules in Figure 4, such that every assumption used is a freshnesses from Δ, and for every use of $(\mathbf{ax_A})$ for some A, that A is an element of T. Say that Δ ***entails*** *$t = u$ or $t = u$* ***is derivable from*** *Δ.*

For example,

- In the theory with axioms $a = b$ and $[a]X = [b]Y$, the derivations

  $$\frac{}{b = c}\,(\mathbf{ax_{a=b}}) \qquad\qquad \frac{}{[b]b = [a]a}\,(\mathbf{ax_{[a]X=[b]Y}})$$

  are valid. We take $\pi = (a\ b\ c)$ and any σ, and $\pi = (a\ b)$ and $\sigma = [b/X,\ a/Y]$, respectively.

$$\frac{}{t = t}\,(\mathbf{refl}) \qquad \frac{t = u}{u = t}\,(\mathbf{symm}) \qquad \frac{t = u \quad u = v}{t = v}\,(\mathbf{tran}) \qquad \frac{a\#t \quad b\#t}{(a\ b)\cdot t = t}\,(\mathbf{perm})$$

$$\frac{t = u}{[a]t = [a]u}\,(\mathbf{cong[]}) \qquad \frac{t = u}{\mathsf{f}(t_1,\ldots,t,\ldots,t_n) = \mathsf{f}(t_1,\ldots,u,\ldots,t_n)}\,(\mathbf{congf})$$

$$\frac{\Delta^\pi\sigma}{t^\pi\sigma = u^\pi\sigma}\,(\mathbf{ax_{\Delta\vdash t=u}}) \qquad \begin{array}{c}[a\#X_1,\ldots,a\#X_n] \qquad \Delta\\ \vdots \\ \dfrac{t = u}{t = u}\,(\mathbf{fr}) \quad (n \geq 1,\ a \notin t, u, \Delta)\end{array}$$

FIGURE 4. Derivation rules of nominal algebra

- We cannot use axiom $\mathsf{f}_1(a,b) = \mathsf{f}_2$ to derive that $\mathsf{f}_1(a,a) = \mathsf{f}_2$, because no permutation can identify a with b (assuming term-formers f_1 and f_2).
- Taking C to be $a\#X \vdash [a]X = [b]X$, of the derivations

$$\frac{\dfrac{}{a\#b}\,(\#\mathbf{ab})}{[a]b = [b]b}\,(\mathbf{ax_C}) \qquad \frac{a\#a}{[a]a = [b]a}\,(\mathbf{ax_C})$$

  the left one is valid, but the right one is *not*, because $a\#a$ is not derivable.

In ($\mathbf{ax_{\Delta\vdash t=u}}$), unknowns get arbitrarily *instantiated* and atoms get arbitrarily *permuted*. Thus in an axiom two distinct unknowns represent *any two terms* but two distinct atoms represent *any two distinct atoms*.

In (**fr**) square brackets denote *discharge* in the sense of natural deduction (as in implication introduction [27]); Δ denotes the other assumptions of the derivation of $t = u$.

In a sequent style presentation of nominal algebra, (**fr**) would be

$$\frac{\Delta, a\#X_1,\ldots,a\#X_n \vdash t = u}{\Delta \vdash t = u} \quad (n \geq 1, a \notin t, u, \Delta).$$

To see how (**fr**) increases the power of the system consider a theory with just one axiom $a\#T \vdash T = a$, which we call S. With (**fr**) we can derive $\vdash T = a$. Without (**fr**) we can derive $\vdash b = a$ but we *cannot* derive $\vdash T = a$ (proof omitted).

$$\frac{\dfrac{[a\#T]^1}{T = a}\,(\mathbf{ax_S})}{T = a}\,(\mathbf{fr})^1 \qquad \frac{\dfrac{}{a\#b}\,(\#\mathbf{ab})}{b = a}\,(\mathbf{ax_S})$$

In the left derivation, the superscript number one ^{1} is an annotation associating the instance of the rule (**fr**) with the assumption it discharges in the derivation. This is standard natural deduction notation.

Useful properties of nominal algebra include:

Lemma 4.5. *For any π:*

(1) *if $\Delta \vdash a\#t$ then $\Delta \vdash \pi(a)\#\pi\cdot t$;*

(2) *if $\Delta \vdash_{\mathsf{T}} t = u$ then $\Delta \vdash_{\mathsf{T}} \pi\cdot t = \pi\cdot u$.*

Lemma 4.6. *If $\Delta \vdash_{\mathsf{T}} t = u$ then $\Delta \vdash_{\mathsf{T}} v[t/X] = v[u/X]$.*

Proof. By an easy induction on the structure of v. $\square$

Theorem 4.7. *For any Δ', Δ, σ, if $\Delta' \vdash \Delta\sigma$ then:*
(1) *if $\Delta \vdash a\#t$ then $\Delta' \vdash a\#t\sigma$;*
(2) *if $\Delta \vdash_{\mathsf{T}} t = u$ then $\Delta' \vdash_{\mathsf{T}} t\sigma = u\sigma$.*

This has as an easy corollary:

Corollary 4.8. *If $\Delta \subseteq \Delta'$ then:*
(1) *if $\Delta \vdash a\#t$ then $\Delta' \vdash a\#t$;*
(2) *if $\Delta \vdash_{\mathsf{T}} t = u$ then $\Delta' \vdash_{\mathsf{T}} t = u$.*

The reader may wonder why we do not use the following alternative axiom rule:

$$\frac{\pi \cdot \Delta\sigma}{\pi \cdot t\sigma = \pi \cdot u\sigma}\,(\mathbf{ax}'_{\mathbf{\Delta \vdash t=u}}).$$

This is just a matter of taste. Using $(\mathbf{ax}'_{\mathbf{\Delta \vdash t=u}})$, atoms in the substitution σ are renamed according to permutation π. This we personally find rather mind-bending. For example, from the axiom $[a]X = [b]X$ it is immediate that $\vdash [b]a = [a]a$ is derivable using $(\mathbf{ax}_{\mathbf{[a]X=[b]X}})$ where we choose $\pi = (b\ a)$ and $\sigma = [a/X]$. If we use $(\mathbf{ax}'_{\mathbf{[a]X=[b]X}})$ we must choose $\pi = (b\ a)$ and $\sigma = [b/X]$.

4.3. **The theories CORE and SUB.** We have seen examples of one-and-a-halfth-order logic derivations. We have seen that structural rules and freshnesses play an important rôle in the control of the explicit meta-variables.

In the rest of this section we formally define the valid judgements of the form $\Delta \vdash_{\mathsf{SUB}} t = u$. In [17], we have shown that SUB really is capture-avoiding substitution, and that this theory is decidable. In this section, we will only show some examples. In Section 7, we will show that equality in SUB coincides with the usual notion of capture-avoiding substitution, such as the one used in first-order logic with equality (FOL).

Write CORE for the theory with no axioms. This is a theory of α-equivalence on nominal terms.

Write SUB for the theory with the axioms in Figure 5. Here a, b are distinct atoms, P, Q, R are distinct unknowns of sort $\mathbb{F}$, T, U, V are distinct unknowns of sort $\mathbb{T}$, and X is an unknown of the appropriate sort.

There may also be axioms, which we shall not dwell on, to distribute substitutions through any other term-formers such as λ, $+$, and so on. These cause no extra issues; if the term-former takes terms of abstraction sort the equality should include a freshness side-condition in the same style as $(\forall\mapsto)$.

The following two lemmas are useful:

Lemma 4.9 (α-conversion)**.** $b\#Z \vdash_{\mathsf{CORE}} X[a \mapsto T] = ((b\ a) \cdot X)[b \mapsto T]$

Proof. De-sugaring, we must derive $\mathsf{sub}([a]X, T) = \mathsf{sub}([b](b\ a) \cdot X, T)$ from $b\#X$.

$$\dfrac{\dfrac{\dfrac{}{a\#[a]X}\,(\#[]\mathbf{a}) \qquad \dfrac{b\#X}{b\#[a]X}\,(\#[]\mathbf{b})}{[a]X = [b](b\ a) \cdot X}\,(\mathbf{perm})}{\mathsf{sub}([a]X, T) = \mathsf{sub}([b](b\ a) \cdot X, T)}\,(\mathbf{congf})$$

$\square$

$$
\begin{array}{rlrcl}
(\mathbf{var}\mapsto) & & a[a\mapsto T] & = & T \\
(\#\mapsto) & a\#X \vdash & X[a\mapsto T] & = & X \\
(\supset\mapsto) & & (P\supset Q)[a\mapsto T] & = & (P[a\mapsto t])\supset(Q[a\mapsto t]) \\
(\approx\mapsto) & & (U\approx V)[a\mapsto T] & = & (U[a\mapsto t])\approx(V[a\mapsto t]) \\
(\forall\mapsto) & b\#T \ \vdash & (\forall[b]P)[a\mapsto T] & = & \forall[b](P[a\mapsto T]) \\
(\mathbf{sub}\mapsto) & b\#T \ \vdash & X[b\mapsto U][a\mapsto T] & = & X[a\mapsto T][b\mapsto U[a\mapsto T]] \\
(\mathbf{ren}\mapsto) & b\#X \ \vdash & X[a\mapsto b] & = & (b\ a)\cdot X
\end{array}
$$

FIGURE 5. Axioms of SUB

Lemma 4.10. $\vdash_{\mathsf{SUB}} X[a\mapsto a]=X$ *is derivable.*

Proof.

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{a\#[a]X}\,(\#[]\mathbf{a}) \qquad \dfrac{[b\#X]^1}{b\#[a]X}\,(\#[]\mathbf{b})}{[b](b\ a)\cdot X=[a]X}\,(\mathbf{perm})}{[a]X=[b](b\ a)\cdot X}\,(\mathbf{symm})}{X[a\mapsto a]=((b\ a)\cdot X)[b\mapsto a]}\,(\mathbf{congf}) \qquad \dfrac{\dfrac{[b\#X]^1}{a\#(b\ a)\cdot X}\,(\#\mathbf{X})}{((b\ a)\cdot X)[b\mapsto a]=X}\,(\mathbf{ax}_{\mathbf{ren}\mapsto})}{\dfrac{X[a\mapsto a]=X}{X[a\mapsto a]=X}\,(\mathbf{fr})^1}\,(\mathbf{tran})
$$

□

Lemma 4.11. $a\#U, b\#T \vdash_{\mathsf{SUB}} X[a\mapsto T][b\mapsto U]=X[b\mapsto U][a\mapsto T]$ *is derivable.*

Proof. By $(\mathbf{tran})$, it suffices to derive $X[a\mapsto T][b\mapsto U]=X[b\mapsto U][a\mapsto T[b\mapsto U]]$ and $X[b\mapsto U][a\mapsto T[b\mapsto U]] = X[b\mapsto U][a\mapsto T]$ from assumptions $a\#U$ and $b\#T$. The former follows from axiom $(\mathbf{sub}\mapsto)$ and assumption $a\#U$. By $(\mathbf{congf})$, the latter follows from $T[b\mapsto U]=T$, which follows from the assumption $b\#T$ by axiom $(\#\mapsto)$. □

For a dedicated investigation of SUB see elsewhere [17].

Now we recall Figure 3.

- The side-condition $a\#P \vdash_{\mathsf{SUB}} P\supset Q=(P\supset Q)[a\mapsto a]$ in the derivation of $P, \forall[a](P\supset Q)\vdash_{a\#P} \forall[a]Q$ is an instance of Lemma 4.10.
- The side-condition $a\#P, b\#P, Q\vdash_{\mathsf{SUB}} \forall[b](a\ b)\cdot P=\forall[a]P$ in the derivation of $P, Q\vdash_{\mathsf{a\#P}} \forall[a]P$ is an instance of Lemma 4.9.

In Figure 6 we show derivations of a few of the more interesting side-conditions used in Figure 3. The final derivation is given in two parts, one of which we call Π, for typographic reasons.

5. PROOF-THEORETICAL RESULTS

This section shows two important properties of the sequent calculus for one-and-a-halfth order logic:

- In derivations, atoms may be *permuted* and unknowns may be *instantiated*. We will call these properties equivariance and substitution.

$$\dfrac{\dfrac{a\#P}{P[a\mapsto T]=P}\,(\mathbf{ax}_{\#\mapsto})}{P=P[a\mapsto T]}\,(\mathbf{symm}) \qquad \dfrac{\dfrac{\dfrac{\dfrac{c\#P}{c\#[a]P}\,(\#[]\mathbf{b})}{c\#\forall[a]P}\,(\#\mathbf{f})}{c\#[b]\forall[a]P}\,(\#[]\mathbf{b})}{c\#\forall[b]\forall[a]P}\,(\#\mathbf{f})$$

$$\dfrac{\dfrac{\dfrac{\dfrac{}{a\#[a]P}\,([]\mathbf{a}) \qquad \dfrac{b\#P}{b\#[a]P}\,([]\mathbf{b})}{[b](b\ a)\cdot P=[a]P}\,(\mathbf{perm})}{[a]P=[b](b\ a)\cdot P}\,(\mathbf{symm}) \qquad \dfrac{\dfrac{\dfrac{b\#P}{P[a\mapsto b]=(b\ a)\cdot P}\,(\mathbf{ax}_{\mathbf{ren}\mapsto})}{(b\ a)\cdot P=P[a\mapsto b]}\,(\mathbf{symm})}{[b](b\ a)\cdot P=[b](P[a\mapsto b])}\,(\mathbf{cong}[])}{\dfrac{[a]P=[b](P[a\mapsto b])}{\forall[a]P=\forall[b](P[a\mapsto b])}\,(\mathbf{congf})}\,(\mathbf{tran})$$

$$\dfrac{\dfrac{\dfrac{}{a\#c}\,(\#\mathbf{ab})}{(\forall[a]P)[b\mapsto c]=\forall[a](P[b\mapsto c])}\,(\mathbf{ax}_{\forall\mapsto})}{\forall[a](P[b\mapsto c])=(\forall[a]P)[b\mapsto c]}\,(\mathbf{symm})$$

$$\Pi \quad = \quad \dfrac{\dfrac{\dfrac{}{a\#[a](P[b\mapsto a])}\,(\#[]\mathbf{a}) \qquad \dfrac{\dfrac{\dfrac{c\#P}{c\#[b]P}\,(\#\mathbf{b}) \qquad \dfrac{}{c\#a}\,(\#\mathbf{ab})}{c\#P[b\mapsto a]}\,(\#\mathbf{f})}{c\#[a](P[b\mapsto a])}\,(\#[]\mathbf{b})}{[c](((c\ a)\cdot P)[b\mapsto c])=[a](P[b\mapsto a])}\,(\mathbf{perm})}{[a](P[b\mapsto a])=[c](((c\ a)\cdot P)[b\mapsto c])}\,(\mathbf{symm})$$

$$\dfrac{\Pi \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{c\#P}{P[a\mapsto c]=(c\ a)\cdot P}\,(\mathbf{ax}_{\mathbf{ren}\mapsto})}{(c\ a)\cdot P=P[a\mapsto c]}\,(\mathbf{symm})}{[b]((c\ a)\cdot P)=[b](P[a\mapsto c])}\,(\mathbf{cong}[])}{((c\ a)\cdot P)[b\mapsto c]=P[a\mapsto c][b\mapsto c]}\,(\mathbf{congf}) \qquad \dfrac{\text{Lemma 4.11}}{P[a\mapsto c][b\mapsto c]=P[b\mapsto c][a\mapsto c]}}{((c\ a)\cdot P)[b\mapsto c]=P[b\mapsto c][a\mapsto c]}\,(\mathbf{tran})}{[c](((c\ a)\cdot P)[b\mapsto c])=[c](P[b\mapsto c][a\mapsto c])}\,(\mathbf{cong}[])}{\dfrac{\dfrac{[a](P[b\mapsto a])=[c](P[b\mapsto c][a\mapsto c])}{[c](P[b\mapsto c][a\mapsto c])=[a](P[b\mapsto a])}\,(\mathbf{symm})}{\forall[c](P[b\mapsto c][a\mapsto c])=\forall[a](P[b\mapsto a])}\,(\mathbf{congf})}\,(\mathbf{tran})$$

FIGURE 6. Derivations of side-conditions

- The *cut-elimination* property of first-order predicate logic is preserved by the extension to one-and-a-halfth-order logic.

5.1. **Equivariance and substitution.** Fix a freshness context Δ, a context and cocontext Φ and Ψ, and a derivation Π.

We extend notation for permutation and substitution actions t^π, $\pi \cdot t$ and $t\sigma$ to contexts Φ and cocontexts Ψ, writing Φ^π, $\pi \cdot \Phi$ and $\Phi\sigma$, and Ψ^π, $\pi \cdot \Psi$ and $\Psi\sigma$ for the result of applying the actions to the terms in the syntax of Φ and Ψ.

Write Π^π and $\pi \cdot \Pi$ for the derivation obtained from Π by translating each part according to the following table:

part of Π	in Π^π replaced by	in $\pi \cdot \Pi$ replaced by
sequent $\Phi \vdash_\Delta \Psi$	$\Phi^\pi \vdash_{\Delta^\pi} \Psi^\pi$	$\pi \cdot \Phi \vdash_\Delta \pi \cdot \Psi$
equality side-condition $\Delta \vdash_{\mathsf{SUB}} \phi = \psi$	$\Delta^\pi \vdash_{\mathsf{SUB}} \phi^\pi = \psi^\pi$	$\Delta \vdash_{\mathsf{SUB}} \pi \cdot \phi = \pi \cdot \psi$
freshness side-condition $\Delta \vdash a\#\Phi, \Psi$	$\Delta^\pi \vdash \pi(a)\#\Phi^\pi, \Psi^\pi$	$\Delta \vdash \pi(a)\#\pi \cdot \Phi, \pi \cdot \Psi$
side-condition $a \notin \Phi, \Psi, \Delta$	$\pi(a) \notin \Phi^\pi, \Psi^\pi, \Delta^\pi$	$a \notin \Phi, \Psi, \Delta$

So Π^π renames every part of Π, while $\pi \cdot \Pi$ renames everything except for freshness contexts and the side-condition of the (**Fr**) rule.

Theorem 5.1. *If Π is a valid derivation of $\Phi \vdash_\Delta \Psi$ then Π^π is a valid derivation of $\Phi^\pi \vdash_{\Delta^\pi} \Psi^\pi$.*

Call this property **meta-level equivariance**.

Proof. The statement

'Π is a valid derivation of $\Phi \vdash_\Delta \Psi$'

has four parameters and so by ZFA equivariance (Appendix A) is invariant under permuting atoms (in the values of) those parameters. The result follows.[2] □

Theorem 5.2. *If Π is a valid derivation of $\Phi \vdash_\Delta \Psi$ then $\pi \cdot \Pi$ is a valid derivation of $\pi \cdot \Phi \vdash_\Delta \pi \cdot \Psi$.*

Call this property **object-level equivariance**.

Proof. By induction on Π. Base cases (**Ax**) and ($\bot$**L**) are direct. We consider the inductive cases in turn:

(1) The case of ($\supset$**R**): Suppose $\Phi \vdash_\Delta \Psi, \phi \supset \psi$ is derived using ($\supset$**R**). Then there exists a derivation Π' of $\phi, \Phi \vdash_\Delta \Psi, \psi$, and by the inductive hypothesis $\pi \cdot \Pi'$ is a derivation of $\pi \cdot \phi,\ \pi \cdot \Phi \vdash_\Delta \pi \cdot \Psi,\ \pi \cdot \psi$. Then

$$\dfrac{\begin{array}{c}\vdots\ \pi \cdot \Pi' \\ \pi \cdot \phi,\ \pi \cdot \Phi \vdash_\Delta \pi \cdot \Psi,\ \pi \cdot \psi\end{array}}{\pi \cdot \Phi \vdash_\Delta \pi \cdot \Psi,\ \pi \cdot \phi \supset \pi \cdot \psi}\,(\supset\mathbf{R})$$

is the required derivation $\pi \cdot \Pi$ of $\pi \cdot \Phi \vdash_\Delta \pi \cdot \Psi,\ \pi \cdot \phi \supset \pi \cdot \psi$.

The cases ($\supset$**L**), ($\forall$**L**), ($\approx$**L**) and ($\approx$**R**) are similar.

(2) The case of ($\forall$**R**): Suppose $\Phi \vdash_\Delta \Psi,\ \forall[a]\psi$ is derived using ($\forall$**R**). Then $\Delta \vdash a\#\Phi, \Psi$ holds and Π' is a derivation of $\Phi \vdash_\Delta \Psi, \psi$. By Lemma 4.5 $\Delta \vdash a\#\pi \cdot \Phi,\ \pi \cdot \Psi$, and by inductive hypothesis $\pi \cdot \Pi'$ is a derivation of $\pi \cdot \Phi \vdash_\Delta \pi \cdot \Psi,\ \pi \cdot \psi$. We conclude that $\pi \cdot \Phi \vdash_\Delta \pi \cdot \Psi,\ \forall[a]\pi \cdot \psi$ is derivable by extending $\pi \cdot \Pi'$ with ($\forall$**R**), as required.

[2] A proof by induction on derivations is possible, but longer.

The cases (**StructL**), (**StructR**) and (**Cut**) are similar.

(3) The case of (**Fr**): Suppose $\Phi \vdash_{\Delta} \Psi$ is derived using (**Fr**). Then Π' is a derivation of $\Phi \vdash_{\Delta, a\#X_1,\dots,a\#X_n} \Psi$ where $a \notin \Phi, \Psi, \Delta$. By meta-level equivariance (Theorem 5.1) also $\Pi'^{(a'\ a)}$ is a derivation of $\Phi \vdash_{\Delta, a'\#X_1,\dots,a'\#X_n} \Psi$, where a' is chosen fresh (i.e. $a' \notin a, \Phi, \Psi, \Delta, \pi$).[3]

By ZFA equivariance (Appendix A) validity of the property

'Π' has the inductive hypothesis'

is itself invariant under permuting atoms. So

'$\Pi'^{(a'\ a)}$ has the inductive hypothesis'

is also valid.

By inductive hypothesis $\pi \cdot \Pi'^{(a'\ a)}$ is a derivation of $\pi \cdot \Phi \vdash_{\Delta, a'\#X_1,\dots,a'\#X_n} \pi \cdot \Psi$. Since $a' \notin \Phi, \Psi, \Delta, \pi$ we may deduce $\pi \cdot \Phi \vdash_{\Delta} \pi \cdot \Psi$ using (**Fr**), as required.

$\square$

Write $\Pi(\sigma, \Delta')$ for the **substitution action** on derivations. We inductively define it on the structure of Π as follows:

- If Π concludes with a rule (**R**) different from (**Fr**), it is of the form

$$\dfrac{\begin{array}{c}\vdots\,\Pi_1\\ \Phi_1 \vdash_{\Delta} \Psi_1\end{array} \quad \cdots \quad \begin{array}{c}\vdots\,\Pi_k\\ \Phi_k \vdash_{\Delta} \Psi_k\end{array}}{\Phi \vdash_{\Delta} \Psi}\,(\mathbf{R}) \quad (\mathit{cond})$$

where $k \in \{0,1,2\}$ and *cond* is $\Delta \vdash_{\mathsf{SUB}} \phi = \psi$, $\Delta \vdash a\#\Phi'$ or empty.

Then $\Pi(\sigma, \Delta')$ is

$$\dfrac{\begin{array}{c}\vdots\,\Pi_1(\sigma,\Delta')\\ \Phi_1\sigma \vdash_{\Delta'} \Psi_1\sigma\end{array} \quad \cdots \quad \begin{array}{c}\vdots\,\Pi_k(\sigma,\Delta')\\ \Phi_k\sigma \vdash_{\Delta'} \Psi_k\sigma\end{array}}{\Phi\sigma \vdash_{\Delta'} \Psi\sigma}\,(\mathbf{R}) \quad (\mathit{cond}')$$

where cond' is $\Delta' \vdash_{\mathsf{SUB}} \phi\sigma = \psi\sigma$, $\Delta' \vdash a\#\Phi'\sigma$ or empty, respectively.

- Otherwise, the derivation concludes in

$$\dfrac{\begin{array}{c}\vdots\,\Pi'\\ \Phi \vdash_{\Delta, a\#X_1,\dots,a\#X_n} \Psi\end{array}}{\Phi \vdash_{\Delta} \Psi}\,(\mathbf{Fr}) \quad (n \geq 1, a \notin \Phi, \Psi, \Delta).$$

Let $Y_1, \dots, Y_m$ be all unknowns mentioned in $\sigma(X_i)$, for $1 \leq i \leq n$, choose the atom a' to be fresh (i.e. $a' \notin a, \Phi, \Psi, \Delta, \Delta', \sigma$), and let $\Delta'' = \Delta', a'\#Y_1, \dots, a'\#Y_m$. Then:

  – If $m \geq 1$, then $\Pi(\sigma, \Delta')$ is

$$\dfrac{\begin{array}{c}\vdots\,\Pi'^{(a'\ a)}(\sigma,\Delta'')\\ \Phi\sigma \vdash_{\Delta''} \Psi\sigma\end{array}}{\Phi\sigma \vdash_{\Delta'} \Psi\sigma}\,(\mathbf{Fr}) \quad (m \geq 1, a' \notin \Phi\sigma, \Psi\sigma, \Delta'),$$

  – If $m = 0$, then $\Pi(\sigma, \Delta')$ is just $\Pi'^{(a'\ a)}(\sigma, \Delta'')$; no extra freshness assumptions need to be introduced. Note that $\Delta'' = \Delta'$ in this case.

[3] If already $a \notin \pi \cdot \Phi, \pi \cdot \Psi, \Delta$ then renaming a to a' in Π is not strictly necessary.

So σ is consistently applied throughout the formula contexts occurring in Π, Δ' replaces Δ, and (**Fr**) may generate slightly different freshness assumptions.

Theorem 5.3. *Suppose that $\Delta' \vdash \Delta\sigma$, and suppose that Π is a valid derivation of $\Phi \vdash_{\Delta} \Psi$. Then $\Pi(\sigma, \Delta')$ is a valid derivation of $\Phi\sigma \vdash_{\Delta'} \Psi\sigma$.*

Call this property **meta-level substitution**.

Proof. By induction on Π, similar to the proof of Theorem 5.2.

The cases ($\forall$**R**), (**StructL**), (**StructR**) and (**Cut**) use meta-level substitution on freshness and equality (Theorem 4.7).

We treat the case of (**Fr**) in more detail. Suppose $\Phi \vdash_{\Delta} \Psi$ is derived using (**Fr**). Then Π' is a derivation of $\Phi \vdash_{\Delta, a\#X_1,\dots,a\#X_n} \Psi$ where $a \not\in \Phi, \Psi, \Delta$.

Let $Y_1, \dots, Y_m$ be all the unknowns mentioned in $\sigma(X_i)$, $1 \leq i \leq n$, choose a' is fresh (i.e. $a' \not\in a, \Phi, \Psi, \Delta, \Delta', \sigma$), and let $\Delta'' = \Delta', a'\#Y_1, \dots, a'\#Y_m$.

By meta-level equivariance (Theorem 5.1) $\Pi'^{(a'\ a)}$ is a derivation of $\Phi \vdash_{\Delta, a'\#X_1,\dots,a'\#X_n} \Psi$, and by ZFA equivariance (Appendix A) we retain the inductive hypothesis for this derivation. We can easily verify that $\Delta'' \vdash (\Delta, a'\#X_1, \dots, a'\#X_n)\sigma$, so by the inductive hypothesis $\Pi'^{(a'\ a)}(\sigma, \Delta'')$ is a valid derivation of $\Phi\sigma \vdash_{\Delta''} \Psi\sigma$.

We proceed by case distinction on m:

- Suppose $m \geq 1$. Since $a' \not\in \Phi\sigma, \Psi\sigma, \Delta'$ we may extend $\Pi'^{(a'\ a)}(\sigma, \Delta'')$ with (**Fr**) to obtain our required derivation $\Pi(\sigma, \Delta')$ of $\Phi\sigma \vdash_{\Delta'} \Psi\sigma$.
- Suppose $m = 0$. By definition $\Pi(\sigma, \Delta')$ is $\Pi'^{(a'\ a)}(\sigma, \Delta'')$. Since $\Delta'' = \Delta'$, it is a derivation of $\Phi\sigma \vdash_{\Delta'} \Psi\sigma$, as required.

□

A useful corollary of Theorem 5.3 is the following:

Corollary 5.4. *If $\Phi \vdash_{\emptyset} \Psi$ is derivable for closed Φ and Ψ, then there is a derivation of $\Phi \vdash_{\emptyset} \Psi$ that does not mention unknowns.*

Proof. Suppose Π is a derivation of $\Phi \vdash_{\emptyset} \Psi$, which possibly mentions unknowns. Let σ be the substitution that maps all unknowns in the derivation to closed terms as follows:

- each unknown P of sort $\mathbb{F}$ is mapped to $\bot$;
- each unknown T of sort $\mathbb{T}$ is mapped to a', where a' is an atom that does not occur anywhere in the derivation.

By meta-level substitution (Theorem 5.3), $\Pi(\sigma, \emptyset)$ is a valid derivation of $\Phi \vdash_{\emptyset} \Psi$. This derivation does not mention unknowns, as can be verified by an easy induction on the structure of the definition of meta-level substitution. □

5.2. **Cut-elimination.** Call the **depth** of a derivation the greatest number of derivation steps not counting rules (**Fr**), (**StructL**) and (**StructR**) between its conclusion and its leaves, over all paths. We do not count nominal algebra derivations of freshnesses and equalities that occur as side-conditions. For example, the last two derivations of Figure 3 both have depth 4.

The following results are not normally problematic but we have internalised both α-equivalence *and* being fresh — so renaming and freshening must be represented in the derivation.

Lemma 5.5. *If $\Phi \vdash_{\Delta} \Psi$ and $\Delta \subseteq \Delta'$ then $\Phi \vdash_{\Delta'} \Psi$. The derivation has the same depth as the original one, and no more instances of cut.*

Call this property **freshness weakening**.

Proof. By straightforward induction on the structure of the derivation. The cases $(\forall\mathbf{R})$, $(\mathbf{StructL})$, $(\mathbf{StructR})$ and $(\mathbf{Cut})$ use Corollary 4.8 to weaken the freshness context of the side-conditions. The case of $(\mathbf{Fr})$ uses ZFA equivariance. $\square$

It is also easy to see that Π^π, $\pi \cdot \Pi$ and $\Pi(\sigma, \Delta')$ preserve the depth and number of instances of Π.

Write $\mathcal{U}(\mathit{Stuff})$ for the unknowns $X, Y, Z, \dots$ mentioned in the *Stuff*.

Lemma 5.6. *If $a \notin u$ and $a\#\mathcal{U}(u) \subseteq \Delta$ then $\Delta \vdash a\#u$.*

Lemma 5.7. *If $\Phi \vdash_\Delta \Psi$ and $\Phi \subseteq \Phi'$ and $\Psi \subseteq \Psi'$ then $\Phi' \vdash_\Delta \Psi'$. The new derivation has the same depth as the original one, and no more instances of cut.*

Call this property **formula weakening**.

Proof. We work by strong induction on the *pair* of the depth of the derivation and its structure, lexicographically ordered. The conditions on preserving depth and number of cuts can easily be verified from the structure of the reasoning which follows, and we do not mention them further.

(1) The case of $(\mathbf{StructL})$: Suppose $\phi, \Phi \vdash_\Delta \Psi$ is derived using $(\mathbf{StructL})$, and assume the inductive hypothesis on all strictly lesser derivations. So $\phi', \Phi \vdash_\Delta \Psi$ and $\Delta \vdash_{\mathsf{SUB}} \phi' = \phi$ are derivable for some ϕ'. This derivation has the same depth as, and a lesser structure than that of $\phi, \Phi \vdash_\Delta \Psi$, so we may use the inductive hypothesis to derive $\phi', \Phi' \vdash_\Delta \Psi'$. By $(\mathbf{StructL})$ we obtain $\phi, \Phi' \vdash_\Delta \Psi'$ as required.

(2) The case of $(\forall\mathbf{R})$: Suppose $\Phi \vdash_\Delta \Psi, \forall[a]\psi$ is derived using $(\forall\mathbf{R})$ and suppose the inductive hypothesis of all strictly lesser derivations.

By assumption $\Phi \vdash_\Delta \Psi, \psi$ has a derivation of strictly lesser depth, and also $\Delta \vdash a\#\Phi, \Psi$ holds.

Choose a' fresh (i.e. $a' \notin a, \psi, \Phi', \Psi', \Delta$) and $\Delta' = \Delta, a'\#\mathcal{U}(\Phi', \Psi', \Delta, \psi)$. Then $\Delta' \vdash a\#\Phi, \Psi$ by Corollary 4.8, and $\Phi \vdash_{\Delta'} \Psi,\ \psi$ by freshness weakening (Lemma 5.5).[4] Then by object-level equivariance (Theorem 5.2) also $(a'\ a) \cdot \Phi \vdash_{\Delta'} (a'\ a) \cdot \Psi,\ (a'\ a) \cdot \psi$. Using $(\mathbf{perm})$, by means of $(\mathbf{StructL})$ and $(\mathbf{StructR})$, we obtain $\Phi \vdash_{\Delta'} \Psi,\ (a'\ a) \cdot \psi$. By inductive hypothesis (the derivation still has strictly lesser depth) there exists a derivation Π of $\Phi' \vdash_{\Delta'} \Psi',\ (a'\ a) \cdot \psi$. By Lemma 5.6 also $\Delta' \vdash a'\#\Phi', \Psi'$, and by simple calculations we observe $\Delta' \vdash_{\mathsf{SUB}} \forall[a'](a'\ a) \cdot \psi = \forall[a]\psi$ (we use $(\mathbf{perm})$, and the freshness information we have assumed of a').

Now we can conclude $\Phi' \vdash_\Delta \Psi',\ \forall[a]\psi$ as follows:

$$\dfrac{\dfrac{\dfrac{\begin{matrix}\vdots\,\Pi\\ \Phi' \vdash_{\Delta'} \Psi',\ (a'\ a) \cdot \psi\end{matrix}}{\Phi' \vdash_{\Delta'} \Psi',\ \forall[a'](a'\ a) \cdot \psi}\,(\forall\mathbf{R})}{\Phi' \vdash_{\Delta'} \Psi',\ \forall[a]\psi}\,(\mathbf{StructR})}{\Phi' \vdash_{\Delta} \Psi',\ \forall[a]\psi}\,(\mathbf{Fr})$$

---

[4] It appears convenient to prove freshness weakening first separately; we do not want to weaken Φ and Ψ to Φ' and Ψ' until we have renamed a to a', in a moment.

$$\frac{\Phi \vdash_{\vartriangle} \Psi,\ \phi}{\neg\phi,\ \Phi \vdash_{\vartriangle} \Psi}\,(\neg\mathbf{L}) \qquad \frac{}{\Phi \vdash_{\vartriangle} \Psi, \top}\,(\top\mathbf{R}) \qquad \frac{\psi,\ \Phi \vdash_{\vartriangle} \Psi}{\Phi \vdash_{\vartriangle} \Psi,\ \neg\psi}\,(\neg\mathbf{R})$$

$$\frac{\phi,\ \phi',\ \Phi \vdash_{\vartriangle} \Psi}{\phi \wedge \phi',\ \Phi \vdash_{\vartriangle} \Psi}\,(\wedge\mathbf{L}) \qquad \frac{\Phi \vdash_{\vartriangle} \Psi,\ \psi \qquad \Phi \vdash_{\vartriangle} \Psi,\ \psi'}{\Phi \vdash_{\vartriangle} \Psi,\ \psi \wedge \psi'}\,(\wedge\mathbf{R})$$

$$\frac{\phi,\ \Phi \vdash_{\vartriangle} \Psi \qquad \phi',\ \Phi \vdash_{\vartriangle} \Psi}{\phi \vee \phi',\ \Phi \vdash_{\vartriangle} \Psi}\,(\vee\mathbf{L}) \qquad \frac{\Phi \vdash_{\vartriangle} \Psi,\ \psi,\ \psi'}{\Phi \vdash_{\vartriangle} \Psi,\ \psi \vee \psi'}\,(\vee\mathbf{R})$$

$$\frac{\Phi \vdash \Psi,\ \phi,\ \phi' \quad \phi,\ \phi',\ \Phi \vdash_{\vartriangle} \Psi}{\phi \Leftrightarrow \phi',\ \Phi \vdash_{\vartriangle} \Psi}\,(\Leftrightarrow\mathbf{L}) \qquad \frac{\psi,\ \Phi \vdash_{\vartriangle} \Psi,\ \psi' \quad \psi',\ \Phi \vdash_{\vartriangle} \Psi,\ \psi}{\Phi \vdash \Psi,\ \psi \Leftrightarrow \psi'}\,(\Leftrightarrow\mathbf{R})$$

$$\frac{\phi,\, \Phi \vdash_{\vartriangle} \Psi}{\exists[a]\phi,\, \Phi \vdash_{\vartriangle} \Psi}\,(\exists\mathbf{L}) \quad (\Delta \vdash a\#\Phi, \Psi) \qquad \frac{\Phi \vdash_{\vartriangle} \Psi,\, \phi[a \mapsto t]}{\Phi \vdash_{\vartriangle} \Psi,\, \exists[a]\phi}\,(\exists\mathbf{R})$$

FIGURE 7. Admissible sequent rules for one-and-a-halfth-order logic

(3) The case of (**Fr**): Suppose $\Phi \vdash_{\Delta, a\#X_1,\dots,a\#X_n} \Psi$ where $a \notin \Phi, \Psi, \Delta$. We use ZFA equivariance (Appendix A) to rename a to some $a' \notin \Phi', \Psi', \Delta$ in the whole derivation to obtain one of $\Phi \vdash_{\Delta, a'\#X_1,\dots,a'\#X_n} \Psi$. We can now apply the inductive hypothesis (which, as discussed above, by equivariance is preserved by the permutative renaming) to weaken to Φ' and Ψ', and finish off with (**Fr**).

The other cases are easy or similar. □

Recall the sugar from Subsection 2.1.

Corollary 5.8. *The rules of Figure 7 are all admissible.*

Proof. We consider just the case of $(\neg\mathbf{R})$. Suppose we have derived $\psi,\ \Phi \vdash_{\vartriangle} \Psi$. Then by Lemma 5.7 there also exists a derivation of $\psi,\ \Phi \vdash_{\vartriangle} \Psi,\ \bot$. Extending that derivation with $(\supset\mathbf{R})$ we obtain a derivation of $\Phi \vdash_{\vartriangle} \Psi,\ \neg\psi$ as required.

The cases $(\wedge\mathbf{R})$, $(\vee\mathbf{L})$, $(\Leftrightarrow\mathbf{L})$, $(\Leftrightarrow\mathbf{R})$ and $(\exists\mathbf{L})$, are similar. Remaining cases are by directly extending derivations. In the case of $(\exists\mathbf{R})$, we use (**StructL**) to replace the $(\neg\phi)[a \mapsto t]$ by $\neg(\phi[a \mapsto t])$. □

Some of the above admissible rules will turn out useful later (Subsection 6.3).

Write $\Phi[b \mapsto u]$ for the elementwise application of the substitution to the elements of formula context Φ.

Lemma 5.9. *If $\Phi \vdash_{\vartriangle} \Psi$ then $\Phi[b \mapsto u] \vdash_{\vartriangle} \Psi[b \mapsto u]$. The depth of the derivation does not increase, and neither does the number of cuts it contains.*

Call this property **object-level substitution**.

Proof. Analogous to the proof of Lemma 5.7. □

Lemma 5.10. (**Fr**) *may be commuted down through all other rules. The transformations involved do not increase the depth of a derivation or its number of cuts.*

Proof. We consider the various possibilities in turn:

- Suppose (**Fr**) is followed by ($\supset$**L**) as follows:

$$\dfrac{\dfrac{\dfrac{\vdots\,\Pi_1}{\Phi \vdash_{\Delta,a\#X_1,\ldots,a\#X_n} \Psi,\, \phi}}{\Phi \vdash_{\Delta} \Psi,\, \phi}\,(\mathbf{Fr}) \qquad \dfrac{\vdots\,\Pi_2}{\psi,\, \Phi \vdash_{\Delta} \Psi}}{\phi \supset \psi,\, \Phi \vdash_{\Delta} \Psi}\,(\supset\mathbf{L})$$

  Here $a \notin \Phi, \Delta, \Psi, \phi$.

  Suppose we are unlucky and a is mentioned in ψ. Choose a fresh atom a' (i.e. $a' \notin \Phi, \Delta, a, \Psi, \phi, \psi$). By meta-level equivariance (Theorem 5.1), $\Pi_1^{(a'\ a)}$ is a valid derivation of $\Phi \vdash_{\Delta,a'\#X_1,\ldots,a'\#X_n} \Psi,\, \phi$. Also, by freshness weakening (Lemma 5.5), there is a derivation Π_2' of $\psi,\, \Phi \vdash_{\Delta,a'\#X_1,\ldots,a'\#X_n} \Psi$.

  We can now put our derivation together:

$$\dfrac{\dfrac{\dfrac{\vdots\,\Pi_1^{(a'\ a)}}{\Phi \vdash_{\Delta,a'\#X_1,\ldots,a'\#X_n} \Psi,\, \phi} \qquad \dfrac{\vdots\,\Pi_2'}{\psi,\, \Phi \vdash_{\Delta,a'\#X_1,\ldots,a'\#X_n} \Psi}}{\phi \supset \psi,\, \Phi \vdash_{\Delta,a'\#X_1,\ldots,a'\#X_n} \Psi}\,(\supset\mathbf{L})}{\phi \supset \psi,\, \Phi \vdash_{\Delta} \Psi}\,(\mathbf{Fr})$$

  The new derivation preserves the depth and number of cuts of the original derivation, since $\Pi_1^{(a'\ a)}$ does so by definition and Π_2' does so by Lemma 5.5.
- All other cases are similar or simpler. For example if (**Fr**) is followed by ($\approx$**L**) it is not immediate that we may swap the derivation rules round, since perhaps t in ($\approx$**L**) mentions a and t' does not. As in the previous case we may rename atoms in the derivation and then commute.

□

Theorem 5.11 (Cut-elimination)**.** *If $\Phi \vdash_{\Delta} \Psi$ is derivable in the sequent calculus for one-and-a-halfth-order logic, then there exists a derivation of $\Phi \vdash_{\Delta} \Psi$ which does not mention* (**Cut**)*.*

Proof. The commutation cases and essential cases are standard [23, 39]; we use Lemma 5.7 for the essential case for $\supset$; the non-standard case of (**Fr**) is handled by Lemma 5.10. The essential case for $\forall$ is handled by Lemma 5.9. □

Corollary 5.12. *The sequent calculus of one-and-a-halfth-order logic is* ***consistent****, i.e. $\vdash_{\Delta}$ can never be derived.*

Proof. By contradiction. Suppose $\vdash_{\Delta}$ is derivable, then by Theorem 5.11 a cut-free derivation exists. Let Π be the shortest derivation of $\vdash_{\Delta}$ for all possible Δ. We check through all possible derivation rules and see by their syntax-directed nature that the derivation must conclude in (**Fr**). But then we have a shorter derivation of some $\vdash_{\Delta'}$, which is a contradiction. □

6. An equational axiomatisation of one-and-a-halfth-order logic

6.1. **Theory** FOL**.** We now give an axiomatic presentation of one-and-a-halfth-order logic in nominal algebra from Section 4.

Definition 6.1. *Let theory* FOL *be given by the axioms of theory* SUB *(Figure 5) plus the axioms of Figure 8.*

$$
\begin{array}{rll}
(\mathbf{MP}) & \top \supset P & = P \\
(\mathbf{M}) & ((((P \supset Q) \supset (\neg R \supset \neg S)) \supset R) \supset T) & \\
 & \qquad\qquad \supset ((T \supset P) \supset (S \supset P)) & = \top \\
(\mathbf{Q1}) & \forall[a]P \supset P[a \mapsto T] & = \top \\
(\mathbf{Q2}) & \forall[a](P \wedge Q) \Leftrightarrow \forall[a]P \wedge \forall[a]Q & = \top \\
(\mathbf{Q3}) & a\#P \ \vdash\ \forall[a](P \supset Q) \Leftrightarrow P \supset \forall[a]Q & = \top \\
 & & \\
(\mathbf{E1}) & U \approx T \wedge P[a \mapsto T] \supset P[a \mapsto U] & = \top \\
(\mathbf{E2}) & T \approx T & = \top
\end{array}
$$

FIGURE 8. Axioms of FOL

We now discuss the axioms in Figure 8:

- We read (**MP**) as 'Modus Ponens' and it expresses the principle 'if P is true and P implies Q is true, then Q is true'. In the literature this principle is also called *detachment*.

  In (**M**) recall that $\neg\phi$ is sugar for $\phi \supset \bot$. This axiom from [32], along with Modus Ponens, is sufficient to derive all rules of classical propositional logic — otherwise known as *boolean logic* [7].

  (**M**) stands for Meredith, the author of this axiom for classical predicate logic, which he discovered in the 1940s. Machine-checked verification of this fact is online [33]. Succinct axioms for propositional logic continue to provide innocent fun in some circles [31]; a good survey is here [22].
- Axioms (**Q1**) – (**Q3**) add quantifiers, we call them **quantifier axioms**; (**Q3**) exploits freshness conditions.

  These axioms appear in the literature [12, page 5 (2)]. What is new here is that our axioms are *not* axiom-schemes; they are *individual axioms* (three, to be precise). We do this using the unknowns of nominal terms, and abstractions and freshness conditions to express capture-avoidance conditions which usually must be expressed — at the meta-level — for every instance. Note how these axioms are faithful to the usual syntactic form of the axiom-schemes found in the literature.
- Axioms (**E1**) and (**E2**) add object-level equality, we call them **equational axioms**.

  Again, we are able to represent by two axioms what might otherwise be two infinite axiom-schemes.

We now make the connection between the axioms in Figure 8 in the context of nominal algebra, and the sequent rules in Figure 2.

6.2. **Sequent derivability implies FOL derivability.** Let **classical propositional logic** be the entailment relation obtained by considering rules (**Ax**), ($\bot$**L**), ($\supset$**L**), and ($\supset$**R**) from Figure 2 and removing Δ.

Theorem 6.2. *If $\phi \Leftrightarrow \psi$ is derivable in classical propositional logic, then $\Delta \vdash_{\mathsf{FOL}} \phi = \psi$ in the nominal algebra theory* FOL.

Proof. By machine-checked proofs online [33], (**MP**) and (**M**) suffice to derive all the logical identities of classical propositional logic. □

Corollary 6.3. *The following equalities are all derivable in* FOL*:*

$$\begin{array}{ll}
\Delta \vdash_{\mathsf{FOL}} \phi \vee (\psi \vee \xi) = (\phi \vee \psi) \vee \xi & \Delta \vdash_{\mathsf{FOL}} \phi \wedge (\psi \wedge \xi) = (\phi \wedge \psi) \wedge \xi \\
\Delta \vdash_{\mathsf{FOL}} \phi \vee \psi = \psi \vee \phi & \Delta \vdash_{\mathsf{FOL}} \phi \wedge \psi = \psi \wedge \phi \\
\Delta \vdash_{\mathsf{FOL}} \phi \vee (\psi \wedge \phi) = \phi & \Delta \vdash_{\mathsf{FOL}} \phi \wedge (\psi \vee \phi) = \phi \\
\Delta \vdash_{\mathsf{FOL}} \phi \vee (\psi \wedge \xi) = (\phi \vee \psi) \wedge (\phi \vee \xi) & \Delta \vdash_{\mathsf{FOL}} \phi \wedge (\psi \vee \xi) = (\phi \wedge \psi) \vee (\phi \wedge \xi) \\
\Delta \vdash_{\mathsf{FOL}} \phi \vee \neg\phi = \top & \Delta \vdash_{\mathsf{FOL}} \phi \wedge \neg\phi = \bot
\end{array}$$

Proof. The reader will recognise these as the equalities of boolean algebra. It is known that equality in boolean algebra characterises precisely logical equivalence in classical propositional logic [7]. By Theorem 6.2 the equality of FOL includes equalities between all formulae that are provably logically equivalent in classical propositional logic (it suffices to use (**MP**) and (**M**)). The result follows. □

We shall say we work by **elementary calculations (in propositional logic)** when we use Corollary 6.3 to transform formulae according to standard identities in classical propositional logic.

Lemma 6.4. $\vdash_{\mathsf{FOL}} \forall[a]\bot = \bot$ *is derivable.*

Proof. We derive $\neg\forall[a]\bot = \top$ as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{a\#\bot}\,(\#\mathbf{f})}{\bot[a\mapsto a]=\bot}\,(\mathbf{ax}_{\#\mapsto})}{\bot=\bot[a\mapsto a]}\,(\mathbf{symm})}{(\forall[a]\bot)\supset\bot=(\forall[a]\bot)\supset\bot[a\mapsto a]}\,(\mathbf{congf}) \qquad \dfrac{}{(\forall[a]\bot)\supset\bot[a\mapsto a]=\top}\,(\mathbf{ax_{Q1}})}{(\forall[a]\bot)\supset\bot=\top}\,(\mathbf{tran})$$

It follows that $\vdash_{\mathsf{FOL}} \neg\neg\forall[a]\bot = \neg\top$ and so by elementary calculations in propositional logic the result follows. □

An informal reading of Lemma 6.4 is that any semantics for $\mathbb{T}$ in FOL should be *non-empty*, for if $\mathbb{T}$ were empty then (intuitively) $\forall[a]\bot = \top$, so $\forall[a]\bot = \bot$ should not be derivable.

$\mathbb{T}$ is non-empty because it is populated by atoms (in the derivation above, we use the fact that it is populated by a). Thanks to the substitution action, atoms behave like 'object-level variable symbols'. Normally sorts of terms are populated by variable symbols, but this feature of the syntax does not show in the semantics, and that affects the notion of derivability: $(\forall x.\bot) \Leftrightarrow \top$ may be derivable. We see that in one-and-a-halfth-order logic terms are populated by unknowns *and* atoms. Although atoms represent variable symbols, the derivation above suggests that any semantics for one-and-a-halfth-order logic must differ from a 'standard' semantics, and give atoms denotational reality. One such semantics is given by a standard semantics for nominal algebra [18] in nominal sets [21]; developing other denotations-containing-variables is very much of current research interest.

We need some meta-level properties.

Lemma 6.5. *For any formulae* ϕ, ψ*:*

(1) $\Delta \vdash_{\mathsf{FOL}} \phi \wedge \psi = \top$ *if and only if* $\Delta \vdash_{\mathsf{FOL}} \phi = \top$ *and* $\Delta \vdash_{\mathsf{FOL}} \psi = \top$.

(2) $\Delta \vdash_{\mathsf{FOL}} \phi \Leftrightarrow \psi = \top$ *if and only if* $\Delta \vdash_{\mathsf{FOL}} \phi = \psi$.

Proof. By elementary calculations in propositional logic. □

We also need some *scope extrusion* properties.

Lemma 6.6. *The following are derivable:*

(1) $a\#P \vdash_{\mathsf{FOL}} \forall[a](P \supset Q) = P \supset \forall[a]Q$.
(2) $a\#P \vdash_{\mathsf{FOL}} \forall[a](\neg P) = \neg P$.
(3) $a\#P \vdash_{\mathsf{FOL}} \forall[a]P = P$.
(4) $a\#P \vdash_{\mathsf{FOL}} \forall[a](P \vee Q) = P \vee \forall[a]Q$.
(5) $a\#P \vdash_{\mathsf{FOL}} \forall[a](P \wedge Q) = P \wedge \forall[a]Q$.

Proof. By part 2 of Lemma 6.5, the first part is an instance or axiom (**Q3**). Since $\neg P \equiv P \supset \bot$, the second part follows by part 1 of this lemma and Lemma 6.4. The third and fourth part are corollaries of the first two parts, since $\vdash_{\mathsf{FOL}} P = \neg\neg P$ and $P \vee Q \equiv \neg P \supset Q$. The last part is a corollary of axiom (**Q2**) and part 3 of this lemma. ☐

We are now in a position to derive the following 'sequent-like' properties of our Hilbert-style FOL:

Lemma 6.7. *For all formulae $\phi, \phi', \psi, \psi', \theta, \varepsilon$, atoms a, terms $t, t' : \mathbb{T}$, and unknowns $X_1, \ldots, X_n$:*

(1) $\Delta \vdash_{\mathsf{FOL}} \phi \wedge \theta \supset \varepsilon \vee \phi \;=\; \top$
(2) $\Delta \vdash_{\mathsf{FOL}} \bot \wedge \theta \supset \varepsilon \;=\; \top$
(3) *if* $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee \phi \;=\; \top$ *and* $\Delta \vdash_{\mathsf{FOL}} \psi \wedge \theta \supset \varepsilon \;=\; \top$
then $\Delta \vdash_{\mathsf{FOL}} (\phi \supset \psi) \wedge \theta \supset \varepsilon \;=\; \top$
(4) *if* $\Delta \vdash_{\mathsf{FOL}} \phi \wedge \theta \supset \varepsilon \vee \psi \;=\; \top$
then $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee (\phi \supset \psi) \;=\; \top$
(5) *if* $\Delta \vdash_{\mathsf{FOL}} \phi[a \mapsto t] \wedge \theta \supset \varepsilon \;=\; \top$
then $\Delta \vdash_{\mathsf{FOL}} \forall[a]\phi \wedge \theta \supset \varepsilon \;=\; \top$
(6) *if* $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee \psi \;=\; \top$ *and* $\Delta \vdash a\#\theta, \varepsilon$
then $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee \forall[a]\psi \;=\; \top$
(7) *if* $\Delta \vdash_{\mathsf{FOL}} \phi[a \mapsto t'] \wedge \theta \supset \varepsilon \;=\; \top$
then $\Delta \vdash_{\mathsf{FOL}} (t' \approx t) \wedge \phi[a \mapsto t] \wedge \theta \supset \varepsilon \;=\; \top$
(8) $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee (t \approx t) \;=\; \top$
(9) *if* $\Delta \vdash_{\mathsf{FOL}} \phi' \wedge \theta \supset \varepsilon \;=\; \top$ *and* $\Delta \vdash_{\mathsf{SUB}} \phi' = \phi$
then $\Delta \vdash_{\mathsf{FOL}} \phi \wedge \theta \supset \varepsilon \;=\; \top$
(10) *if* $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee \psi' \;=\; \top$ *and* $\Delta \vdash_{\mathsf{SUB}} \psi' = \psi$
then $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee \psi \;=\; \top$
(11) *if* $\Delta, a\#X_1, \ldots, a\#X_n \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \;=\; \top$ *and* $a \notin \theta, \varepsilon, \Delta$
then $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \;=\; \top$
(12) *if* $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee \phi \;=\; \top,\;\; \Delta \vdash_{\mathsf{FOL}} \phi' \wedge \theta \supset \varepsilon \;=\; \top$
and $\Delta \vdash_{\mathsf{SUB}} \phi = \phi'$ *then* $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \;=\; \top$

Proof. The first four parts follow by elementary calculations in propositional logic.

For part 5, suppose that $\Delta \vdash_{\mathsf{FOL}} \phi[a \mapsto t] \wedge \theta \supset \varepsilon \;=\; \top$. By axiom (**Q1**) we know $\Delta \vdash_{\mathsf{FOL}} \forall[a]\phi \supset \phi[a \mapsto t] = \top$. By Lemma 6.5 we obtain

$$\Delta \vdash_{\mathsf{FOL}} (\forall[a]\phi \supset \phi[a \mapsto t]) \wedge (\phi[a \mapsto t] \wedge \theta \supset \varepsilon) \;=\; \top.$$

Using further elementary calculations we conclude $\Delta \vdash_{\mathsf{FOL}} \forall[a]\phi \wedge \theta \supset \varepsilon \ = \top$ as required.

For part 6, suppose that $\Delta \vdash a\#\theta, \varepsilon$ and $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee \psi \ = \top$. Using (**cong**[]) and (**congf**) we obtain $\Delta \vdash_{\mathsf{FOL}} \forall[a](\theta \supset \varepsilon \vee \psi) \ = \forall[a]\top$. We use Lemma 6.6 and (**tran**) to conclude $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee \forall[a]\psi \ = \top$.

Parts 7 and 8 use axioms (**E1**) and (**E2**), respectively. Parts 9 and 10 follow by (**tran**) and (**congf**), since $\Delta \vdash_{\mathsf{SUB}} \phi' = \phi$ implies $\Delta \vdash_{\mathsf{FOL}} \phi' = \phi$. Part 11 is immediate using (**fr**).

Part 12: Since $\Delta \vdash_{\mathsf{SUB}} \phi = \phi'$ implies $\Delta \vdash_{\mathsf{FOL}} \phi = \phi'$, we may suppose

$$\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon \vee \phi = \top \qquad \text{and} \qquad \Delta \vdash_{\mathsf{FOL}} \phi \wedge \theta \supset \varepsilon = \top.$$

By Lemma 6.5 we obtain $\Delta \vdash_{\mathsf{FOL}} (\theta \supset \varepsilon \vee \phi) \wedge (\phi \wedge \theta \supset \varepsilon) = \top$. By elementary calculations in propositional logic $\Delta \vdash_{\mathsf{FOL}} \phi \wedge \theta \supset \varepsilon = \theta \supset \varepsilon \vee \neg\phi$, so we conclude $\Delta \vdash_{\mathsf{FOL}} (\theta \supset \varepsilon \vee \phi) \wedge (\theta \supset \varepsilon \vee \neg\phi) = \top$. By further calculations we reduce this to $\Delta \vdash_{\mathsf{FOL}} \theta \supset \varepsilon = \top$. $\square$

For any one-and-a-halfth-order logic context $\Phi = \{\phi_1, \dots, \phi_n\}$, define its **conjunctive form** $\Phi^\wedge$ to be

- $\top$ when $n = 0$, and
- $\phi_1 \wedge \cdots \wedge \phi_n$ when $n > 0$.

Analogously define the **disjunctive form** $\Phi^\vee$ to be

- $\bot$ when $n = 0$, and
- $\phi_1 \vee \cdots \vee \phi_n$ when $n > 0$.

The order of the ϕ_i is irrelevant; we promise never to do anything such that it matters.

Theorem 6.8. *If $\Phi \vdash_\Delta \Psi$ is derivable in one-and-a-halfth-order logic then $\Delta \vdash_{\mathsf{FOL}} \Phi^\wedge \supset \Psi^\vee = \top$ in the nominal algebra theory* FOL*.*

Proof. We inductively transform a one-and-a-halfth-order logic derivation of $\Phi \vdash_\Delta \Psi$ into a nominal algebra derivation of $\Phi^\wedge \supset \Psi^\vee = \top$ from Δ in theory FOL.

For every rule (**R**), the derivation has the following format:

$$\frac{\Pi_1 \quad \cdots \quad \Pi_k}{\Phi \vdash_\Delta \Psi}\,(\mathbf{R}) \quad (\mathit{cond})$$

Here $k \in \{0, 1, 2\}$, Π_i are derivations of $\Phi_i \vdash_{\Delta_i} \Psi_i$, $1 \leq i \leq k$, and *cond* is a (possibly trivial) side-condition (the non-trivial cases are (**Fr**), (**Cut**), (**StructL**), (**StructR**), and ($\forall$**R**)).

So $\Phi_i \vdash_{\Delta_i} \Psi_i$ are derivable, then by inductive hypothesis $\Delta_i \vdash_{\mathsf{FOL}} {\Phi_i}^\wedge \supset {\Psi_i}^\vee = \top$ holds. We use this together with *cond* to prove $\Delta \vdash_{\mathsf{FOL}} \Phi^\wedge \supset \Psi^\vee \ = \top$. For each inference rule (**R**), this is an instance of a part of Lemma 6.7.

For example, if (**R**) is (**Cut**) then $\Delta \vdash_{\mathsf{FOL}} \Phi^\wedge \supset \Psi^\vee \ = \top$ should follow from

$$\Delta \vdash_{\mathsf{FOL}} \Phi^\wedge \supset \Psi^\vee \vee \phi \ = \top, \ \ \Delta \vdash_{\mathsf{FOL}} \phi' \wedge \Phi^\wedge \supset \Psi^\vee \ = \top, \ \text{ and } \ \Delta \vdash_{\mathsf{SUB}} \phi = \phi'.$$

This is an instance of part 12 of Lemma 6.7, using $\theta \equiv \Phi^\wedge$ and $\varepsilon \equiv \Psi^\vee$.

And if (**R**) is ($\forall$**R**) then $\Delta \vdash_{\mathsf{FOL}} \Phi^\wedge \supset \Psi^\vee \vee \forall[a]\psi = \top$ should follow from

$$\Delta \vdash_{\mathsf{FOL}} \Phi^\wedge \supset \Psi^\vee \vee \psi = \top \ \text{ and } \ \Delta \vdash a\#\Phi^\wedge, \Psi^\vee.$$

This is an instance of part 6, again using $\theta \equiv \Phi^\wedge$ and $\varepsilon \equiv \Psi^\vee$. $\square$

6.3. **FOL derivability implies sequent derivability.** We now show that the sequent presentation of one-and-a-halfth-order logic (Figure 2) can mimic the axioms of nominal algebra theory FOL (Figures 4 and 8). In the proofs in this subsection we will use the admissible sequent rules from Corollary 5.8 (Figure 7) without comment.

Lemma 6.9. *For all formulae $\phi, \psi, \rho, \theta, \varepsilon$, terms $t, u : \mathbb{T}$ and atoms a, the following are derivable in one-and-a-halfth-order logic:*

(1) $\vdash (\top \supset \phi) \Leftrightarrow \phi$
(2) $\vdash ((((\phi \supset \psi) \supset (\neg\rho \supset \neg\theta)) \supset \rho) \supset \varepsilon) \supset ((\varepsilon \supset \phi) \supset (\theta \supset \phi))$
(3) $\vdash \forall[a]\phi \supset \phi[a \mapsto t]$
(4) $\vdash \forall[a](\phi \wedge \psi) \Leftrightarrow \forall[a]\phi \wedge \forall[a]\psi$
(5) *if* $\Delta \vdash a\#\phi$ *then* $\vdash_{\vartriangle} \forall[a](\phi \supset \psi) \Leftrightarrow \phi \supset \forall[a]\psi$
(6) $\vdash u \approx t \wedge \phi[a \mapsto t] \supset \phi[a \mapsto u]$
(7) $\vdash t \approx t$

Proof. We give details of parts 5 and 6. The derivation of part 6 is completely syntax directed:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{\phi[a \mapsto u] \vdash \phi[a \mapsto u]}\,(\mathbf{Axiom})}{u \approx t, \phi[a \mapsto t] \vdash \phi[a \mapsto u]}\,(\approx\mathbf{L})}{u \approx t \wedge \phi[a \mapsto t] \vdash \phi[a \mapsto u]}\,(\wedge\mathbf{L})}{\vdash u \approx t \wedge \phi[a \mapsto t] \supset \phi[a \mapsto u]}\,(\supset\mathbf{R})$$

For part 5, we assume $\Delta \vdash a\#\phi$. By $(\Leftrightarrow\mathbf{R})$, it suffices to derive

(a) $\phi, \forall[a](\phi \supset \psi) \vdash_{\vartriangle} \forall[a]\psi$, and
(b) $\phi \supset \forall[a]\psi \vdash_{\vartriangle} \forall[a](\phi \supset \psi)$.

We show that (a) is derivable, showing derivability of (b) follows similar lines:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\phi \vdash_{\vartriangle} \psi, \phi}\,(\mathbf{Axiom}) \quad \dfrac{}{\phi, \psi \vdash_{\vartriangle} \psi}\,(\mathbf{Axiom})}{\phi, \phi \supset \psi \vdash_{\vartriangle} \psi}\,(\supset\mathbf{L})}{\phi, (\phi \supset \psi)[a \mapsto a] \vdash_{\vartriangle} \psi}\,(\mathbf{StructL}) \quad (\Delta \vdash_{\mathsf{SUB}} (\phi \supset \psi)[a \mapsto a] = \phi \supset \psi)}{\phi, \forall[a](\phi \supset \psi) \vdash_{\vartriangle} \psi}\,(\forall\mathbf{L})}{\phi, \forall[a](\phi \supset \psi) \vdash_{\vartriangle} \forall[a]\psi}\,(\forall\mathbf{R}) \quad (\Delta \vdash a\#\phi,\ \forall[a](\phi \supset \psi)).$$

The side-condition on $(\mathbf{StructL})$ is an instance of Lemma 4.10. The freshness side-conditions are straightforward from the definition of freshness derivability. $\square$

Lemma 6.10. *In the sequent calculus of one-and-a-halfth-order logic:*

- *bi-implication $\Leftrightarrow$ is an equivalence relation (a reflexive symmetric transitive relation), i.e. the following rules are admissible:*

$$\dfrac{}{\Phi \vdash_{\vartriangle} \Psi, \phi \Leftrightarrow \phi} \qquad \dfrac{\Phi \vdash_{\vartriangle} \Psi, \phi \Leftrightarrow \psi}{\Phi \vdash_{\vartriangle} \Psi, \psi \Leftrightarrow \phi} \qquad \dfrac{\Phi \vdash_{\vartriangle} \Psi, \phi \Leftrightarrow \psi \quad \Phi \vdash_{\vartriangle} \Psi, \psi \Leftrightarrow \xi}{\Phi \vdash_{\vartriangle} \Psi, \phi \Leftrightarrow \xi}$$

- *bi-implication is a congruence:*

$$\dfrac{\Phi \vdash_{\vartriangle} \Psi, \phi \Leftrightarrow \psi}{\Phi \vdash_{\vartriangle} \Psi, \xi[\phi/P] \Leftrightarrow \xi[\psi/P]}$$

- $\top$ *is the left and right identity of bi-implication:*

$$\frac{\Phi \vdash_{\vartriangle} \Psi, \phi}{\Phi \vdash_{\vartriangle} \Psi, \top \Leftrightarrow \phi} \qquad \frac{\Phi \vdash_{\vartriangle} \Psi, \top \Leftrightarrow \phi}{\Phi \vdash_{\vartriangle} \Psi, \phi} \qquad \frac{\Phi \vdash_{\vartriangle} \Psi, \phi}{\Phi \vdash_{\vartriangle} \Psi, \phi \Leftrightarrow \top} \qquad \frac{\Phi \vdash_{\vartriangle} \Psi, \phi \Leftrightarrow \top}{\Phi \vdash_{\vartriangle} \Psi, \phi}$$

Proof. By easy calculations using the derivation rules in Figure 2 (and the admissible rules in Figure 7). In the congruence case we use induction on the structure of ξ. □

Lemma 6.11. *For all sorts τ, terms $t, u : \tau$, unknowns $X : \tau$, formulae ϕ, and freshness contexts Δ:*

$$\textit{if} \quad \Delta \vdash_{\mathsf{FOL}} t = u \quad \textit{then} \quad \vdash_{\vartriangle} \phi[t/X] \Leftrightarrow \phi[u/X]$$

Proof. By induction on the structure of FOL derivations of $t = u$ from Δ.
(**refl**): $\vdash_{\vartriangle} \phi[t/X] \Leftrightarrow \phi[t/X]$ follows by reflexivity of $\Leftrightarrow$.
(**symm**): $\vdash_{\vartriangle} \phi[u/X] \Leftrightarrow \phi[t/X]$ follows from $\vdash_{\vartriangle} \phi[t/X] \Leftrightarrow \phi[u/X]$ by symmetry of $\Leftrightarrow$. By inductive hypothesis this follows from the assumption.
(**tran**): $\vdash_{\vartriangle} \phi[t/X] \Leftrightarrow \phi[v/X]$ follows from $\vdash_{\vartriangle} \phi[t/X] \Leftrightarrow \phi[u/X]$ and $\vdash_{\vartriangle} \phi[u/X] \Leftrightarrow \phi[v/X]$ by transitivity of $\Leftrightarrow$. By the inductive hypothesis these follow from the assumptions.
(**cong**[]): By the inductive hypothesis $\vdash_{\vartriangle} \psi[t/Y] \Leftrightarrow \psi[u/Y]$ for any Y and ψ. We must show $\vdash_{\vartriangle} \phi[[a]t/X] \Leftrightarrow \phi[[a]u/X]$, which is syntactically equivalent to

$$\vdash_{\vartriangle} \phi[[a]Z/X][t/Z] \Leftrightarrow \phi[[a]Z/X][u/Z],$$

where Z is an unknown (of appropriate sort) that does not occur in ϕ. This follows directly from the inductive hypothesis, taking $\psi \equiv \phi[[a]Z/X]$ and $Y \equiv Z$.
(**congf**): Analogous to the previous case.
(**perm**): we show $\vdash_{\vartriangle} \phi[(a\ b) \cdot t/X] \Leftrightarrow \phi[t/X]$ as follows:

$$\dfrac{\dfrac{}{\vdash_{\vartriangle} \phi[(a\ b) \cdot t/X] \Leftrightarrow \phi[(a\ b) \cdot t/X]}\,(\mathbf{Ax})}{\vdash_{\vartriangle} \phi[(a\ b) \cdot t/X] \Leftrightarrow \phi[t/X]}\,(\mathbf{StructR})$$

where $\Delta \vdash_{\mathsf{SUB}} \phi[(a\ b) \cdot t/X] \Leftrightarrow \phi[(a\ b) \cdot t/X] = \phi[(a\ b) \cdot t/X] \Leftrightarrow \phi[t/X]$ is the side-condition of (**StructR**). By (**congf**) and congruence Lemma 4.6, this follows from the assumption $\Delta \vdash_{\mathsf{SUB}} (a\ b) \cdot t = t$.
(**fr**): so $\Delta \vdash_{\mathsf{FOL}} t = u$ is derived from $\Delta, a\#X_1, \ldots, a\#X_n \vdash_{\mathsf{FOL}} t = u$, where $a \notin t, u, \Delta$. We must show $\vdash_{\vartriangle} \phi[t/X] \Leftrightarrow \phi[u/X]$. We cannot apply (**Fr**) directly, since ϕ might mention a. Using ZFA equivariance we rename a to $a' \notin t, u, \phi, \Delta$ while preserving the inductive hypothesis, to obtain $\vdash_{\Delta, a'\#X_1, \ldots, a'\#X_n} \phi[t/X] \Leftrightarrow \phi[u/X]$. We conclude $\vdash_{\vartriangle} \phi[t/X] \Leftrightarrow \phi[u/X]$ by (**Fr**), since $a' \notin \phi[t/X] \Leftrightarrow \phi[u/X], \Delta$.
($\mathbf{ax_A}$): We work by cases:

- If A is an axiom of SUB (from Figure 5) then it is of the form $\Delta \vdash_{\mathsf{SUB}} t = u$. By congruence Lemma 4.6 we know $\Delta \vdash_{\mathsf{SUB}} \phi[t/X] = \phi[u/X]$. We must show $\vdash_{\vartriangle} \phi[t/X] \Leftrightarrow \phi[u/X]$. By ($\Leftrightarrow$**R**), this follows from $\phi[t/X] \vdash_{\vartriangle} \phi[u/X]$ and $\phi[u/X] \vdash_{\vartriangle} \phi[t/X]$. The former can be derived:

$$\dfrac{\dfrac{}{\phi[t/X] \vdash_{\vartriangle} \phi[t/X]}\,(\mathbf{Axiom})}{\phi[t/X] \vdash_{\vartriangle} \phi[u/X]}\,(\mathbf{StructR}) \quad (\Delta \vdash_{\mathsf{SUB}} \phi[t/X] = \phi[u/X])$$

  The latter derivation is analogous, using (**StructL**).

- If A is an axiom from Figure 8 then the derivation is of the form
$$\frac{\Pi}{\phi^\pi\sigma = \psi^\pi\sigma}\,(\mathbf{ax}_{\mathbf{\Delta}'\vdash\phi=\psi})$$
where Π is a derivation of $\Delta'^\pi\sigma$. We must show $\vdash_\Delta \xi[\phi^\pi\sigma/P] \Leftrightarrow \xi[\psi^\pi\sigma/P]$. By congruence of $\Leftrightarrow$ (Lemma 6.10) this follows from $\vdash_\Delta \phi^\pi\sigma \Leftrightarrow \psi^\pi\sigma$. In case $\psi \equiv \top$, this follows from $\vdash_\Delta \phi^\pi\sigma$ by right identity of $\Leftrightarrow$. For each axiom the remaining proof obligation is an instance of part of Lemma 6.9, using the assumption $\Delta \vdash \Delta'^\pi\sigma$.

$\square$

Lemma 6.12. *If* $\vdash_\Delta \Phi^\wedge \supset \Psi^\vee$ *then* $\Phi \vdash_\Delta \Psi$.

Proof. By (**Cut**) $\Phi \vdash_\Delta \Psi$ follows from $\Phi \vdash_\Delta \Psi,\ \Phi^\wedge \supset \Psi^\vee$ and $\Phi^\wedge \supset \Psi^\vee,\ \Phi \vdash_\Delta \Psi$. The former follows from the assumption using weakening (Lemma 5.7). The latter follows by ($\supset$**L**) and a simple induction on the size of Φ and Ψ. $\square$

Sequent derivability is equivalent to FOL derivability:

Theorem 6.13. $\Phi \vdash_\Delta \Psi$ *if and only if* $\Delta \vdash_{\mathsf{FOL}} \Phi^\wedge \supset \Psi^\vee = \top$.

Proof. The left-to-right part is Theorem 6.8.

For the right-to-left part, assume $\Delta \vdash_{\mathsf{FOL}} \Phi^\wedge \supset \Psi^\vee = \top$. Then by Lemma 6.11 $\vdash_\Delta \Phi^\wedge \supset \Psi^\vee \Leftrightarrow \top$ is derivable. By right identity of $\Leftrightarrow$, also $\vdash_\Delta \Phi^\wedge \supset \Psi^\vee$. By Lemma 6.12 we obtain $\Phi \vdash_\Delta \Psi$, as required. $\square$

This theorem has some nice corollaries.

Corollary 6.14. *For any* Δ, ϕ, ψ*:*
$$\Delta \vdash_{\mathsf{FOL}} \phi = \psi \textit{ if and only if } \phi \vdash_\Delta \psi \textit{ and } \psi \vdash_\Delta \phi.$$

Proof. By Theorem 6.13 $\phi \vdash_\Delta \psi$ and $\psi \vdash_\Delta \phi$ are equivalent to $\Delta \vdash_{\mathsf{FOL}} \phi \supset \psi = \top$ and $\Delta \vdash_{\mathsf{FOL}} \psi \supset \phi = \top$. They are equivalent to $\Delta \vdash_{\mathsf{FOL}} \phi \Leftrightarrow \psi = \top$, by part 1 of Lemma 6.5. Finally, by part 2 of that lemma, this is equivalent to $\Delta \vdash_{\mathsf{FOL}} \phi = \psi$. $\square$

Corollary 6.15. FOL *is consistent, i.e.* $\Delta \vdash_{\mathsf{FOL}} \top = \bot$ *does not hold for any* Δ.

Proof. By contradiction. Suppose $\Delta \vdash_{\mathsf{FOL}} \top = \bot$. Using elementary calculations in propositional logic, also $\Delta \vdash_{\mathsf{FOL}} \top \supset \bot = \top$. Note that $\top \equiv \emptyset^\wedge$ and $\bot \equiv \emptyset^\vee$, so by Theorem 6.13 $\vdash_\Delta$ is derivable, which contradicts consistency of one-and-a-halfth-order logic (Corollary 5.12). $\square$

7. First-order logic with equality

In this section we show how first-order logic can be considered as the fragment of one-and-a-halfth-order logic without unknowns or explicit substitutions.

7.1. **Ground terms.** Call a nominal term **ground** if it does not mention unknowns (it is closed) and it does not mention sub (the substitution term-former). Ground terms of sort $\mathbb{T}$ and $\mathbb{F}$ are inductively characterised by
$$\begin{array}{lcl} t & ::= & a \mid \mathsf{f}_{i_1}(t_1,\ldots,t_{n_1}) \mid \ldots \\ \phi & ::= & \bot \mid \phi \supset \phi \mid \forall[a]\phi \mid t \approx t \mid \mathsf{f}_{j_1}(t_1,\ldots,t_{n_1}) \mid \ldots. \end{array}$$
Here i_1, i_2, … are the indexes of the object-level term-formers and j_1, j_2, … are the indexes of the predicate term-formers other than $\bot$, $\supset$, $\forall$, and $\approx$, as mentioned

in Subsection 2.1. We will not mention these term-formers again; they merely give rise to a few more cases in proofs.

For example, a and $\forall[a](a \approx b)$ are ground; $\forall[a]P$ and $a[a \mapsto a]$ are not.

We may write the term $\forall[a]\phi$ where ϕ is ground just as $\forall a.\phi$ (consistent with standard notation). Recall that a in $\forall[a]P$ may *not* be renamed in general, e.g. to $\forall[b]P$. Intuitively P represents an unknown formula which *might* mention a (if we know $b\#P$ we can at least rename to $\forall[b](b\ a).P$). To emphasise this we retained the notation $[a]\_$ until now. In $\forall a.\phi$ where ϕ is ground, we know all atoms in ϕ and this issue does not arise.

Write $fn(t)$ and $fn(\phi)$ for the **free names** of ground terms $t : \mathbb{T}$ and $\phi : \mathbb{F}$ respectively, inductively defined by:

$$fn(a) = \{a\}$$

$$fn(\bot) = \emptyset \qquad fn(\phi \supset \psi) = fn(\phi) \cup fn(\psi)$$
$$fn(\forall a.\phi) = fn(\phi) \setminus \{a\} \qquad fn(t_1 \approx t_2) = fn(t_1) \cup fn(t_2)$$

Lemma 7.1. *For ground terms $t : \mathbb{T}$ and $\phi : \mathbb{F}$:*

(1) $\vdash a\#t$ *if and only if* $a \notin fn(t)$.
(2) $\vdash a\#\phi$ *if and only if* $a \notin fn(\phi)$.

Proof. By simple induction on freshness derivations of on the one hand, and by induction on the definition of fn on the other. □

Define α**-equivalence** $=_\alpha$ as the least congruence on formulae such that

$$\frac{(a\ b)\cdot\phi =_\alpha \psi}{\forall a.\phi =_\alpha \forall b.\psi} \quad (b \notin fn(\phi)).$$

The reader might have expected the clause for $\forall$ to read something like

$$\frac{\phi_c =_\alpha \psi_c}{\forall a.\phi_a =_\alpha \forall b.\psi_b}$$

where here ϕ_c is informal notation for ϕ with every a replaced throughout by a freshly chosen c, and similarly for ψ_c. The two notions of α-equivalence are identical [13]. The definition we adopt gives a closer match to how equality is defined in nominal algebra (specifically to (**perm**)).

CORE is a theory of α-equivalence:

Theorem 7.2. *For ground terms $t, u : \mathbb{T}$ and $\phi, \psi : \mathbb{F}$:*

(1) $\vdash_{\mathsf{CORE}} t = u$ *if and only if* $t =_\alpha u$.
(2) $\vdash_{\mathsf{CORE}} \phi = \psi$ *if and only if* $\phi =_\alpha \psi$.

Proof. By known arguments of nominal results [18, 21]. □

For ground terms $t, u : \mathbb{T}$ and $\phi : \mathbb{F}$, write $u[\![a \mapsto t]\!]$ and $\phi[\![a \mapsto t]\!]$ for u and ϕ **with a replaced by t**, inductively defined by:

$$a[\![a \mapsto t]\!] \equiv t \qquad b[\![a \mapsto t]\!] \equiv b$$

$$\bot[\![a \mapsto t]\!] \equiv \bot \qquad (\phi \supset \psi)[\![a \mapsto t]\!] \equiv \phi[\![a \mapsto t]\!] \supset \psi[\![a \mapsto t]\!]$$
$$(\forall a.\phi)[\![a \mapsto t]\!] \equiv \forall a.\phi$$
$$(\forall b.\phi)[\![a \mapsto t]\!] \equiv \forall b'.\phi[\![b \mapsto b']\!][\![a \mapsto t]\!] \quad (b' \text{ fresh})$$
$$(t_1 \approx t_2)[\![a \mapsto t]\!] \equiv t_1[\![a \mapsto t]\!] \approx t_2[\![a \mapsto t]\!]$$

Here we make some arbitrary choice of b' for each ϕ, b, and b', and fix it for the rest of this paper.

On closed terms we can interpret all occurrences of term-former sub by capture-avoiding substitution.

Definition 7.3. *Define the **translation** $\_^{\mathrm{I}}$ of closed terms to ground terms inductively on closed terms by:*

$$a^{\mathrm{I}} \equiv a \qquad ([a]t)^{\mathrm{I}} \equiv [a](t^{\mathrm{I}}) \qquad (t[a \mapsto u])^{\mathrm{I}} \equiv t^{\mathrm{I}}[\![a \mapsto u^{\mathrm{I}}]\!].$$
$$\mathsf{f}(t_1, \ldots, t_n)^{\mathrm{I}} \equiv \mathsf{f}(t_1^{\mathrm{I}}, \ldots, t_n^{\mathrm{I}}) \quad (\mathsf{f} \neq \mathsf{sub})$$

The following nontrivial results are proved elsewhere [16] (the tools are also provided in [17]).

Lemma 7.4. *For closed terms t, u, if $\vdash a\#t$ then $a \notin fn(t^{\mathrm{I}})$.*

Lemma 7.5. *For closed terms t, $\vdash_{\mathsf{SUB}} t = t^{\mathrm{I}}$.*

Theorem 7.6. *For closed terms t, u, $\vdash_{\mathsf{SUB}} t = u$ if and only if $t^{\mathrm{I}} =_\alpha u^{\mathrm{I}}$.*

7.2. **Derivability in First-Order Logic.**

Definition 7.7. *A **Gentzen sequent** is a pair $\Phi \vdash \Psi$ of finite sets of* ground formulae *Φ and Ψ. The **valid judgements** of Gentzen's sequent calculus for first-order logic are the Gentzen sequents inductively specified by the rules in Figure 9.*

Here $fn(\Phi, \Psi)$ stands for the union of all $fn(\phi)$, $\phi \in \Phi, \Psi$. Furthermore, we take formulae up to α-equivalence, e.g. if $\mathsf{p} : (\mathbb{T})\mathbb{F}$ is a predicate term-former (such as $\mathsf{issocrates}$*) then $\forall a.\mathsf{p}(a) \vdash \forall b.\mathsf{p}(b)$ follows directly by* (**Ax**) *since $\forall a.\mathsf{p}(a) =_\alpha \forall b.\mathsf{p}(b)$.*

Lemma 7.8. *If $\Phi \vdash \Psi$ is derivable in Gentzen's sequent calculus then $\Phi \vdash_\emptyset \Psi$ is derivable in the sequent calculus for one-and-a-halfth-order logic.*

Proof. The statement of this lemma is a *little* bit vague, since we take formulae up to α-equivalence when we define Gentzen style derivability, but we do not take formulae up to α-equivalence in one-and-a-halfth-order logic (but we have structural rules (**StructL**) and (**StructR**) instead). We ignore this issue, and suppose that some arbitrary choice of representative closed nominal terms is made for us.

Suppose $\Phi \vdash \Psi$ is derivable in Genzten's sequent calculus. By induction on derivations of $\Phi \vdash \Psi$ we construct a derivation of $\Phi \vdash_\emptyset \Psi$. Each rule translates to its one-and-a-halfth-order counterpart, where we note the following:

- If at the meta-level in the Gentzen system we used α-conversion, or just if we wish to change representatives, then we can 'patch' the derivation in one-and-a-halfth-order logic with structural rules (**StructL**) and (**StructR**). The side-conditions follow by Theorem 7.2.

$$\frac{}{\phi,\ \Phi \vdash \Psi,\ \phi}\,(\mathbf{Ax}) \qquad \frac{}{\bot,\ \Phi \vdash \Psi}\,(\bot\mathbf{L})$$

$$\frac{\Phi \vdash \Psi,\ \phi \quad \psi,\ \Phi \vdash \Psi}{\phi \supset \psi,\ \Phi \vdash \Psi}\,(\supset\mathbf{L}) \qquad \frac{\phi,\ \Phi \vdash \Psi,\ \psi}{\Phi \vdash \Psi,\ \phi \supset \psi}\,(\supset\mathbf{R})$$

$$\frac{\phi[\![a \mapsto t]\!],\ \Phi \vdash \Psi}{\forall a.\phi,\ \Phi \vdash \Psi}\,(\forall\mathbf{L}) \qquad \frac{\Phi \vdash \Psi,\ \phi}{\Phi \vdash \Psi,\ \forall a.\phi}\,(\forall\mathbf{R}) \quad (a \notin \mathit{fn}(\Phi, \Psi))$$

$$\frac{\phi[\![a \mapsto t']\!],\ \Phi \vdash \Psi}{t' \approx t,\ \phi[\![a \mapsto t]\!],\ \Phi \vdash \Psi}\,(\approx\mathbf{L}) \qquad \frac{}{\Phi \vdash \Psi,\ t \approx t}\,(\approx\mathbf{R})$$

FIGURE 9. Gentzen's sequent calculus for first-order logic

- For the case of ($\forall$**L**) we need an extra use of (**StructL**) to manage the substitution. The side-condition $\vdash_{\mathsf{SUB}} \phi[a \mapsto t] = \phi[\![a \mapsto t]\!]$ follows by Lemma 7.5.
  The case of ($\approx$**L**) is similar.
- For the case of ($\forall$**R**) the side-conditions $\vdash a\#\Phi, \Psi$ follow from the assumption $a \notin \mathit{fn}(\Phi, \Psi)$ by Lemma 7.1.

□

If Φ is a closed formula context, write $\Phi^{\mathbb{I}}$ for the ground context $\{\phi^{\mathbb{I}} \mid \phi \in \Phi\}$.

Theorem 7.9. *For* closed Φ, Ψ, $\Phi \vdash_{\emptyset} \Psi$ *is derivable in one-and-a-halfth-order logic, if and only if* $\Phi^{\mathbb{I}} \vdash \Psi^{\mathbb{I}}$ *is derivable in Gentzen's sequent calculus.*

Proof. For the right-to-left direction, suppose $\Phi^{\mathbb{I}} \vdash \Psi^{\mathbb{I}}$ is derivable in Gentzen's sequent calculus. By Lemma 7.8, $\Phi^{\mathbb{I}} \vdash_{\emptyset} \Psi^{\mathbb{I}}$ is derivable in one-and-a-halfth-order logic. We also know that $\vdash_{\mathsf{SUB}} \phi^{\mathbb{I}} = \phi$ for each $\phi \in \Phi, \Psi$, by Lemma 7.5 and equational rule (**symm**). We use this to extend the derivation of $\Phi^{\mathbb{I}} \vdash_{\emptyset} \Psi^{\mathbb{I}}$ with applications of (**StructL**) or (**StructR**) for each $\phi^{\mathbb{I}} \in \Phi^{\mathbb{I}}, \Psi^{\mathbb{I}}$ to obtain one of $\Phi \vdash_{\emptyset} \Psi$, as required.

The left-to-right direction is by induction on derivations of $\Phi \vdash_{\emptyset} \Psi$. By Corollary 5.4 we assume that these derivations do not mention unknowns, and by Theorem 5.11 we assume that they do not mention (**Cut**).

We consider the rules in turn:

- (**StructL**) and (**StructR**) are *facts*, by Theorem 7.6.
- (**Fr**) is impossible since the derivation does not mention unknowns.
- The other rules translate directly to their first-order counterparts. For the case of ($\forall$**R**) we use Lemma 7.4.

□

Corollary 7.10. *For ground* Φ, Ψ, $\Phi \vdash_{\emptyset} \Psi$ *is derivable in one-and-a-halfth-order logic, if and only if* $\Phi \vdash \Psi$ *is derivable in Gentzen's sequent calculus.*

Proof. This is a direct instance of Theorem 7.9, since $\Phi^{\mathbb{I}} \equiv \Phi$ and $\Psi^{\mathbb{I}} \equiv \Psi$ when Φ and Ψ are ground. □

Corollary 7.11. *For ground* ϕ, ψ, $\vdash_{\mathsf{FOL}} \phi = \psi$ *in nominal algebra if and only if* $\phi \vdash \psi$ *and* $\psi \vdash \phi$ *are derivable in Gentzen's sequent calculus.*

Proof. By Corollaries 6.14 and 7.10. □

8. Conclusions

Explicitly representing meta-variables has a long pedigree.

Monadic second-order logic [6] enriches first-order logic with variables ranging over predicates of arity one representing unknown unary predicates, or if we prefer 'unknown sets'; the stronger second-order and higher-order logics [44, 40] represent unknowns as variables of function type. Representing unknowns as function variables has some distinctive features inherited from their intended functional semantics.

First, you have to choose the arity of your unknown in advance, e.g. function variable $f : \overbrace{\mathbb{T} \to \cdots \to \mathbb{T}}^{n} \to \mathbb{F}$ can be interpreted as an unknown n-ary predicate — but *which* n? Thus these logics distribute 'unknown predicates' across many types.

Second and perhaps more importantly, instantiation of these variables avoids capture. This has a side-effect that it is not possible to represent an unknown predicate uniformly across contexts. For example to represent the unknown ϕ in the context we can intuitively express as $\forall a.\phi$, it suffices to write $\forall\lambda a.f\,a$ where $f : \mathbb{T} \to \mathbb{F}$; but to represent the unknown ϕ in the context $\forall a.\forall b.\phi$ we must write $\forall\lambda a.\forall\lambda b.f\,a\,b$ — f must take a higher type. Since for any given f we must choose a type for it, it is not possible to directly represent the context we might write as $\mathsf{Qs}\phi$, where Qs represents an *unknown* context (of quantifiers).

Instantation of unknowns in nominal algebra does not avoid capture so that $\forall[a]P$ accurately reflects our intention when we write $\forall a.\phi$ where ϕ may be instantiated in a capturing manner. Furthermore P *still* represents ϕ in $\forall[a]\forall[b]P$.

This suggests an application to the problem of incomplete proofs. Incomplete proofs arise naturally in proof-search in a human-assisted theorem-prover such as Isabelle [36] or COQ [28]. Here the theorem-prover acts as a program to manipulate proofs which are incomplete both in the sense of having holes, and in the sense of occurring in an unknown context, and the human assistant guides the system to fill in these holes until a complete proof emerges. A quantifier introduction rule binds the quantified variable in the derivation above it (this is usually expressed by a freshness condition). In the presence of incomplete proofs it is necessary to somehow represent this binding over an *as yet unknown* derivation. We believe that nominal terms with their unknowns and abstractions are a good match for the incomplete proofs and quantifiers binding in it. As the incomplete proof is 'filled in' by the human assistant and the system, a capturing substitution should be made for the unknown. An exciting application of our technology would be an investigation of how well (if at all) a system related to one-and-a-halfth-order logic but with proof-terms can represent this process. A step in this direction would be to investigate a Curry-Howard correspondence [43] for one-and-a-halfth-order logic.

For this it might be convenient to use an intuitionistic flavour of one-and-a-halfth-order logic. Note that there is no technical barrier to this. An intuitionistic one-and-a-halfth-order logic can be defined in the usual way by restricting the sequents in Figure 2 to have a single conclusion.

Unknowns also lend a very distinctive style to our nominal algebra theory FOL (see Figures 5 and 8), and to the sequent rules of one-and-a-halfth-order logic (see Figure 2). These rules accurately reflect common practice in the handling of meta-variables, see [12, 27] and the examples of the Introduction, and as a result we have been able to import first-order proof theory quite directly into our augmented setting.

It is not possible to make a direct comparison between one-and-a-halfth-order logic and second-order logic. The second-order theorem $\forall P.((\forall P.P) \supset P)$ cannot be expressed in one-and-a-halfth-order logic, because one-and-a-halfth-order logic has no quantification over predicates; this gives it a first-order flavour. On the other hand the one-and-a-halfth-order logic theorem $\forall[a]P \vdash_{\mathsf{a\#P}} P$ cannot be expressed in second-order logic, because that logic cannot directly express freshness conditions. To the authors' knowledge no-one has been able to give a truly satisfactory account of the precise relation of nominal-style unknowns, and higher-order variables.

Our work is one (more) element in a very long line of investigations into algebraic logic [1]; for example cylindric [26], polyadic [25], and quantifier [38] algebra. There too, unknowns are syntax representing unknown elements quantified universally at top level, and abstraction $[a]$- (our notation) is clearly visible, e.g. as the c_i of cylindric algebra. However, these systems do not have as a goal to reflect the *proof-theory* of first-order logic. One-and-a-halfth-order logic does this. The connection with the algebraic method is then via Nominal Algebra and our axiomatisation of derivability in one-and-a-halfth-order logic.

It is possible to represent the *syntax* of first-order logic in a so-called *framework* logical system, at 'object-level', i.e. as *an inductive datatype* — the so-called *deep embedding*. Then meta-variables are easily representable as meta-variables of the framework. This path is taken by Higher-Order Abstract Syntax [37], Fraenkel-Mostowski syntax [21], the Theory of Contexts [34] and much other research. This enterprise is quite different from that undertaken in this paper; one-and-a-halfth-order logic is about extending the syntax of the logic itself so it contains something which behaves very much like a meta-variable ranging over unknown formulae, without losing logical properties such as cut-elimination.

It is also possible to represent the *semantics* of first-order logic as theory in a framework, for example as a pair of types $\mathbb{T}$ and $\mathbb{F}$ along with functions between them like $\supset: (\mathbb{F} \times \mathbb{F}) \to \mathbb{F}$ or $\forall : (\mathbb{T} \to \mathbb{F}) \to \mathbb{F}$. This is called a *shallow embedding*.

In [36] the case is made for Isabelle and for its higher-order logic framework as an efficient basis for shallow embeddings, and for conducting mathematics in these embeddings. For example, a shallow embedding of first-order logic called Isabelle/FOL exists in Isabelle's higher-order logic framework Isabelle/Pure. A shallow embedding requires only a logic with a sufficiently powerful judgement form. Our axiomatisation of one-and-a-halfth-order logic in nominal algebra, is a shallow embedding.

The technical content of this paper is therefore in two halves with equal status: a case study of a shallow embedding of first-order logic using nominal algebra as a logical framework, *and* a study of logic proof-theory in the presence of nominal-style unknowns.

Logic proof-theory with nominal-style unknowns is new, and seems to work quite well; the rules of Figure 2 are clean and close to informal practice. We can now try to extend this system with abstraction over meta-variables; we come back to this idea in a moment.

What of nominal algebra as a logical framework; how does it size up for example compared to higher-order approaches? Time and further research will decide just where nominal algebra fits in but we can make some preliminary observations.

Nominal algebra is *not* higher-order — this means that it is *not* based on terms up to $\alpha\beta$-equivalence; it is based on nominal terms. Unification on nominal terms is decidable [42] and this is not the case of unification up to $\alpha\beta$-equivalence. Thus, using nominal algebra avoids some of the complexity of higher-order unification. This paper and others [17, 10] show how nominal terms can express fully-functional logics, and in a very direct manner 'ϵ away from informal practice'.

The authors see no difficulties in principle with axiomatising substructural logics such as linear logic [24], bunched implications [35], relevance logics [9], and so on; if the logic is susceptible to a (nominal) algebraic treatment then it can be axiomatised in nominal algebra. This is not always the case of a shallow embedding in a higher-order logical framework, because the structural rules of the connectives and quantifiers of that logic may 'infect' those of the object-logic.

There is also a more subtle foundational issue. Formal proofs of basic theoretical properties such as Church-Rosser in systems such as Isabelle are important and technically informative, however there is some legitimate concern about their *foundational* status, seeing as the correctness of the basic framework is a higher-order logic and so depends amongst other things on… Church-Rosser. Nominal algebra is not higher-order and the correctness of meta-level properties does not appear to depend directly on Church-Rosser and other properties of λ-calculus syntax.

The technical tools used in this paper were developed based on work on Nominal Unification by the first author with Urban and Pitts [42], which introduced the theory of nominal terms up to CORE (our terminology). This was extended with Fernández [10] to Nominal Rewriting, a theory of rewriting on nominal terms, again up to CORE, and recently investigated with the second author, as a general framework of nominal algebra [18, 17].

For future work we are particularly interested in the following topics:

We can return to theory and be inspired by higher-order logic to ask whether we could permit abstraction over meta-variables, introducing an infinite hierarchy of stronger meta-variables such that at each level a meta-variable of higher level behaves to the lower level as an unknown X behaves to an atom a. This idea is explored in the arena of operational semantics by a λ-calculus called NEW calculus of contexts [14], and a rewriting framework called hierarchical nominal rewriting [15]. This might recover some or all of the power which one-and-a-halfth-order logic lacks compared to higher-order logic, but in a different way. In short, we envisage two- three- four- and ω-and-a-halfth-order logic. This would involve interesting extensions to the 'nominal theme'. Another direction is to allow unknowns ranging over derivations of sequents, which may have interesting interactions with $(\forall\mathbf{R})$, which would abstract in such an unknown.

The semantics of one-and-a-halfth-order logic are interesting and raise the questions 'what is an appropriate semantics for X', and 'what is an appropriate semantics for a'? Note that it is not possible to directly evaluate X to an element of a set underlying domain, because intuitively X 'can mention a'. Thus we can use domains in which atoms *can* appear (à la Fraenkel-Mostowski sets [18, 21] or other approaches [4, 11]). The simplest solution, and perhaps the best one, is to evaluate X to terms (a 'substitutional semantics' [30, Section 2] faithful to its intuition as an 'unknown term') and then evaluate atoms a to elements of a set underlying domain.

References

[1] H. Andréka, I. Németi, and I. Sain. Algebraic logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 2nd Edition*, volume 2, pages 133–249. Kluwer, 2001.

[2] J. Barwise. *Admissible Sets and Structures: an approach to definability theory*. Perspectives in mathematical logic. Springer, 1975.

[3] J. Barwise. An introduction to first-order logic. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 5–46. North Holland, 1977.

[4] M. Beeson. Lambda logic. In *2nd Int'l Joint Conf. on Automated Reasoning (IJCAR 2004)*, volume 3097 of *LNCS*, pages 460–474. Springer, 2004.

[5] J. Bell and M. Machover. *A course in mathematical logic*. North-Holland, 1977.

[6] C. Bruno. The expression of graph properties in some fragments of monadic second-order logic. *DIMACS Series in Discrete Mathematics*, 31, 1997.

[7] S. Burris and H. Sankappanavar. *A Course in Universal Algebra*. Springer, 1981.

[8] A. Degtyarev and A. Voronkov. Equality reasoning in sequent-based calculi. In J. A. Robinson and A. Voronkov, editors, *Handbook of Automated Reasoning*, pages 611–706. Elsevier and MIT Press, 2001.

[9] M. Dunn and G. Restall. Relevance logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 2nd Edition*, volume 6, pages 1–128. Kluwer, 2002.

[10] M. Fernández and M. J. Gabbay. Nominal rewriting. *Information and Computation*, 2005. In press.

[11] K. Fine. *Reasoning with Arbitrary Objects*. Blackwell, 1985.

[12] D. M. Gabbay and G. Malod. Naming worlds in modal and temporal logic. *Journal of Logic, Language and Information*, 11(1):29–65, 2002.

[13] M. J. Gabbay. *A Theory of Inductive Definitions with alpha-Equivalence*. PhD thesis, Cambridge, UK, 2000.

[14] M. J. Gabbay. A new calculus of contexts. In *PPDP '05: Proc. of the 7th ACM SIGPLAN int'l conf. on Principles and Practice of Declarative Programming*, pages 94–105. ACM Press, 2005.

[15] M. J. Gabbay. Hierarchical nominal rewriting. In *LFMTP'06: Logical Frameworks and Meta-Languages: Theory and Practice*, pages 32–47, 2006.

[16] M. J. Gabbay and A. Mathijssen. Capture-avoiding substitution as a nominal algebra (journal version). Submitted Formal Aspects of Computing.

[17] M. J. Gabbay and A. Mathijssen. Capture-avoiding substitution as a nominal algebra. In *ICTAC'2006: 3rd Int'l Colloquium on Theoretical Aspects of Computing*, volume 4281 of *LNCS*, pages 198–212, 2006.

[18] M. J. Gabbay and A. Mathijssen. A formal calculus for informal equality with binding. In *WoLLIC'07: 14th Workshop on Logic, Language, Information and Computation*, volume 4576 of *LNCS*, pages 162–176, 2007. To appear.

[19] M. J. Gabbay and A. Mathijssen. Nominal algebra. Technical report, Heriott-Watt, 2007.

[20] M. J. Gabbay and A. M. Pitts. A new approach to abstract syntax involving binders. In *14th Annual Symposium on Logic in Computer Science*, pages 214–224. IEEE Computer Society Press, 1999.

[21] M. J. Gabbay and A. M. Pitts. A new approach to abstract syntax with variable binding. *Formal Aspects of Computing*, 13(3–5):341–363, 2001.

[22] A. Gareau and R. Padmanabhan. Two axioms for implication algebras. *Notre Dame Journal of Formal Logic*, 2006.

[23] G. Gentzen. Untersuchungen über das logische schließen [Investigations into logical deduction]. *Mathematische Zeitschrift 39*, pages 176–210,405–431, 1935. Translated in [41], pages 68–131.

[24] J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50:1–102, 1987.

[25] P. Halmos. Algebraic logic, ii. homogeneous locally finite polyadic boolean algebras of infinite degree. *Fundamenta Mathematicae*, 43:255–325, 1956.

[26] L. Henkin, J. D. Monk, and A. Tarski. *Cylindric Algebras.* North Holland, 1971 and 1985. Parts I and II.

[27] W. Hodges. Elementary predicate logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 2nd Edition*, volume 1, pages 1–131. Kluwer, 2001.

[28] G. Huet, G. Kahn, and C. Paulin-Mohring. The Coq proof assistent, a tutorial. `http://pauillac.inria.fr/coq/doc/tutorial.html`. LogiCal Project.
[29] P. T. Johnstone. *Notes on logic and set theory*. Cambridge University Press, 1987.
[30] H. Leblanc. Alternatives to standard first-order semantics. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 2nd Edition*, volume 2, pages 53–132. Kluwer, 2001.
[31] W. McCune, R. Veroff, B. Fitelson, K. Harris, A. Feist, and L. Wos. Short single axioms for boolean algebra. *Journal of Automated Reasoning*, 1(29):1–16, 2002.
[32] C. A. Meredith. Single axioms for the systems (c,n), (c,o) and (a,n) of the two-valued propositional calculus. *The Journal of Computing Systems*, 3:155–164, 1953.
[33] Metamath.Org. Derivation of classical propositional logic from meredith's axiom. `us.metamath.org/mpegif/meredith.html`.
[34] M. Miculan. Developing (meta)theory of lambda-calculus in the theory of contexts. *ENTCS*, 1(58), 2001.
[35] P. W. O'Hearn and D. J. Pym. The logic of bunched implications. *Bulletin of Symbolic Logic*, 2(5):215–244, 1999.
[36] L. C. Paulson. Isabelle: the next 700 theorem provers. In P. Odifreddi, editor, *Logic and Computer Science*, pages 361–386. Academic Press, 1990.
[37] F. Pfenning and C. Elliot. Higher-order abstract syntax. In *PLDI '88: Proc. of the ACM SIGPLAN 1988 conf. on Programming Language design and Implementation*, pages 199–208. ACM Press, 1988.
[38] C. Pinter. A simple algebra of first-order logic. *Notre Dame Journal of Formal Logic*, 14(3):361–366, 1973.
[39] D. Prawitz. *Natural Deduction: A Proof Theoretical Study*. Almqvist and Wiksell, Stockholm, 1965.
[40] S. Shapiro. Systems between first-order and second-order logics. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 2nd Edition*, volume 1, pages 131–188. Kluwer, 2001.
[41] M. Szabo, editor. *Collected Papers of Gerhard Gentzen*. North Holland, 1969.
[42] C. Urban, A. M. Pitts, and M. J. Gabbay. Nominal unification. *Theoretical Computer Science*, 323(1–3):473–497, 2004.
[43] P. Urzyczyn and M. Sørensen. *Lectures on the Curry-Howard isomorphism*, volume 149 of *Studies in Logic*. Elsevier, 2006.
[44] J. van Benthem. Higher-order logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 2nd Edition*, volume 1, pages 189–244. Kluwer, 2001.

Appendix A. Equivariance

We use atoms in this paper — we introduce them when we write 'Fix a countably infinite collection of **atoms** a, b, c' in Section 2.

We can represent atoms as, say, numbers $0, 1, 2, 3, \ldots$, or as sets $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots$. In principle we might 'accidentally' use some property of atoms specific to their implementation, such as $a \leq b$ or $a \in b$. However, we know we do not do this, because we consider atoms to be ... atomic.

By explicitly bearing this in mind, we can rename atoms. This is *equivariance*, proved below, which we use freely in this paper to give structural inductive proofs (renaming atoms in inductive hypotheses where convenient) while remaining fully formal.

Without equivariance, to be fully formal proofs by induction on measures such as term length or derivation depth would be necessary. These are longer, harder to read, and are rarely given in full detail outside of a theorem-prover.

To give a precise statement and formal proof of equivariance and how it can be used to rename atoms in proofs by structural induction on the syntax of terms and derivation-trees, it is well-worth a brief look at foundations so that we can prove a few simple theorems about a foundation in which atoms really *are* atomic.

ZFA stands for 'Zermelo-Fraenkel set theory with atoms'. ZFA is a set theory, and set theories are just axioms in first-order logic with equality (similar to the one we defined in Section 7). Equivariance was introduced in [20]. It is a meta-mathematical property, meaning that it refers to the assertions proved in this paper.

We do *not* mean that equivariance refers to the terms of sort $\mathbb{F}$, nor the formulae of first-order logic, nor sequents of one-and-a-halfth-order logic. We mean that it refers to the assertions written *in english in this paper* — *about* one-and-a-halfth-order logic, nominal algebra, and so on.

ZFA features a set $\mathbb{A}$ of *atoms* $a, b, c, \ldots$. The original motivation of atoms was to address the question 'what set is equal to Plato?'. Obviously Plato is not any of the 'normal' sets of set theory, such as the empty set $\emptyset$ or the set containing just the empty set $\{\emptyset\}$, and so on. The answer was to accommodate 'the real world' by introducing it *en masse* into the set model, as atoms.

As far as the set theory is concerned atoms are atomic objects with no internal (set-)structure, so it is quite natural to use these to model variable symbols. This idea appears already in [2]. As discussed, the advantage of this over an atom-less foundation such as ZF, is that equivariance is guaranteed, as we shall see.

For the language of ZFA set theory, in addition to the basic language of first-order logic with equality, we assume:

- A binary predicate symbol $\in$ called *set inclusion*.
- A constant term-former $\mathbb{A}$ called *the set of atoms*.

We use standard sugar of classical logic (similar to the sugar mentioned in Subsection 2.1).

Definition A.1. ***ZFA set theory*** *has the axioms in Figure 10.*

Here, ϕ ranges over all predicates, $\phi[y/x]$ denotes the predicate obtained by capture-avoiding substitution of x by y, and $F(y)$ represents a function on the sets universe (strictly speaking, this is itself sugar, which is briefly described in Corollary A.4). Furthermore, we use the following sugar:

$$
\begin{array}{lll}
x = \{z \mid z \in x\} & \textit{is sugar for} & \forall y.(\forall z.(z \in x \Leftrightarrow z \in y) \supset x = y) \\
y = \{z \in x \mid \phi\} & \textit{is sugar for} & \forall z.(z \in y \Leftrightarrow (z \in x \wedge \phi)) \\
z = \{F(y) \mid y \in x\} & \textit{is sugar for} & \forall u.(u \in z \Leftrightarrow \exists y.(F(y) = u \wedge y \in x)) \\
z = \{x, y\} & \textit{is sugar for} & \forall u.(u \in z \Leftrightarrow (u = x \vee u = y)) \\
z = \{y \mid \exists y'.(y \in y' \wedge y' \in x)\} & \textit{is sugar for} & \forall y.(y \in z \Leftrightarrow \exists y'.(y \in y' \wedge y' \in x)) \\
z = \{y \mid y \subseteq x\} & \textit{is sugar for} & \forall y.(y \in z \Leftrightarrow \forall y'.(y' \in y \supset y' \in x)) \\
\emptyset \in x & \textit{is sugar for} & \exists z.(z \in x \wedge \forall z'.z' \notin z) \\
y \cup \{z\} \in x & \textit{is sugar for} & \exists u.(u \in x \wedge \forall u'.(u' \in u \Leftrightarrow u \in y \vee u = z))
\end{array}
$$

The syntactic sugar used in set theory is very rich; further details can be found elsewhere [29].

Note that atoms are **extensionally equal** to the empty set. This means that the same elements are related to them by $\in$: $a \in \mathbb{A} \supset \forall x.x \notin a$, but $a \neq \emptyset$ (and $\emptyset$ is not an atom). $\emptyset$ is not an atom, and is distinguished from atoms in that it is the unique empty set which is not in the set of atoms.

We can define a **permutation action** on ZFA sets by:

$$
(a\ b)a = b \qquad (a\ b)b = a \qquad (a\ b)c = c \quad (c \neq a, b)
$$
$$
(a\ b)X = \{(a\ b)x \mid x \in X\} \quad (X \notin \mathbb{A})
$$

This definition is by *ϵ-induction*. Definition by ϵ-induction is a standard method in set theory [29]; if it is true that a property holds of $y \in \mathbb{A}$, and *if* it holds of all

(**Sets**)	$\forall x.((\exists y.y \in x) \supset x \notin \mathbb{A})$	
(**Extensionality**)	$\forall x.(x \notin \mathbb{A} \supset x = \{z \mid z \in x\})$	
(**Collection**)	$\forall x.\exists y.(y \notin \mathbb{A} \land y = \{z \in x \mid \phi\})$	(y not free in ϕ)
($\in$-**Induction**)	$\forall x.(\forall y.(y \in x \supset \phi[y/x]) \supset \phi) \supset \forall x.\phi$	
(**Replacement**)	$\forall x.\exists z.(z \notin \mathbb{A} \land z = \{F(y) \mid y \in x\})$	
(**Pairset**)	$\forall x.\forall y.\exists z.(z = \{x, y\})$	
(**Union**)	$\forall x.\exists z.(z \notin \mathbb{A} \land z = \{y \mid \exists y'.(y \in y' \land y' \in x)\})$	
(**Powerset**)	$\forall x.\exists z.(z = \{y \mid y \subseteq x\})$	
(**Infinity**)	$\exists x.(\emptyset \in x \land \forall y.(y \in x \supset y \cup \{y\} \in x))$	

FIGURE 10. Axioms of ZFA set theory

$x \in y$ *then* it holds of all y, then that property holds of all y. Written informally: *sets are well-founded trees with daughter-of given by set inclusion* $\in$.

We read $(a\ b)X = \{(a\ b)x \mid x \in X\}$ as

> $(a\ b)$ acting on a set is equal to the set of $(a\ b)$ acting on the elements of that set

or even more succinctly

> **The permutation action is pointwise.**

Write $\circ$ for functional composition. So $\pi \circ \pi'$ maps a to $\pi(\pi'(a))$.

Lemma A.2. $\pi(\pi' z) = (\pi \circ \pi')z$.

Proof. The proof is by ϵ-induction.

- By definition if $a \in \mathbb{A}$ then $\pi(\pi' a) = (\pi \circ \pi')a$.
- Suppose $Z \notin \mathbb{A}$. Then by definition of the permutation action and by the inductive hypothesis

$$\begin{aligned}\pi(\pi' Z) &= \{\pi(\pi' u) \mid u \in Z\} = \{(\pi \circ \pi')u \mid u \in Z\} \\ &= (\pi \circ \pi')\{u \mid u \in Z\} = (\pi \circ \pi')Z.\end{aligned}$$

□

Recall that ϕ ranges over predicates of ZFA. Write $\phi(x_1, \ldots, x_n)$ to range over predicates which mention at most $x_1, \ldots, x_n$ as free variable symbols.

Theorem A.3. *If $\phi(x_1, \ldots, x_n)$ is a predicate of ZFA set theory then*

$$\phi(x_1, \ldots, x_n) \Leftrightarrow \phi(\pi x_1, \ldots, \pi x_n)$$

is always provable.

As a corollary, $\phi(x_1, \ldots, x_n)$ and $\phi(\pi x_1, \ldots, \pi x_n)$ are interchangeable in proof and in validity on models.

Proof. We work by induction on the syntax of ϕ.

- $x \in y$ implies $\pi x \in \pi y$ follows directly from the fact that $\pi y = \{\pi y' \mid y' \in y\}$ by definition. The reverse implication is easy using π^{-1}, the inverse of π.
- Similarly, $x = y$ if and only if $\pi x = \pi y$.
- If $\phi(x_1, \ldots, x_n)$ is of the form $\phi_1 \supset \phi_2$ then we may easily use the inductive hypothesis.

- Likewise if $\phi(x_1, \ldots, x_n)$ is $\bot$, or is of the form $\forall z.\phi'$, we may easily use the inductive hypothesis.
- $\pi\mathbb{A} = \mathbb{A}$ is provable, so $x \in \mathbb{A}$ if and only if $\pi x \in \mathbb{A}$, and $\mathbb{A} \in y$ if and only if $\mathbb{A} \in \pi y$, and similarly $x = \mathbb{A}$ if and only if $\pi x = \mathbb{A}$ and $\mathbb{A} = y$ if and only if $\mathbb{A} = \pi y$.

The result follows. □

As a corollary we have:

Corollary A.4. *If $F(x_1, \ldots, x_n)$ is function (not a function-set) from the set universe to itself then*

$$\pi(F(x_1, \ldots, x_n)) = F(\pi x_1, \ldots, \pi x_n)$$

is always provable.

Proof. In set theory, we specify F using a predicate $\phi(x_1, \ldots, x_n, z)$ such that

$$\forall x_1, \ldots, x_n.(\exists z.\phi(x_1, \ldots, x_n, z) \ \wedge \forall z, z'.(\phi(x_1, \ldots, x_n, z) \wedge \phi(x_1, \ldots, x_n, z') \supset z = z')).$$

□

Dept. of Computer Science, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, UK

E-mail address: murdoch.gabbay@gmail.com

Dept. of Mathematics and Computer Science, Eindhoven University of Technology (TU/e), P.O. Box 513, 5600 MB Eindhoven, The Netherlands

E-mail address: aad.mathijssen@gmail.com