The Explicit Definition of Quantifiers via Hilbert's ε is Confluent and Terminating

Claus-Peter Wirth
Dept. of Computer Sci., Saarland Univ., 66123 Saarbrücken, Germany
wirth@logic.at

Abstract

We investigate the elimination of quantifiers in first-order formulas via Hilbert's epsilon-operator (or -binder), following Bernays' explicit definitions of the existential and the universal quantifier symbol by means of epsilon-terms. This elimination has its first explicit occurrence in the proof of the first epsilon-theorem in Hilbert–Bernays in 1939. We think that there is a lacuna in this proof w.r.t. this elimination, related to the erroneous assumption that explicit definitions always terminate. Surprisingly, to the best of our knowledge, nobody ever published a confluence or termination proof for this elimination procedure. Even myths on non-confluence and the openness of the termination problem are circulating. We show confluence and termination of this elimination procedure by means of a direct, straightforward, and easily verifiable proof, based on a theorem on how to obtain termination from weak normalization.

Keywords: Hilbert–Bernays Proof Theory, History of Proof Theory, Hilbert's epsilon, Quantifier Elimination, (Weak) Normalization, (Strong) Termination, (Local) Confluence.

1 Introduction

1.1 The Explicit Historical Source of the Problem

With "Hilbert–Bernays" we will designate the "bible of proof theory", i.e. the two-volume monograph *Grundlagen der Mathematik* (*Foundations of Mathematics*) in its two editions [Hilbert & Bernays, 1934; 1939] and [Hilbert & Bernays, 1968; 1970].

On p.19f. of [Hilbert & Bernays, 1939], as well as on p. 20 of the second edition [Hilbert & Bernays, 1970], we read:

> "Unser zweiter vorbereitender Schritt besteht in der Ausschaltung der All- und Seinszeichen. Wie im vorigen Abschnitt gezeigt wurde, können wir die Anwendung der Grundformeln (a), (b) und der Schemata (α), (β) des Prädikatenkalkuls mit Hilfe der ε-Formel und der expliziten Definitionen (ε_1), (ε_2) entbehrlich machen[1]. Führen wir diese Ausschaltung der Grundformeln und Schemata für die Quantoren an der zu betrachtenden Ableitung der Formel $\mathfrak{E}$ aus und ersetzen wir hernach jeden Ausdruck $(\mathfrak{v})\,\mathfrak{A}(\mathfrak{v})$ durch $\mathfrak{A}\big(\varepsilon_{\mathfrak{v}}\,\overline{\mathfrak{A}(\mathfrak{v})}\,\big)$, jeden Ausdruck $(E\,\mathfrak{v})\,\mathfrak{A}(\mathfrak{v})$ durch $\mathfrak{A}\big(\varepsilon_{\mathfrak{v}}\,\mathfrak{A}(\mathfrak{v})\big)$, so gehen die aus (ε_1), (ε_2) durch Einsetzung gewonnenen Formeln in solche über, die durch Einsetzung aus der Formel $A \sim A$ entstehen. Die Quantoren werden durch dieses Verfahren gänzlich ausgeschaltet, so daß *nunmehr gebundene Variablen ausschließlich in Verbindung mit dem ε-Symbol auftreten, und der Beweiszusammenhang nur durch Wiederholungen, Einsetzungen, Umbenennung gebundener Variablen und Schlußschemata stattfindet.*"

> "Our second preparatory step consists in the elimination of the universal and existential quantifier symbols. As shown in the previous section, we can dispense with the application of Formulas (a), (b) and Schemata (α), (β) of the predicate calculus if we use the ε-formula and the explicit definitions (ε_1), (ε_2). If we apply this elimination of basic formulas und schemata for the quantifiers to the formula $\mathfrak{E}$ under consideration, and afterwards replace every expression $(\mathfrak{v})\,\mathfrak{A}(\mathfrak{v})$ with $\mathfrak{A}\big(\varepsilon_{\mathfrak{v}}\,\overline{\mathfrak{A}(\mathfrak{v})}\,\big)$, every expression $(E\,\mathfrak{v})\,\mathfrak{A}(\mathfrak{v})$ with $\mathfrak{A}\big(\varepsilon_{\mathfrak{v}}\,\mathfrak{A}(\mathfrak{v})\big)$, then the formulas obtained from (ε_1), (ε_2) by substitution are turned into formulas obtained by substitution from the formula $A \sim A$. By this procedure, the quantifiers are completely eliminated, so that *bound variables may occur only in combination with the ε-symbol, and the interconnections of the proof may consist only of repetitions, substitutions, renaming of bound variables, and inference schemata.*"

Note that the "A" is not a meta-variable here (such as "$\mathfrak{A}$" is a meta-variable for a formula, and "$\mathfrak{v}$" for a *bound individual variable*), but a concrete object-level formula variable. In a proof step called *substitution* either such a *formula variable* (which is always free) or a *free individual variable* is replaced everywhere in a formula with an arbitrary formula or term, respectively. Furthermore, note that "Schlußschema" ("inference schema") is nothing but a short name for the inference schema of *modus ponens.*

Moreover, note that Note 1 actually occurs only in the second edition and reads "[1]Vgl. S.15." ("[1]Cf. p.15."). Neither on Page 15 — nor anywhere else in the volumes — can we find any further information, however, regarding the following immediate questions:

- In which order are the final replacements of the two explicitly mentioned forms of expressions to be applied in the elimination of quantifiers?
- Or are such eliminations independent of the order of the replacements in the sense that they always yield unique normal forms?

What we can actually find on Page 15 are the mentioned "explicit definitions (ε_1), (ε_2)", which describe the rewrite relation of these replacements. In the more modern notation we prefer for this paper, these explicit definitions read:

$$\exists x.\ A \quad\Leftrightarrow\quad A\{x \mapsto \varepsilon x.\ A\} \tag{ε_1}$$

$$\forall x.\ A \quad\Leftrightarrow\quad A\{x \mapsto \varepsilon x.\ \neg A\} \tag{ε_2}$$

Note that x is a meta-variable for *individual variables* (in the original: a concrete object-level, bound individual variable), and A is a meta-variable for formulas (in the original: a concrete object-level, singulary formula variable). The original version of (ε_1) literally reads: $(Ex)\,A(x) \sim A\big(\varepsilon_x\,A(x)\big)$.

Note that the formulas considered here and in what follows are always first-order formulas, extended with ε-terms and possibly also with free (second-order) *formula variables.* For our considerations in this paper, it does not matter whether we include such formula variables into our first-order formulas or not.

1.2 Subject Matter

What we will study in this paper is the question how the elimination of first-order quantifiers via their explicit definitions can take place.

Here we should recall that, in *explicit definitions* (contrary to recursive definitions), the symbol to be defined (here: $\exists$ or $\forall$), occurring on the left-hand side of an equation (the *definiendum*) must not re-occur in the term on the right-hand side (*definiens*).

In this standard terminology, (ε_1) and (ε_2) classify as explicit definitions, because $\exists$ and $\forall$ do not occur on the right-hand sides — at least not explicitly.

It is commonplace knowledge that (contrary to recursive or implicit definitions) explicit definitions are analytic (i.e. not synthetic) in the sense that they cannot contribute anything essential to our knowledge base — simply because any notion introduced by an explicit definition can be eliminated from any language (at least in principle) after replacing all *definienda* with their respective *definientia.*

For first-order terms the eliminability is indeed trivial, even for non-right-linear equations such as

$$\mathsf{russell}(x) = \mathsf{mbp}(x, x),$$

where the number of occurrences of defined symbols in x is doubled when rewriting with this equation; i.e., if $n(t)$ denotes the number of explicitly defined symbols in the term t, then $n(\mathsf{russell}(t)) = n(t) + 1$, whereas $n(\mathsf{mbp}(t, t)) \geq 2 * n(t)$.

The termination of a stepwise elimination by applying one equation after the other — until no defined symbols remain — does not crucially depend on whether we rewrite the defined symbols in t before we apply the equation for the defined term $\mathsf{russell}(t)$ or after. Indeed, the difference this alternative can make is only a duplication of the rewrite steps required for the normalization of t.

This argumentation, however, does not straightforwardly apply to our definitions (ε_1), (ε_2). Indeed, the instance of the first occurrence of the meta-variable A on the right-hand side is modified by a substitution that may introduce an arbitrarily large number of copies of the instance of A.

We will show in this paper, however, that rewriting of an arbitrary formula F with (ε_1), (ε_2) is always confluent and terminating. This means that, no matter in which order we eliminate the quantifiers, a resulting quantifier-free formula will always be obtained, and that this formula is a unique normal form for F.

1.3 A Lacuna in Hilbert–Bernays?

The fact that this rewriting is innermost terminating has been well known before, but none of the experts on Hilbert's ε we consulted knew about the strong termination (i.e. termination independent of any rewriting strategy), and one of them even claimed that the rewriting would not be confluent.

As the proofs of the ε-theorems of [Hilbert & Bernays, 1939] show, Paul Bernays (1888–1977) was well aware of the influence of strategies on elimination procedures. The mathematical technology of the 1930s, however, makes it most unlikely that he could easily show the strong termination — let alone consider it to be trivial in the context of a textbook (such as Hilbert–Bernays).

Moreover, the actual formula language of Hilbert–Bernays strongly suggests an outermost strategy: A non-outermost rewriting typically requires the instantiation of A to formulas containing variables that are bound by the outer quantifiers and epsilons. Such an instantiation is not permitted in Hilbert–Bernays, however, because these additional variables must come from a set of variables different from the free individual variables, which are called *bound* individual variables and which are not permitted to occur free in a substitution for A. Thus, for an innermost rewriting in the predicate calculus of Hilbert–Bernays, we have to resort to multiple tacit applications of Rule (δ') for a complete reconstruction of the whole outer part of the formula in each innermost rewrite step; for Rule (δ') see e.g. Page 109 in [Hilbert & Bernays, 1968; 2016b].

All in all, the fact that neither the innermost rewriting strategy nor Rule (δ') is mentioned in this context in [Hilbert & Bernays, 1939] makes it most likely that Bernays just relied here on his learning that explicit definitions always admit an elimination, which is actually not the case in general for higher-order definitions.

1.4 Alternative Proofs by Applying Theories of First- or Higher-Order Rewriting?

In this paper, we will approach our results directly, without applying the theory of first- or higher-order rewrite systems. Other options for obtaining the crucial termination result could be:

1. To map the first-order terms with quantifiers and epsilons to quantifier- and epsilon-free first-order terms, to find a first-order term rewriting system that admits the transitive reduction of the images of any original reduction, and to prove the termination of the first-order term rewriting system, using the powerful theorems and methods to establish termination of first-order term rewriting systems (or even some of the software systems that may show first-order termination automatically, cf. e.g. [Winkler &al., 2013]).

2. To apply some results on termination of higher-order rewriting systems.

3. To map the first-order terms with quantifiers and epsilons to Church's simply-typed λ-calculus (which is known to be terminating), such that the images of each original reduction admit the transitive reduction in simply-typed λ-calculus.

Let us look at second-order formulations of (ε_1), partly because the original formulation of Hilbert's ε as found in [Ackermann, 1925] and [Hilbert, 1926; 1928] is already a second-order one without binders, and partly to develop options 2 and 3 a bit further.

If we use i to designate the sort (basic type) of individuals and o to designate the sort of formulas (as standard in Church's simply-typed λ-calculus), then the ε gets the typing of $\varepsilon : (i \to o) \to i$, and for a second-order variable $A : i \to o$ and the existential operator $\Sigma : (i \to o) \to o$, we get

$$\Sigma A = A(\varepsilon A),$$

or in η-expanded form

$$\Sigma\lambda x.(Ax) = A(\varepsilon\lambda x.(Ax)).$$

To implement these equations according to option 2, we have to pick one of the three competing higher-order rewriting frameworks, namely *combinatory reduction systems (CRSs)* [Klop, 1980], [Klop &al., 1993], *higher-order rewrite systems* [Nipkow, 1991], [Raamsdonk, 1999], and *algebraic-functional systems* [Jouannaud & Okada, 1991]. We pick the CRS framework because it is the oldest and most popular one (also admitting extension to conditional rewriting straightforwardly, cf. [Wirth, 2009, Note 9]).

In CRS syntax (cf. e.g. [Klop &al., 1993, § 11]), the η-expanded rule reads

$$\Sigma[x](A(x)) = A(\varepsilon[x](A(x))),$$

where x is a variable, A is a singulary *meta-variable* (not only a top-level one, but also w.r.t. the special technical terms used for CRSs, i.e. a meta-variable for a special variable that must not occur in the terms in the range of the rewrite relation), Σ and ε are singulary function symbols (i.e. 1-ary constant symbols), and $[x]$ is an abstraction operator, binding the variable x. In this notation, we indeed have a CRS *rewrite rule* with the intended rewrite relation. We can formulate (ε_2) in a similar way, resulting in a two-rule CRS that is *orthogonal* (called "regular" in [Klop, 1980]), i.e. non-overlapping ("non-ambiguous") and left-linear. Thus, according to [Klop &al., 1993, Corollary 13.6] ([Klop, 1980, Theorem II.3.11]), the rewrite relation is confluent.

As it is obvious that this rewrite relation is weakly normalizing (as it is innermost terminating), its termination (strong normalization) follows from Theorem II.5.9.3 of [Klop, 1980, p.168], provided that we can show our rewrite relation to be non-erasing. This means that we have to show that the set of free variables is invariant under rewrite steps. Note that the instance of A may contain free variables (such as y), but even if the instance of A is, say, $\underline{\lambda}[x](y = y)$ (i.e. the quantifier is vacuous, binding a variable that does not occur in its scope), it seems that the deletion of the second occurrence of A in the right-hand side does not matter, because all occurrences of free variables are preserved by the first occurrence of A in the right-hand side.

This argumentation, however, forgets that CRSs come without β-reduction. So we may need the rule $(\lambda[x](A(x)))B = A(B)$ in addition, which would render the CRS erasing. On the other hand, $\underline{\lambda}$ is different from λ (although some crucial underlining of λ is missing in [Klop &al., 1993]) and part and parcel of the substitution framework for "meta-variables" in [Klop &al., 1993]; this means we should get along without the β-rule for λ, provided that we write existential quantification in our formulas as, say, "$\Sigma[x]$" instead of "$\Sigma\lambda x.$".

If the latter is indeed the case, and if our understanding of [Klop, 1980] is the right one, then confluence and termination can be established by applying the theory of CRSs.

As the contacted experts on higher-order rewriting did not want to help settling these questions (and no answer was found in [Raamsdonk, 2001], [Ketema & Raamsdonk, 2004] either), and as the effort to familiarize oneself (again) with the most fascinating and outstanding work documented in the PhD thesis [Klop, 1980] is considerable and disproportionate for our subject matter, we will present here a straightforward and efficiently verifiable proof of termination and confluence of the reduction relation defined directly on first-order terms with quantifiers and epsilons.

To implement option 3, however, we could to take ε as a constant with the above typing and the mapping to Church's simply-typed λ-calculus could replace the previous constant Σ with the λ-term $\lambda A.\ (A(\varepsilon A))$ of the same type as Σ before. Then reduction by the first of the above equations could be done by a first β-reduction, and a second β-reduction on the λ-term A could be used to reduce $A(\varepsilon A)$, such that an original reduction step with (ε_1) results in two β-reduction steps after the mapping to simply-typed λ-calculus. Although this proof plan is most promising, it is not easily accessible in the sense that a mathematician could verify it without a careful formalization of lots of technical and syntactic details. Moreover, as Bernays in the 1930s could not have known about the termination of simply-typed λ-calculus — first shown by Tait [1967] — this is not a proof plan he could have followed (though he was in correspondence with Church and visiting the Institute for Advanced Study in Princeton during session 1935/36).

Finally, note that — compared to options 1–3 — our direct and efficiently verifiable procedure is not only more informative on the concrete structure of the particular subject matter, but also the stronger, more concise, and historiographically more relevant evidence against myths on Hilbert–Bernays with regard to non-confluence and openness of the termination question.

2 Background and Tools

2.1 Basic Notions and Notation

We follow standard mathematical writing style, cf. [Gillman, 1987].

We try to be self-contained in this this paper. In case we should omit some required information, we refer the reader to the survey [Klop, 1980, § I.5] on abstract rewrite systems.

Let '$\mathbf{N}$' denote the set of natural numbers, and '$<$' the ordering on $\mathbf{N}$. Let $\mathbf{N}_{+} := \{\, n \in \mathbf{N} \mid 0 \neq n \,\}$.

For classes R, A, and B we define:

$\mathrm{dom}(R)$	$:= \{\, a \mid \exists b.\ (a,b) \in R \,\}$	*domain*
${}_{A}\!\upharpoonleft\! R$	$:= \{\, (a,b) \in R \mid a \in A \,\}$	*(domain-) restriction to* A
$\langle A \rangle R$	$:= \{\, b \mid \exists a \in A.\ (a,b) \in R \,\}$	*image of* A, i.e. $\langle A \rangle R = \mathrm{ran}({}_{A}\!\upharpoonleft\! R)$

And the dual ones:

$\mathrm{ran}(R)$	$:= \{\, b \mid \exists a.\ (a,b) \in R \,\}$	*range*
$R\!\upharpoonright_{B}$	$:= \{\, (a,b) \in R \mid b \in B \,\}$	*range-restriction to* B
$R\langle B \rangle$	$:= \{\, a \mid \exists b \in B.\ (a,b) \in R \,\}$	*reverse-image of* B, i.e. $R\langle B \rangle = \mathrm{dom}(R\!\upharpoonright_{B})$

We use 'id' for the identity function, and '$\circ$' for the composition of binary relations. Functions are (right-) unique relations, and so the meaning of "$f \circ g$" is extensionally given by $(f \circ g)(x) = g(f(x))$.

Let $\longrightarrow$ be a binary relation. $\longrightarrow$ is a relation *on* A if

$$\mathrm{dom}(\longrightarrow) \cup \mathrm{ran}(\longrightarrow) \subseteq A.$$

$\longrightarrow$ is *irreflexive* if $\mathrm{id} \cap \longrightarrow\ = \emptyset$. It is A*-reflexive* if ${}_{A}\!\upharpoonleft\!\mathrm{id} \subseteq\ \longrightarrow$. Speaking of a *reflexive* relation we refer to the largest A that is appropriate in the local context, and referring to this A we write $\xrightarrow{0}$ to ambiguously denote ${}_{A}\!\upharpoonleft\!\mathrm{id}$. With $\xrightarrow{1}\ :=\ \longrightarrow$, and $\xrightarrow{n+1}\ :=\ \xrightarrow{n} \circ \longrightarrow$ for $n \in \mathbf{N}_{+}$, $\xrightarrow{m}$ denotes the m-step relation for $\longrightarrow$. The *transitive closure* of $\longrightarrow$ is $\xrightarrow{+}\ := \bigcup_{n \in \mathbf{N}_{+}} \xrightarrow{n}$. The *reflexive closure* of $\longrightarrow$ is $\xrightarrow{=}\ := \bigcup_{n \in \{0,1\}} \xrightarrow{n}$. The *reflexive transitive closure* of $\longrightarrow$ is $\xrightarrow{*}\ := \bigcup_{n \in \mathbf{N}} \xrightarrow{n}$. The *reverse* of $\longrightarrow$ is $\longleftarrow\ := \{\ (b,a) \mid (a,b) \in\ \longrightarrow \}$.

v and w are called *joinable w.r.t.* $\longrightarrow$ if $v {\downarrow} w$, i.e. if $v \xrightarrow{*} \circ \xleftarrow{*} w$. $\longrightarrow$ is *locally confluent* if $v {\downarrow} w$ for any v, w with $v \longleftarrow \circ \longrightarrow w$; it is *confluent* if $v {\downarrow} w$ for any v, w with $v \xleftarrow{*} \circ \xrightarrow{*} w$. a' is an $\longrightarrow$*-normal form of* a if $a \xrightarrow{*} a' \notin \mathrm{dom}(\longrightarrow)$.

A sequence $(s_i)_{i \in \mathbf{N}}$ is *non-terminating in* $\longrightarrow$ if $s_i \longrightarrow s_{i+1}$ for all $i \in \mathbf{N}$. $\longrightarrow$ is *terminating* if there are no non-terminating sequences in $\longrightarrow$. A relation R (on A) is *well-founded* if any non-empty class B ($\subseteq A$) has an R-minimal element, i.e. $\exists a \in B.\ \neg \exists a' \in B.\ a' R\, a$. Note that well-foundedness of $\longleftarrow$ immediately entails termination of $\longrightarrow$ (via the range of the non-terminating sequence), but the converse requires a weak form of the Axiom of Choice to construct the non-terminating sequence, cf. e.g. [Moore & Wirth, 2014, § 4.1].

Corollary 2.1 *If a binary relation is well-founded, so is its transitive closure.*

2.2 A Generalized Theorem as the Main Tool

The following Theorem 2.2 is a generalization of Jan Willem Klop's Theorem I.5.18 [Klop, 1980, p. 53], which can be obtained again from Theorem 2.2 by the specialization $\longrightarrow_0 := \emptyset$.

Theorem 2.2
Let $\longrightarrow_0$ and $\longrightarrow_1$ be two binary relations.
Set $\longrightarrow_2 := \overset{*}{\longrightarrow}_0 \circ \longrightarrow_1$.
Set $\longrightarrow_3 := \longrightarrow_0 \cup \longrightarrow_1$.
Let $a \in \mathrm{dom}(\longrightarrow_3)$. Let a' be an $\longrightarrow_3$-normal form of a. Set $A := \langle\{a\}\rangle\overset{*}{\longrightarrow}_3$.
Set $\longrightarrow_4 := {}_A{\upharpoonright}\longrightarrow_3$. If

1. *$\longleftarrow_0{\upharpoonright}_A$ is well-founded;*

2. *there is an upper bound $n \in \mathbf{N}$ on the length of $\longrightarrow_2$-derivations starting from a and reaching a' by $\overset{*}{\longrightarrow}_0$; more formally, this means that we have $m \leq n$ for any $m \in \mathbf{N}$ and any sequence $b_0, \ldots, b_m$ with $a = b_0$, $b_i \longrightarrow_2 b_{i+1}$ for each $i \in \{0, \ldots, m-1\}$, and $b_m \overset{*}{\longrightarrow}_0 a'$;*

3. *for all b_1, b_2 with $b_1 \longleftarrow_4 \circ \longrightarrow_1 b_2$, we have $b_1 \overset{*}{\longrightarrow}_4 \circ \overset{*}{\longleftarrow}_4 b_2$; and*

4. *for all b_1, b_2 with $b_1 \longleftarrow_4 \circ \longrightarrow_0 b_2$, we have $b_1 \overset{*}{\longrightarrow}_4 \circ \overset{=}{\longleftarrow}_4 b_2$;*

then $\longleftarrow_4$ is well-founded.

Proof of Theorem 2.2
Claim 1: For all b_1, b_2 and $n \in \mathbf{N}$ with $b_1 \longleftarrow_4 \circ \overset{n}{\longrightarrow}_0 b_2$, we have $b_1 \overset{*}{\longrightarrow}_4 \circ \overset{=}{\longleftarrow}_4 b_2$.
Proof of Claim 1: By induction on n. In case of $b_1 \longleftarrow_4 \circ \overset{0}{\longrightarrow}_0 b_2$, we have $b_1 \longleftarrow_4 b_2$.
In case of $b_1 \longleftarrow_4 \circ \overset{n}{\longrightarrow}_0 b_2 \longrightarrow_0 b_3$, by induction hypothesis we have $b_1 \overset{*}{\longrightarrow}_4 b_4 \overset{=}{\longleftarrow}_4 b_2$ for some $b_4 \in A$. In case of $b_4 = b_2$, we have $b_1 \overset{*}{\longrightarrow}_4 b_4 \longrightarrow_0 b_3$, and thus $b_1 \overset{*}{\longrightarrow}_4 b_3$. Otherwise, we have $b_4 \longleftarrow_4 b_2$, and thus $b_4 \overset{*}{\longrightarrow}_4 b_5 \overset{=}{\longleftarrow}_4 b_3$ for some b_5 by item 4, i.e. the desired $b_1 \overset{*}{\longrightarrow}_4 b_5 \overset{=}{\longleftarrow}_4 b_3$. Q.e.d. (Claim 1)

Set $B := \{\ b \in A \mid b \overset{*}{\longrightarrow}_4 a'\ \}$.
By item 2, we can define a function $l : B \to \{\ m \in \mathbf{N} \mid m \leq n\ \}$ via

$$l(b) := \max\{\ m \in \mathbf{N} \mid b \overset{m}{\longrightarrow}_2 \circ \overset{*}{\longrightarrow}_0 a'\ \}.$$

Claim 2: For all $b \in B$ with $b \overset{*}{\longrightarrow}_4 b'$, we have $b' \in B$.

Proof of Claim 2: By induction on $k := l(b)$ in $<$. The induction hypothesis is that for all $b'' \in B$ with $b'' \overset{*}{\longrightarrow}_4 b'''$ and $l(b'') < k$, we have $b''' \in B$. Note that (for $b'' \in B$) $b'' \longrightarrow_4 b'''$ implies $l(b''') \leq l(b'')$. Thus, by another induction on the length of derivations, the induction conclusion follows from the induction hypothesis and the proposition that for all $b'' \in B$ with $b'' \longrightarrow_4 b'''$ and $l(b'') = k$, we have $b''' \in B$.
So let us assume $b \in B$ and $b \longrightarrow_4 b'$. Then, using the induction hypothesis, we have to show $b' \in B$, for which it suffices to show $b' \overset{*}{\longrightarrow}_4 a'$.
By our assumption, we have $b \overset{*}{\longrightarrow}_4 a'$, which falls into at least one of the following two cases:

$b \overset{*}{\longrightarrow}_0 a'$: By Claim 1: $b' \overset{*}{\longrightarrow}_4 \circ \overset{=}{\longleftarrow}_4 a'$. Because $a' \notin \mathrm{dom}(\longrightarrow_3)$, and a fortiori also $a' \notin \mathrm{dom}(\longrightarrow_4)$, we actually have $b' \overset{*}{\longrightarrow}_4 a'$.

$b \overset{*}{\longrightarrow}_0 \hat{b} \longrightarrow_1 b''' \overset{*}{\longrightarrow}_4 a'$ for some $\hat{b}$, b''' : Again by Claim 1, we get $b' \overset{*}{\longrightarrow}_4 b'''' \overset{=}{\longleftarrow}_4 \hat{b}$ for some $b'''' \in A$.
In case of $b'''' = \hat{b}$, we have $b' \overset{*}{\longrightarrow}_4 b'''' \longrightarrow_1 b''' \overset{*}{\longrightarrow}_4 a'$, i.e. the desired $b' \overset{*}{\longrightarrow}_4 b'''' \longrightarrow_4 b''' \overset{*}{\longrightarrow}_4 a'$.
Otherwise we have $b'''' \longleftarrow_4 \hat{b}$. Thus, by item 3, there is some b'' with $b'''' \overset{*}{\longrightarrow}_4 b'' \overset{*}{\longleftarrow}_4 b'''$. Because of $b \overset{*}{\longrightarrow}_0 \hat{b} \longrightarrow_1 b''' \overset{*}{\longrightarrow}_4 a'$ we have $b''' \in B$ and $l(b''') < l(b)$. Thus, by the induction hypothesis, we get $b'' \in B$, and then the desired $b' \overset{*}{\longrightarrow}_4 b'''' \overset{*}{\longrightarrow}_4 b'' \overset{*}{\longrightarrow}_4 a'$. Q.e.d. (Claim 2)

Claim 3: $A = B$.

Proof of Claim 3: By $a \overset{*}{\longrightarrow}_3 a'$, we also have $a \overset{*}{\longrightarrow}_4 a'$, and so $a \in B$.
Thus, by Claim 2, we get $\langle\!\{a\}\!\rangle \overset{*}{\longrightarrow}_4 \subseteq B$.
All in all, we get: $A = \langle\!\{a\}\!\rangle \overset{*}{\longrightarrow}_3 = \langle\!\{a\}\!\rangle \overset{*}{\longrightarrow}_4 \subseteq B \subseteq A$. Q.e.d. (Claim 3)

By Claim 4, we get $l : A \to \{\, m \in \mathbf{N} \mid m \leq n \,\}$. Now for every b_1, b_2 with $b_1 \longleftarrow_4 b_2$, we have $b_1, b_2 \in A$ and, moreover, $(l(b_1), b_1)$ is strictly smaller than $(l(b_2), b_2)$ in the lexicographic combination of $<$ and $\longleftarrow_0 {\restriction_A}$, which is well-founded by item 1. Indeed, in case of $b_1 \longleftarrow_0 b_2$, we have $l(b_1) \leq l(b_2)$ and $b_1 \longleftarrow_0 {\restriction_A} b_2$, and in case of $b_1 \longleftarrow_1 b_2$, we have $l(b_1) < l(b_2)$.

Q.e.d. (Theorem 2.2)

2.3 Terms, Formulas, Substitutions, Contexts

A straightforward intuitive understanding of terms, formulas, substitutions, and contexts will actually suffice for most working mathematicians to understand the remainder of this paper. For the others, we give an example formalization of these notions here.

Terms and *formulas* are defined inductively as follows:

- An individual variable is a term.
- If A is an n-ary formula variable ($n \in \mathbf{N}$) and $t_1, \ldots, t_n$ are terms, then $A(t_1, \ldots, t_n)$ is a formula.
- If f is an n-ary constant function or predicate symbol ($n \in \mathbf{N}$) and $t_1, \ldots, t_n$ are terms,
  then $\mathrm{f}(t_1, \ldots, t_n)$ is a term or formula, respectively.
  In case of $n = 0$, we simply write "f" instead of "f()".
- If F is a formula, then $\neg F$ is a formula. If F_1 and F_2 are formulas, then $(F_1 \vee F_2)$, $(F_1 \wedge F_2)$, $(F_1 \Rightarrow F_2)$, $\ldots$ are formulas.
- If x is an individual variable and F is a formula,
  then $\varepsilon x.\ F$ is a term and $\exists x.\ F$ and $\forall x.\ F$ are formulas.
  In these terms and formulas, all occurrences of x are *bound*; non-bound occurrences of variables in terms and formulas are called *free,* such as each occurrence of any formula variable, and also of any individual variable y that is not in the scope of a binder on y, such as "$\varepsilon y.$", "$\exists y.$", or "$\forall y.$".

In our definition of terms and formulas we deviate from HILBERT–BERNAYS in not having an extra set of individual variables for bound occurrences, disjoint from the set to be used for free occurrences. So we have only one set of individual variables, but this does not really make any difference here, in particular because we ignore the variable names in the bound occurrences by the following stipulation:

We equate formulas modulo the renaming of bound variables.

A *substitution* is a mapping of individual variables to terms and of n-ary formula variables to expressions of the form $\underline{\lambda}(x_1, \cdots, x_n).\ F$, respectively, where $x_1, \cdots, x_n$ are mutually distinct individual variables and F is a formula. For $n = 0$, we just write "F" instead of "$\underline{\lambda}().\ F$".

Presupposing the above stipulation of considering formulas only up to renaming of bound variables, we now define the result of an application of a substitution σ to terms and formulas inductively as follows. We use postfix notation with highest operator precedence.

- Let x be an individual variable.
  If $x \notin \mathrm{dom}(\sigma)$, then $x\sigma = x$; otherwise $x\sigma = \sigma(x)$, i.e. the value of x under σ.

- Let A be an n-ary formula variable, and let $t_1, \ldots, t_n$ be terms. If $A \notin \mathrm{dom}(\sigma)$, then $(A(t_1, \ldots, t_n))\sigma = A(t_1\sigma, \ldots, t_n\sigma)$. Otherwise $(A(t_1, \ldots, t_n))\sigma$ is the result of the β-reduction of $\sigma(A)(t_1\sigma, \ldots, t_n\sigma)$, i.e., for $\sigma(A) = \underline{\lambda}(x_1, \cdots, x_n).\ F$, the formula $F\sigma'$, where σ' is the substitution $\{x_1 \mapsto t_1\sigma,\ \ldots,\ x_n \mapsto t_n\sigma\}$.

- If f is an n-ary constant function or predicate symbol and $t_1, \ldots, t_n$ are terms, then $(\mathrm{f}(t_1, \ldots, t_n))\sigma = \mathrm{f}(t_1\sigma, \ldots, t_n\sigma)$.

- If F is a formula, then $(\neg F)\sigma = \neg F\sigma$. If F_1 and F_2 are formulas, then $(F_1 \vee F_2)\sigma = (F_1\sigma \vee F_2\sigma)$, $(F_1 \wedge F_2)\sigma = (F_1\sigma \wedge F_2\sigma)$, $(F_1 \Rightarrow F_2)\sigma = (F_1\sigma \Rightarrow F_2\sigma)$, ...

- If x is an individual variable
  — w.l.o.g. neither an element of $\mathrm{dom}(\sigma)$, nor occurring (free) in $\mathrm{ran}(\sigma)$ —
  and F is a formula,
  then $(\varepsilon x.\ F)\sigma = \varepsilon x.\ F\sigma$, $(\exists x.\ F)\sigma = \exists x.\ F\sigma$, $(\forall x.\ F)\sigma = \forall x.\ F\sigma$.

Corollary 2.3 *If X is an individual variable or a nullary formula variable, and σ is a substitution, then for any formula or term G whose free variables are in A:*
$$G\sigma = G({}_{A}|\sigma).$$

By induction on the construction of G_1 we easily get:

Corollary 2.4
For any term or variable G_1, any X and G_2 being either an individual variable and a term, or a nullary formula variable and a formula, and any substitution σ where $X \notin \mathrm{dom}(\sigma)$ and X does not occur (free) in $\mathrm{ran}(\sigma)$:
$$(G_1\{X \mapsto G_2\})\sigma = (G_1\sigma)\{X \mapsto G_2\sigma\}.$$

Finally, let $H_0, \ldots, H_n$ $(n \in \mathbf{N})$ be mutually distinct, nullary formula variables, reserved for the following definition: A *context* written "$G[\cdots]$" (a formula or term with holes) is actually a formula or term G with one single (free) occurrence of each of the formula variables $H_1, \ldots, H_n$. Moreover, "$G[F_1, \ldots, F_n]$" denotes $G\{H_1 \mapsto F_1,\ \ldots,\ H_n \mapsto F_n\}$, for formulas $F_1, \ldots, F_n$.

Corollary 2.5
For any context $G[\cdots]$, and any formula F, and any substitution σ :
$$(G[F])\sigma = G\sigma[F\sigma].$$

3 The Concrete Rewrite Relation

By writing "$\neg^{\forall}$" for "$\neg$" and "$\neg^{\exists}$" for the empty string "", we can unify the two formulas (ε_1) and (ε_2) to the single formula

$$Qx.\ A \quad \Leftrightarrow \quad A\{x \mapsto \varepsilon x.\ \neg^{Q} A\} \tag{ε_Q}$$

for $Q \in \{\exists, \forall\}$, and x a meta-variable for an individual variable, and A a meta-variable for a formula.

Let $\longrightarrow$ be the rewrite relation resulting from rewriting with the equivalence (ε_Q) as a rewrite rule from left to right. Explicitly, this means that $F_1 {\longrightarrow} F_2$ if there are a context $G[\cdots]$, a quantifier symbol Q, an individual variable x, and a formula A, such that $F_1 = G[Qx.\ A]$ and $F_2 = G[A\{x \mapsto \varepsilon x.\ \neg^{Q} A\}]$.

Let $\longrightarrow_0$ and $\longrightarrow_1$ be the partition of $\longrightarrow$ for the case of a *vacuous* quantifier (i.e. for the case that x does not occur in the formula A in (ε_Q)), and for the case that the quantifier is not vacuous.

Let $\longrightarrow_{\mathcal{I}}$ be the innermost rewrite relation given by rewriting with the equivalence (ε_Q).

Let $\twoheadrightarrow\!\!\!\!\!\!\shortmid\shortmid$ be the version of $\longrightarrow$ for the rewriting of parallel redexes. Explicitly, this means that $F_1 \twoheadrightarrow\!\!\!\!\!\!\shortmid\shortmid F_2$ if there are a context $G[\cdots]$ with $n \in \mathbf{N}$ holes, quantifier symbols $Q_1, \ldots, Q_n$, individual variables $x_1, \ldots, x_n$, and formulas $A_1, \ldots, A_n$, such that

$$\begin{array}{rcl} F_1 & = & G[Q_1 x_1.\ A_1,\ \ \ldots,\ \ Q_n x_n.\ A_n], \\ F_2 & = & G[A_1\{x_1 \mapsto \varepsilon x_1.\ \neg^{Q_1} A_1\},\ \ \ldots,\ \ A_n\{x_n \mapsto \varepsilon x_n.\ \neg^{Q_n} A_n\}]. \end{array}$$

From these definitions, we immediately get the following corollaries.

Corollary 3.1 $\longrightarrow_{\mathcal{I}} \subseteq \longrightarrow$.

Corollary 3.2 $\twoheadrightarrow\!\!\!\!\!\!\shortmid\shortmid \subseteq \overset{*}{\longrightarrow}$.

3.1 Local Confluence

Note that the technical terms of the following lemma are clarified and formalized in its proof.

Lemma 3.3 *If we have a peak* $F_1 \longleftarrow F_0 \longrightarrow F_2$ *of local divergence and the redex of the rewrite step to* F_1 *is properly inside the one of the rewrite step to* F_2 *(which is on top of* F_0*), then there are formulas* F_3, F_4 *satisfying all the following items:*

1. $F_1 \longrightarrow F_4 \longleftarrow F_3 \mathrel{\leftarrow\!\!\shortparallel\!\!-} F_2$.

2. *If the initial step to the left is actually applied to a non-vacuous quantifier (i.e. if* $F_1 \longleftarrow_1 F_0$*), then we have* $F_4 \longleftarrow_1 F_3 \mathrel{\leftarrow\!\!\shortparallel\!\!-}_1 F_2$.

3. *If the initial step to the right is actually applied to a non-vacuous quantifier (i.e. if* $F_0 \longrightarrow_1 F_2$*), then we have* $F_1 \longrightarrow_1 F_4$.

4. *If the initial step to the right is actually applied to a vacuous quantifier (i.e. if* $F_0 \longrightarrow_0 F_2$*), then we have* $F_3 = F_2$.

Proof of Lemma 3.3

Suppose we have a peak $F_1 \longleftarrow F_0 \longrightarrow F_2$ of local divergence and the redex of the rewrite step to F_1 is properly inside the one of the rewrite step to F_2, which is on top of F_0. Then F_0 has the form

$$Q_1 x_1.\ G_1[Q_2 x_2.\ G_2]. \tag{F_0}$$

We may in particular assume here that x_2 is different from x_1 and does not occur free in the context $G_1[\cdots]$ if we consider the dots "$\cdots$" to be empty. Moreover we may assume that the formulas F_1 and F_2 are the following:

$$Q_1 x_1.\ G_1[G_2\{x_2 \mapsto \varepsilon x_2.\ \neg^{Q_2} G_2\}]. \tag{F_1}$$

$$\left(\ G_1[Q_2 x_2.\ G_2]\ \right)\left\{\ x_1 \mapsto \varepsilon x_1.\ \neg^{Q_1} G_1[Q_2 x_2.\ G_2]\ \right\}. \tag{F_2}$$

If we rewrite the outermost redex in F_1, we obtain the formula

$$\left(\ G_1[G_2\{x_2 \mapsto \varepsilon x_2.\ \neg^{Q_2} G_2\}]\ \right)\sigma$$

written with the help of the substitution σ given as

$$\left\{\ x_1 \mapsto \varepsilon x_1.\ \neg^{Q_1} G_1[G_2\{x_2 \mapsto \varepsilon x_2.\ \neg^{Q_2} G_2\}]\ \right\}. \tag{σ}$$

If we propagate this substitution, by Corollary 2.5 we obtain a formula given by the context

$$G_1\sigma[\cdots] \tag{C}$$

where we read the dots "$\cdots$" as

$$(G_2\{x_2 \mapsto \varepsilon x_2.\ \neg^{Q_2} G_2\})\sigma.$$

Because x_2 occurs free in none of $\mathrm{dom}(\sigma)$, $G_1[\cdots]$, $G_1[G_2\{x_2 \mapsto \varepsilon x_2.\ \neg^{Q_2} G_2\}]$, $\mathrm{ran}(\sigma)$, by Corollary 2.4 we can propagate σ further to write the inner formula as

$$G_2\sigma\{x_2 \mapsto \varepsilon x_2.\ \neg^{Q_2} G_2\sigma\}. \tag{I}$$

Putting (C) and (I) together again, we can choose formula F_4 with the property $F_1 \longrightarrow F_4$ as follows:

$$G_1\sigma\left[\ G_2\sigma\{x_2 \mapsto \varepsilon x_2.\ \neg^{Q_2} G_2\sigma\}\ \right]. \tag{F_4}$$

If we now rewrite all occurrences of the redex mentioned at the end of the notation of the formula F_2 in parallel, then we obtain the formula

$$\big(\ G_1[Q_2x_2.\ G_2]\ \big)\sigma.$$

Before we can rewrite the remaining redex, we have to propagate σ to obtain a clear description of it. By Corollary 2.5, this results again in a context as given in (C) above, where, however, we now read the "$\cdots$" as

$$Q_2x_2.\ G_2\sigma.$$

Note that, in this formula, the substitution σ has passed the quantifier "$Q_2x_2.$" soundly. Indeed, as mentioned above, x_1 is different from x_2, and x_2 cannot occur free in $\mathrm{ran}(\sigma)$. Putting this formula and its context together again, we can choose as F_3 with the property $F_3 \leftarrow\!\!\!\!+\!\!\!+\!\!\!- F_2$ as follows:

$$G_1\sigma\big[\ Q_2x_2.\ G_2\sigma\ \big]. \tag{F_3}$$

If we now rewrite the remaining redex, we again obtain the formula F_4, as was to be shown for item 1.

For item 2, it suffices to note that, if x_2 occurs free in G_2, then x_2 also occurs free in $G_2\sigma$ because x_1 and x_2 are different.

For item 3, it suffices to note that, if x_1 occurs free in $G_1[Q_2x_2.\ G_2]$, then x_1 also occurs free in $G_1[G_2\{x_2 \mapsto \varepsilon x_2.\ \neg^{Q_2}G_2\}]$.

For item 4, it suffices to note that, if x_1 does not occur free in $G_1[Q_2x_2.\ G_2]$, then both F_2 and F_3 are actually $G_1[Q_2x_2.\ G_2]$. **Q.e.d. (Lemma 3.3)**

As overlaps are trivial and as peaks of local divergence with parallel redexes are joinable in one step at each side trivially, we get as a corollaries of Lemma 3.3(1,4) and Corollary 3.2:

Corollary 3.4 $\longrightarrow$ *is locally confluent.*

Corollary 3.5 *For all F_1, F_2 with $F_1 \longleftarrow \circ \longrightarrow_0 F_2$, we have $F_1 \overset{+}{\longrightarrow} \circ \longleftarrow F_2$.*

3.2 Well-Foundedness

As every $\longrightarrow_0$-step (vacuous quantifiers) and every $\longrightarrow_{\mathcal{I}}$-step (innermost quantifiers) reduces the number occurrences of quantifiers by 1, we have:

Corollary 3.6 $\longleftarrow_0 \cup \longleftarrow_{\mathcal{I}}$ *is well-founded.*

Theorem 3.7 $\longleftarrow$ *is well-founded.*

Proof of Theorem 3.7

Assume that B is a non-empty class. Then there is some $a \in B$. We just have to find an $\longleftarrow$-minimal element in B.

If a is $\longleftarrow$-minimal in B, then we have succeeded. Thus suppose that a is not $\longleftarrow$-minimal in B. Then $a \in \mathrm{dom}(\longrightarrow)$.

Set $A := \langle\!\langle\{a\}\rangle\!\rangle \overset{*}{\longrightarrow}$. Set $\longrightarrow_4 := {}_A\!\rceil\!\longrightarrow$. It now suffices to show that $\longleftarrow_4$ is well-founded (because an $\longleftarrow_4$-minimal element of $A \cap B$ is also an $\longleftarrow$-minimal element of B).

By Corollary 3.6, A has an $\longleftarrow_{\mathcal{I}}$-minimal element a'. As $a' \notin \mathrm{dom}(\longrightarrow)$ by Corollary 3.1, a' is an $\longrightarrow$-normal form of a. To obtain the well-foundedness of $\longleftarrow_4$, we are now going to apply Theorem 2.2.

Set $\longrightarrow_2 := \overset{*}{\longrightarrow}_0 \circ \longrightarrow_1$. Set $\longrightarrow_3 := \longrightarrow_0 \cup \longrightarrow_1$. Then $\longrightarrow = \longrightarrow_3$.

It now suffices to show items 1 to 4 of Theorem 2.2. Item 1 holds by Corollary 3.6. Item 3 holds by Corollary 3.4. Item 4 holds by Corollary 3.5. As the number of occurrences of the ε is invariant under $\longrightarrow_0$ and is increased at least by 1 by every $\longrightarrow_1$-step, it increases at least by 1 by every $\longrightarrow_2$-step. Thus, to satisfy item 2, we can choose the upper bound n to be the number of occurrences of ε in a' (minus the number in a). **Q.e.d. (Theorem 3.7)**

3.3 Confluence

By the Newman Lemma (cf. [Newman, 1942] or, for a formal proof, [Wirth, 2004, § 3.4]), we obtain from Corollary 3.4 and Theorem 3.7:

Theorem 3.8 $\longrightarrow$ *is confluent.*

3.4 On the Length of Derivations

By Theorems 3.7 and 3.8, we now know for certain that the rewrite relation is confluent and terminating (as its reverse is even well-founded), which means that we can eliminate the quantifiers in any order — but this does not mean that this is efficient.

Here is a serious warning to the contrary: The nesting depth of the occurrences of the ε-symbols introduced by the normalization can be exponential in the number of quantifiers in the input formula, and the number of steps of an outermost normalization is even higher and seems to be non-elementary, cf. [WIRTH, 2015, Example 4.7], [WIRTH, 2008, Example 8].

As any innermost rewrite step reduces the number of quantifiers exactly by 1, and as no rewrite step can reduce the number of quantifiers by more than 1, we immediately get:

Theorem 3.9
Let F be a formula with n quantifiers. Innermost rewriting of F by $\longrightarrow_{\mathcal{I}}$ obtains the (unique) $\longrightarrow$-normal form F' of F in exactly n steps, which is the minimal number of steps to reach F' by $\longrightarrow$ from F.

4 Conclusion

With Theorems 3.7 and 3.8, we have shown confluence and termination of the elimination of quantifiers via their explicit definition via HILBERT's ε. This means in particular that any first-order term with quantifiers and epsilons (and formula variables), has a unique normal form w.r.t. this elimination of quantifiers, which has its first explicit occurrence in [HILBERT & BERNAYS, 1939], namely in the proof of the 1st ε-theorem on Page 19f.

Moreover, the directness, self-containedness, and easy verifiability of the proofs should settle the questions on confluence and termination here once and for all — at least for working mathematicians. Formalists and rewriters, however, may see the need to develop a more formal verification of our proof and write a short paper that our results are all trivial in some higher-order rewriting theory. Writing or helping to find a good textbook on higher-order rewriting, however, seems to be in more urgent demand.

Furthermore, we hope that some philosophers will be stimulated by this paper to pick up the subject of the non-triviality of higher-order *explicit definition* and write or help to find a book on that subject.

Finally, the starting point of our interest in the subject, namely the question whether there is a lacuna in HILBERT–BERNAYS as discussed in § 1.3, needs further discussion by the experts on HILBERT's ε and the history of mathematical logic in the 20th century. On basis of our current knowledge, we would clearly answer this question positively.

References

[Ackermann, 1925] Wilhelm Ackermann. Begründung des „tertium non datur“ mittels der Hilbertschen Theorie der Widerspruchsfreiheit. *Mathematische Annalen*, 93:1–36, 1925. Received March 30, 1924. Inauguraldissertation, Göttingen 1924.

[Codish & Middeldorp, 2004] Michael Codish and Aart Middeldorp. 7th Int. Workshop on Termination (WST), 2004. Technical Report AIB–2004–07, RWTH Aachen, Dept. of Computer Sci., 2004. ISSN 0935–3232. http://sunsite.informatik.rwth-aachen.de/Publications/AIB/2004/2004-07.ps.gz.

[Gillman, 1987] Leonard Gillman. *Writing Mathematics Well.* The Mathematical Association of America, 1987.

[Heijenoort, 1971] Jean van Heijenoort. *From* Frege *to* Gödel*: A Source Book in Mathematical Logic, 1879–1931.* Harvard Univ. Press, 1971. 2nd rev. edn. (1st edn. 1967).

[Hilbert & Bernays, 1934] David Hilbert and Paul Bernays. *Grundlagen der Mathematik — Erster Band.* Number XL in Grundlehren der mathematischen Wissenschaften. Springer, 1934. 1st edn. (2nd edn. is [Hilbert & Bernays, 1968]). English translation is [Hilbert & Bernays, 2016a; 2016b].

[Hilbert & Bernays, 1939] David Hilbert and Paul Bernays. *Grundlagen der Mathematik — Zweiter Band.* Number L in Grundlehren der mathematischen Wissenschaften. Springer, 1939. 1st edn. (2nd edn. is [Hilbert & Bernays, 1970]).

[Hilbert & Bernays, 1968] David Hilbert and Paul Bernays. *Grundlagen der Mathematik I.* Number 40 in Grundlehren der mathematischen Wissenschaften. Springer, 1968. 2nd rev. edn. of [Hilbert & Bernays, 1934]. English translation is [Hilbert & Bernays, 2016a; 2016b].

[Hilbert & Bernays, 1970] David Hilbert and Paul Bernays. *Grundlagen der Mathematik II.* Number 50 in Grundlehren der mathematischen Wissenschaften. Springer, 1970. 2nd rev. edn. of [Hilbert & Bernays, 1939].

[Hilbert & Bernays, 2016a] David Hilbert and Paul Bernays. *Grundlagen der Mathematik I — Foundations of Mathematics I, Part A: Title Pages, Prefaces, and §§ 1–2.* Springer, 2016. First English translation and bilingual facsimile edn. of the 2nd German edn. [Hilbert & Bernays, 1968], incl. the annotation and translation of all differences of the 1st German edn. [Hilbert & Bernays, 1934]. Ed. by Claus-Peter Wirth, Jörg Siekmann, Michael Gabbay, Dov Gabbay. Advisory Board: Wilfried Sieg (chair), Irving H. Anellis, Steve Awodey, Matthias Baaz, Wilfried Buchholz, Bernd Buldt, Reinhard Kahle, Paolo Mancosu, Charles Parsons, Volker Peckhaus, William W. Tait, Christian Tapp, Richard Zach. Translated and commented by Claus-Peter Wirth &al. Thoroughly rev. 3rd edn. (1st edn. College Publications, London, 2011; 2nd edn. http://wirth.bplaced.net/p/hilbertbernays, 2013).

[Hilbert & Bernays, 2016b] David Hilbert and Paul Bernays. *Grundlagen der Mathematik I — Foundations of Mathematics I, Part B: §§ 3–5 and Deleted Part I of the*

1st Edn.. Springer, 2016. First English translation and bilingual facsimile edn. of the 2nd German edn. [HILBERT & BERNAYS, 1968], incl. the annotation and translation of all deleted texts of the 1st German edn. [HILBERT & BERNAYS, 1934]. Ed. by CLAUS-PETER WIRTH, JÖRG SIEKMANN, MICHAEL GABBAY, DOV GABBAY. Advisory Board: WILFRIED SIEG (chair), IRVING H. ANELLIS, STEVE AWODEY, MATTHIAS BAAZ, WILFRIED BUCHHOLZ, BERND BULDT, REINHARD KAHLE, PAOLO MANCOSU, CHARLES PARSONS, VOLKER PECKHAUS, WILLIAM W. TAIT, CHRISTIAN TAPP, RICHARD ZACH. Translated and commented by CLAUS-PETER WIRTH &AL. Thoroughly rev. 3rd edn. (1st edn. College Publications, London, 2012; 2nd edn. `http://wirth.bplaced.net/p/hilbertbernays`, 2013).

[HILBERT, 1926] David Hilbert. Über das Unendliche — Vortrag, gehalten am 4. Juni 1925 gelegentlich einer zur Ehrung des Andenkens an WEIERSTRASS von der Westfälischen Math. Ges. veranstalteten Mathematiker-Zusammenkunft in Münster i. W. *Mathematische Annalen*, 95:161–190, 1926. Received June 24, 1925. English translation *On the infinite* by STEFAN BAUER-MENGELBERG with an introduction by JEAN VAN HEIJENOORT in [HEIJENOORT, 1971, pp. 367–392].

[HILBERT, 1928] David Hilbert. Die Grundlagen der Mathematik — Vortrag, gehalten auf Einladung des Mathematischen Seminars im Juli 1927 in Hamburg. *Abhandlungen aus dem mathematischen Seminar der Univ. Hamburg*, 6:65–85, 1928. English translation *The foundations of mathematics* by STEFAN BAUER-MENGELBERG and DAGFINN FØLLESDAL with a short introduction by JEAN VAN HEIJENOORT in [HEIJENOORT, 1971, pp. 464–479].

[JOUANNAUD & OKADA, 1991] Jean-Pierre Jouannaud and Mitsuhiro Okada. A computation-model for executable higher-order algebraic specification languages. 1991. In [LICS, 1991, pp. 350–361].

[KETEMA & RAAMSDONK, 2004] Jeroen Ketema and Femke van Raamsdonk. Erasure and termination in higher-order rewriting. 2004. In [CODISH & MIDDELDORP, 2004, pp. 30–33]. Also available at `www.prove-and-die.org/publ/wst04.pdf`.

[KLOP &AL., 1993] Jan Willem Klop, Vincent van Oostrom, and Femke van Raamsdonk. Combinatory reduction systems: introduction and survey. *Theoretical Computer Sci.*, 121:279–308, 1993.

[KLOP, 1980] Jan Willem Klop. Combinatory reduction systems. Mathematical centre tracts 127, Mathematisch Centrum (since 1983: Centrum Wiskunde & Informatica), Amsterdam, 1980. PhD thesis, Utrecht Univ.., `http://libgen.io/get.php?md5=a80df1dda74dffd97700bda2277e8bc4`.

[LICS, 1991] *Proc. 6th Annual IEEE Symposium on Logic In Computer Sci. (LICS), Amsterdam, 1991*. IEEE Press, 1991. `http://lii.rwth-aachen.de/lics/archive/1991`.

[MIDDELDORP, 2001] Aart Middeldorp, editor. *12th Int. Conf. on Rewriting Techniques and Applications (RTA), Utrecht (The Netherlands), 2001*, number 2051 in Lecture Notes in Computer Science. Springer, 2001.

[MOORE & WIRTH, 2014] J Strother Moore and Claus-Peter Wirth. *Automation of Mathe-*

matical Induction as part of the History of Logic. SEKI-Report SR–2013–02 (ISSN 1437–4447). SEKI Publications, 2014. Rev. edn. July 2014 (1st edn. 2013), ii+107 pp., http://arxiv.org/abs/1309.6226.

[Narendran & Rusinowitch, 1999] Paliath Narendran and Michaël Rusinowitch, editors. *10th Int. Conf. on Rewriting Techniques and Applications (RTA), Trento (Italy), 1999*, number 1631 in Lecture Notes in Computer Science. Springer, 1999.

[Newman, 1942] Max H. A. Newman. On theories with a combinatorial definition of equivalence. *Annals of Mathematics*, 43:223–242, 1942.

[Nipkow, 1991] Tobias Nipkow. Higher-order critical pairs. 1991. In [LICS, 1991, pp. 342–349].

[Raamsdonk, 1999] Femke van Raamsdonk. Higher-order rewriting. 1999. In [Narendran & Rusinowitch, 1999, pp. 221–239].

[Raamsdonk, 2001] Femke van Raamsdonk. On termination of higher-order rewriting. 2001. In [Middeldorp, 2001, pp. 261–275].

[Raamsdonk, 2013] Femke van Raamsdonk, editor. *24th Int. Conf. on Rewriting Techniques and Applications (RTA), Einhoven (The Netherlands), 2013*, number 24 in Leibniz Int. Proc. in Informatics. Dagstuhl (Germany), 2013.

[Tait, 1967] William W. Tait. Intensional interpretations of functionals of finite type I. *J. Symbolic Logic*, 32:198–212, 1967.

[Winkler &al., 2013] Sarah Winkler, Harald Zankl, and Aart Middeldorp. Beyond Peano Arithmetic — automatically proving termination of the Goodstein sequence. 2013. In [Raamsdonk, 2013, pp. 335–351].

[Wirth, 2004] Claus-Peter Wirth. Descente Infinie + Deduction. *Logic J. of the IGPL*, 12:1–96, 2004. http://wirth.bplaced.net/p/d.

[Wirth, 2008] Claus-Peter Wirth. Hilbert's epsilon as an operator of indefinite committed choice. *J. Applied Logic*, 6:287–317, 2008. http://dx.doi.org/10.1016/j.jal.2007.07.009.

[Wirth, 2009] Claus-Peter Wirth. Shallow confluence of conditional term rewriting systems. *J. Symbolic Computation*, 44:69–98, 2009. http://dx.doi.org/10.1016/j.jsc.2008.05.005.

[Wirth, 2015] Claus-Peter Wirth. *A Simplified and Improved Free-Variable Framework for* Hilbert*'s epsilon as an Operator of Indefinite Committed Choice*. SEKI Report SR–2011–01 (ISSN 1437–4447). SEKI Publications, 2015. Rev. edn. June 2015 (1st edn. 2011), ii+82 pp., http://arxiv.org/abs/1104.2444.

Received \jreceived