Annals of Pure and Applied Logic 38 (1988) 123–134
North-Holland

THE SYNTAX OF NONSTANDARD ANALYSIS

Edward NELSON*
Department of Mathematics, Princeton University, Princeton, NJ 08544, USA

Communicated by H. Rogers, Jr.
Received 22 January 1984; revised 30 May 1985

> However, from a formalist point of view we may look at our theory syntactically and may consider that what we have done is to introduce *new deductive procedures* rather than new mathematical entities.
>
> Abraham Robinson [4]

1. The syntax of 'standard'

In [3] I presented a codification of a large part of nonstandard analysis, called Internal Set Theory (IST). Here we will show that proofs in IST may be regarded as abbreviations of proofs within conventional mathematics. That is, we analyze Robinson's 'new deductive procedures'. In doing so, we solve (to the extent that IST codifies nonstandard analysis) a problem posed by him in his retiring presidential address before the Association for Symbolic Logic [5]. Problem 11 reads, in part, "to devise a purely syntactical transformation which correlates standard and nonstandard proofs of the same theorems in a large area, e.g., complex function theory".

In IST a new predicate "x is standard" is adjointed to conventional mathematics, which we take to be Zermelo–Fraenkel set theory with the axiom of choice (ZFC). This new predicate is devoid of semantical content. That is to say, it is meaningless. The three axiom schemes of IST — namely, the transfer principle (T), the idealization principle (I), and the standardization principle (S) — give rules for manipulating the new predicate.

When working in the framework of IST, it is helpful to think of using 'standard' in very much the same way that 'fixed' is used in informal mathematical discourse. For example, let $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$ and consider

$$\textit{there is an } x \textit{ in } \mathbb{R}^+ \textit{ which is smaller than any fixed } \varepsilon \textit{ in } \mathbb{R}^+. \tag{1}$$

This is a common way of expressing a true statement about $\mathbb{R}^+$. But suppose we go on to say that such an x is called infinitesimal; that is, we define a number x in

* This research was supported in part by NSF grant MCS-7801863.

$\mathbb{R}^+$ to be infinitesimal in case

$$x \text{ is smaller than any fixed } \varepsilon \text{ in } \mathbb{R}^+ \tag{2}$$

and rewrite (1) as

$$\text{there is an infinitesimal } x \text{ in } \mathbb{R}^+. \tag{3}$$

The likely response to this would be: You have misunderstood the syntax of the informal statement (1); although it has the form

$$\exists x\, \forall \varepsilon\, (\varepsilon \text{ fixed} \Rightarrow x < \varepsilon), \tag{4}$$

what it means is

$$\forall \varepsilon\, \exists x\, x < \varepsilon. \tag{5}$$

To this the practitioner of IST responds: We agree that (1) and (5) mean the same thing, but that is precisely what (3) and (4) mean. By adjoining the predicate "x is standard" (or "x is fixed") to conventional mathematics, we are able to speak about infinitesimals in the ordinary system $\mathbb{R}^+$ without adjoining any new mathematical entities. Formulas in IST with free variables, such as (2), may not have any meaning in conventional terms, but every statement (formula with no free variables) in IST does have meaning in conventional terms — this is what the reduction algorithm of [3] (which is described more simply in Section 2 below) accomplishes, reducing (4) to (5), for example.

The point of view towards IST which we advocate is that external statements be regarded as abbreviations for certain internal statements. The reduction algorithm codifies the process of finding the equivalent internal statement, treating 'standard' syntactically like 'fixed' to reorder the quantifiers. Proofs in IST differ from standard proofs in that they treat "x is fixed" as a predicate.

2. Reduction of formulas

Let $\mathscr{A}$ be a first-order theory, with propositional connectives $\Rightarrow$ and $\sim$, universal quantifier $\forall$, predicate letters A_j^n, and individual constants a_i. The other propositional connectives and the existential quantifier are introduced as abbreviations, and the notions of formula and statement are defined in the usual way. Later we will give the forms of the logical axioms and rules of inference.

The reduction algorithm increases the logical type of formulas, so we will need a richer language, the language of a model of $\mathscr{A}$. Let $\bar{\mathscr{A}}$ be ZFC together with constants M, $\bar{A}_j^n$, $\bar{a}_i$ and the axioms which assert that these are a model of $\mathscr{A}$. Let $\hat{M}$ be the smallest set such that $M \in \hat{M}$ and whenever $X, Y \in \hat{M}$, then $P_{\text{fin}}(X) \in \hat{M}$ and $X^Y \in \hat{M}$, where $P_{\text{fin}}(X)$ is the set of all finite subsets of X and X^Y is the set of all functions from Y to X. By a formula of $\bar{\mathscr{A}}$ we will always mean a formula in which each variable is restricted to range over a certain element of $\hat{M}$; it will not

be necessary for our purposes to give a detailed formlization of $\tilde{\mathscr{A}}$ as a first-order system with different types of variables.

Let $\mathscr{A}^*$ be $\tilde{\mathscr{A}}$ together with the unary predicate letter 'standard' and the axioms stated below. We use lower case Latin letters to denote n-tuples of variables, including the possibility $n=0$ We use capital Greek letters for formulas of $\mathscr{A}^*$ and lower case Greek letters for internal formulas of $\mathscr{A}^*$ (formulas of $\tilde{\mathscr{A}}$). In addition, we use the rather self-explanatory abbreviations of [3]. The additional axioms of $\mathscr{A}^*$ are the axioms asserting that M, $\tilde{A}_j^n$, and $\tilde{a}_i$ are standard and the axiom schemes

(T) $\quad \forall^{\mathrm{st}} t\,(\forall^{\mathrm{st}} x\, \phi(x, t) \Rightarrow \forall x\, \phi(x, t))$

where $\phi(x, t)$ has no undisplayed free variables,

(I) $\quad \exists x\, \forall^{\mathrm{st}} y\, \phi(x, y) \Leftrightarrow \forall^{\mathrm{st\,fin}} y'\, \exists x\, \forall y \in y'\, \phi(x, y)$,

(S_0) $\quad \forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} y\, \phi(x, y) \Rightarrow \exists^{\mathrm{st}} \bar{y}\, \forall^{\mathrm{st}} x\, \phi(x, \bar{y}(x))$.

We call (S_0) the *restricted standardization principle* because ϕ is restricted to be internal. We will show in Section 4 that the unrestricted standardized principle

(S′) $\quad \forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} y\, \Phi(x, y) \Rightarrow \exists^{\mathrm{st}} \bar{y}\, \forall^{\mathrm{st}} x\, \Phi(x, \bar{y}(x))$

is a consequence.

Let Φ be a formula of $\mathscr{A}^*$. We shall define what is meant by the *partial reduction* Φ' of Φ. As a preliminary, we first rewrite Φ so that every occurrence of "x standard" is replaced by "$\exists^{\mathrm{st}} y\, y = x$" and so that no variable x is quantified twice.

The partial reduction Φ' will always be of the form $\forall^{\mathrm{st}} u\, \exists^{\mathrm{st}} v\, \phi(u, v)$. If Φ is internal we define Φ' to be Φ. Suppose that Φ has the partial reduction $\forall^{\mathrm{st}} u\, \exists^{\mathrm{st}} v\, \phi(u, v)$, and similarly for Φ_1 and Φ_2. We define inductively

$$(\sim\Phi)' \quad \text{to be} \quad \forall^{\mathrm{st}} \bar{v}\, \exists^{\mathrm{st}} u \sim \phi(u, \bar{v}(u)),$$

$$(\Phi_1 \Rightarrow \Phi_2)' \quad \text{to be} \quad \forall^{\mathrm{st}} u_2 \forall^{\mathrm{st}} \bar{v}_1 \exists^{\mathrm{st}} u_1 \exists^{\mathrm{st}} v_2 (\phi_1(u_1, \bar{v}_1(u_1)) \Rightarrow \phi_2(u_2, v_2)),$$

$$(\forall x\, \Phi)' \quad \text{to be} \quad \forall^{\mathrm{st}} u\, \exists^{\mathrm{st\,fin}} v'\, \forall x\, \exists v \in v'\, \phi(u, v),$$

$$(\forall^{\mathrm{st}} x\, \Phi)' \quad \text{to be} \quad \forall^{\mathrm{st}} x\, \forall^{\mathrm{st}} u\, \exists^{\mathrm{st}} v\, \phi(u, v).$$

By (I) and (S_0) it is a theorem of $\mathscr{A}^*$ that $\Phi \Leftrightarrow \Phi'$. Notice that Φ' has the same free variables as Φ and so is a statement if and only if Φ is a statement. If Φ is a statement whose partial reduction Φ' is $\forall^{\mathrm{st}} u\, \exists^{\mathrm{st}} v\, \phi(u, v)$ we define its *reduction* Φ° to be $\forall u\, \exists v\, \phi(u, v)$. By (T), it is a theorem of $\mathscr{A}^*$ that $\Phi \Leftrightarrow \Phi^\circ$. Notice that Φ° is an internal statement; that is, a statement of $\tilde{\mathscr{A}}$.

3. Reduction of proofs

The system $\mathscr{A}^*$ is a conservative extension of $\tilde{\mathscr{A}}$. This is because the additional axioms (T), (I) and (S_0) express properties of adequate ultralimits, and this

construction may be carried out within $\tilde{\mathscr{A}}$ (see for example the Appendix to [3]). This is a semantical approach to the relationship between nonstandard analysis and conventional mathematics.

Our aim in this paper is to give a purely syntactical algorithm which takes a proof in $\mathscr{A}^*$ and converts it step by step into a proof within $\tilde{\mathscr{A}}$. One difficulty which we face at the outset is that formulas with free variables in $\mathscr{A}^*$ have no interpretation within $\tilde{\mathscr{A}}$, and one of the rules of inference — generalization — is used for such formulas. Fortunately there is a formulation (see [2, Exercise 2, p. 62]) of the notion of a first-order theory in which modus ponens is the only rule of inference. In this formulation the proper axioms are any collection of statements and the logical axioms are all closures of formulas of the following form:

(i) $\Phi \Rightarrow (\Psi \Rightarrow \Phi)$,
(ii) $(\Phi \Rightarrow (\Psi \Rightarrow \Lambda)) \Rightarrow ((\Phi \Rightarrow \Psi) \Rightarrow (\Phi \Rightarrow \Lambda))$,
(iii) $(\sim\Psi \Rightarrow \sim\Phi) \Rightarrow ((\sim\Psi \Rightarrow \Phi) \Rightarrow \Psi)$,
(iv) $\forall x\, \Phi(x) \Rightarrow \Phi(t)$ where t is an individual constant or variable,
(v) $\forall y\, (\Phi \Rightarrow \Psi) \Rightarrow (\Phi \Rightarrow \forall y\, \Psi)$ if Φ has no free occurrence of y,
(vi) $\forall y\, (\Phi \Rightarrow \Psi) \Rightarrow (\forall y\, \Phi \Rightarrow \forall y\, \Psi)$.

The usual formulation of the notion of a first-order theory [2] has as its rules of inference modus ponens and generalization, for proper axioms any collection of formulas, and for logical axioms (i)–(v) (rather than the closures of (i)–(vi)). It is easy to see that a statement is a theorem according to the usual formulation if and only if it is a theorem according to the above formulation.

The following theorem is used in reducing modus ponens from $\mathscr{A}^*$ to $\tilde{\mathscr{A}}$:

Theorem 1. *Let Φ_1 and Φ_2 be statements of $\mathscr{A}^*$. Then $(\Phi_1 \Rightarrow \Phi_2)^\circ \Rightarrow (\Phi_1^\circ \Rightarrow \Phi_2^\circ)$ is a theorem of $\tilde{\mathscr{A}}$.*

Proof. $(\Phi_1 \Rightarrow \Phi_2)^\circ$ is

$$\forall u_2\, \forall \tilde{v}_1\, \exists u_1\, \exists v_2\, (\phi_1(u_1, \tilde{v}_1(u_1)) \Rightarrow \phi_2(u_2, v_2)),$$

which is equivalent to

$$\exists \tilde{v}_1\, \forall u_1\, \phi_1(u_1, \tilde{v}_1(u_1)) \Rightarrow \forall u_2\, \exists v_2\, \phi_2(u_2, v_2), \tag{6}$$

while $\Phi_1^\circ \Rightarrow \Phi_2^\circ$ is

$$\forall u_1\, \exists v_1\, \phi_1(u_1, v_1) \Rightarrow \forall u_2\, \exists v_2\, \phi_2(u_2, v_2). \tag{7}$$

Since the consequents of the implications (6) and (7) are the same, to show that (6) $\Rightarrow$ (7) we need only show that the antecedent $\forall u_1\, \exists v_1\, \phi_1(u_1, v_1)$ of (7) implies the antecedent $\exists \tilde{v}_1\, \forall u_1\, \phi_1(u_1, \tilde{v}_1(u_1))$ of (6). But this is a theorem of $\tilde{\mathscr{A}}$; namely, the axiom of choice. □

We will show, in Theorems 2, 3, and 4, that the reduction of every axiom of $\mathscr{A}^*$ is a theorem of $\tilde{\mathscr{A}}$. Then the description of our syntactical algorithm is very simple. Let $\Phi_1, \cdots, \Phi_n$ be a proof in $\mathscr{A}^*$. Thus each Φ_i is a statement of $\mathscr{A}^*$

which is either an axiom of $\mathscr{A}^*$ or follows by modus ponens from Φ_j and Φ_k with $j, k < i$. Suppose by induction on i that for all $j < i$ the statement Φ_j° is a theorem of $\tilde{\mathscr{A}}$. (This is vacuously true for $i = 1$.) We need to show that Φ_i° is a theorem of $\tilde{\mathscr{A}}$. If Φ_i is an axiom of $\mathscr{A}^*$ this is true by Theorem 2, 3, or 4. So suppose that Φ_i follows by modus ponens from Φ_j and Φ_k with $j, k < i$; that is, suppose that Φ_k is $\Phi_j \Rightarrow \Phi_i$. By the induction hypothesis, $(\Phi_j \Rightarrow \Phi_i)^\circ$ is a theorem of $\tilde{\mathscr{A}}$, so by Theorem 1 and modus ponens in $\tilde{\mathscr{A}}$ the statement $\Phi_j^\circ \Rightarrow \Phi_i^\circ$ is a theorem of $\tilde{\mathscr{A}}$. Since by the induction hypothesis Φ_j° is a theorem of $\tilde{\mathscr{A}}$, we have by modus ponens in $\tilde{\mathscr{A}}$ that Φ_i° is a theorem of $\tilde{\mathscr{A}}$, which completes the induction. Thus the entire proof reduces to a proof in $\tilde{\mathscr{A}}$.

We know from semantical considerations that the reduction of every axiom of $\mathscr{A}^*$ is a theorem of $\tilde{\mathscr{A}}$. The purpose of proving this directly in Theorems 2, 3, and 4 is to obtain a better understanding of what makes nonstandard analysis work.

The proof of Theorem 2 shows that the reductions of (T) and of the closures of (I) and (S_0) are trivial theorems of $\tilde{\mathscr{A}}$, being essentially tautologies of the form $\psi \Rightarrow \psi$. This is not surprising, since these axioms are just the rules for forming reductions.

The logical axioms are more interesting. The proof of Theorem 3 shows that the reduction of a tautology ((i), (ii), or (iii)) involves only the consideration of an explicit finite number of cases, and the reduction of the rule of specialization (iv) is trivial.

The only non-trivial case is (vi), considered in Theorem 4. (The axiom (v) is essentially a special case of (vi).) Mathematical reasoning is full of uses of modus ponens with parameters: from $\Phi(y)$ and $\Phi(y) \Rightarrow \Psi(y)$ infer $\Psi(y)$. Since we must deal only with statements, we must reformulate this as: from $\forall y\, \Phi(y)$ and $\forall y\, (\Phi(y) \Rightarrow \Psi(y))$ infer $\forall y\, \Psi(y)$. But (vi) (and modus ponens for statements, as discussed above) is precisely what is needed to reduce this inference from $\mathscr{A}^*$ to $\tilde{\mathscr{A}}$. Thus Theorem 4 is the heart of the matter: a simple inference in nonstandard analysis may, when expressed in standard mathematics, entail a complicated cross-section argument (see the proof of Theorem 4). This point is illustrated by example in Section 7.

For convenience in proving Theorem 2, we split the equivalence (I) into two implications (I_1) and (I_2).

Theorem 2. *The reductions of the additional axioms of $\mathscr{A}^*$, namely* "M is standard", "$\bar{A}_j^{\tau}$ is standard", "$\bar{a}_i$ is standard", *and*

(T) $\quad \forall^{\mathrm{st}} t\, (\forall^{\mathrm{st}} x\, \phi(x, t) \Rightarrow \forall y\, \phi(y, t))$,

(I_1) $\quad \forall w\, (\exists x\, \forall^{\mathrm{st}} y\, \phi(x, y, w) \Rightarrow \forall^{\mathrm{st\,fin}} b'\, \exists a\, \forall b \in b'\, \phi(a, b, w))$,

(I_2) $\quad \forall w\, (\forall^{\mathrm{st\,fin}} y'\, \exists x\, \forall y \in y'\, \phi(x, y, w) \Rightarrow \exists a\, \forall^{\mathrm{st}} b\, \phi(a, b, w))$,

(S_0) $\quad \forall w\, (\forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} y\, \phi(x, y, w) \Rightarrow \exists^{\mathrm{st}} \tilde{b}\, \forall^{\mathrm{st}} a\, \phi(a, \tilde{b}(a), w))$,

where each ϕ is internal with no undisplayed free variables, are theorems of $\tilde{\mathscr{A}}$.

Proof. The reduction of "M is standard" is the theorem $\exists x\, x = M$, and similarly for $\bar{A}^i_n$ and $\bar{a}_i$. The statement (T)° is the theorem $\forall t\, \exists x\, (\phi(x, t) \Rightarrow \forall y\, \phi(y, t))$. Next, $(\mathrm{I}_0)^\circ$ is

$$\forall^{\mathrm{fin}} b'\, \exists^{\mathrm{fin}} y''\, \forall w\, \exists^{\mathrm{fin}} y' \in y''\, (\exists x\, \forall y \in y'\, \phi(x, y, w) \Rightarrow \exists a\, \forall b \in b'\, \phi(a, b, w)). \quad (8)$$

(Strictly speaking, $(\mathrm{I}_1)^\circ$ is the above statement with $\exists x\, \forall y \in y'\, \phi(x, y, w)$ replaced by $\sim\forall x\, \exists y \in y' \sim\phi(x, y, w)$, but the two forms are of course equivalent in $\mathcal{S}\bar{\mathcal{A}}$. We will not make similar comments in the future.) Let $y'' = \{b'\}$ and $y' = b'$. Then (8) becomes the theorem

$$\forall^{\mathrm{fin}} b'\, \forall w\, (\exists x\, \forall y \in b'\, \phi(x, y, w) \Rightarrow \exists a\, \forall b \in b'\, \phi(a, b, w)).$$

We find that $(\mathrm{I}_2)^\circ$ is identical with $(\mathrm{I}_1)^\circ$, and so is a theorem. Finally, $(\mathrm{S}_0)^\circ$ is

$$\forall \bar{a}\, \forall \bar{y}\, \exists^{\mathrm{fin}} x'\, \exists^{\mathrm{fin}} \bar{b}'\, \forall w\, \exists x \in x'\, \exists \bar{b} \in \bar{b}'\, (\phi(x, \bar{y}(x), w) \Rightarrow \phi(\bar{a}(\bar{b}), \bar{b}(\bar{a}(\bar{b})), w)).$$

Let $\bar{b}' = \{\bar{y}\}$, $\bar{b} = \bar{y}$, $x' = \{\bar{a}(\bar{y})\}$, $x = \bar{a}(\bar{y})$. Then this becomes the theorem

$$\forall \bar{a}\, \forall \bar{y}\, \forall w\, (\phi(\bar{a}(\bar{y})), \bar{y}(\bar{a}(\bar{y})), w) \Rightarrow \phi(\bar{a}(\bar{y}), \bar{y}(\bar{a}(\bar{y})), w)). \qquad \square$$

Theorem 3. *The reductions of the closures of* (i), (ii), (iii), *and* (iv) *are theorems of* $\mathcal{S}\bar{\mathcal{A}}$.

Proof. The closure of (i) may be written

$$\forall w\, (\forall^{\mathrm{st}} a\, \exists^{\mathrm{st}} b\, \phi(a, b, w) \Rightarrow (\forall^{\mathrm{st}} c\, \exists^{\mathrm{st}} d\, \psi(c, d, w) \Rightarrow \forall^{\mathrm{st}} e\, \exists^{\mathrm{st}} f\, \phi(e, f, w)),$$

whose reduction is

$$\forall e\, \forall \bar{d}\, \forall \bar{b}\, \exists^{\mathrm{fin}} a'\, \exists^{\mathrm{fin}} c'\, \exists^{\mathrm{fin}} f'\, \forall w\, \exists a \in a'\, \exists c \in c'\, \exists f \in f'$$
$$(\phi(a, \bar{b}(a), w) \Rightarrow (\psi(c, \bar{d}(c), w) \Rightarrow \phi(e, f, w))).$$

Let $a' = \{e\}$, let c' be any non-empty finite set, let $f' = \{\bar{b}(e)\}$, $a = e$, $c \in c'$, and $f = \bar{b}(e)$. Then the statement becomes the closure of a tautology of the form $\phi \Rightarrow (\psi \Rightarrow \phi)$.

The closure of (ii) may be written

$$\forall w\, ((\forall^{\mathrm{st}} a\, \exists^{\mathrm{st}} b\, \phi(a, b, w) \Rightarrow (\forall^{\mathrm{st}} c\, \exists^{\mathrm{st}} d\, \psi(c, d, w) \Rightarrow \forall^{\mathrm{st}} e\, \exists^{\mathrm{st}} f\, \lambda(e, f, w)))$$
$$\Rightarrow ((\forall^{\mathrm{st}} g\, \exists^{\mathrm{st}} h\, \phi(g, h, w) \Rightarrow \forall^{\mathrm{st}} i\, \exists^{\mathrm{st}} j\, \psi(i, j, w))$$
$$\Rightarrow (\forall^{\mathrm{st}} k\, \exists^{\mathrm{st}} l\, \phi(k, l, w) \Rightarrow \forall^{\mathrm{st}} m\, \exists^{\mathrm{st}} n\, \lambda(m, n, w))), \quad (9)$$

whose reduction is

$$\forall m \bar{l} \bar{g} \bar{j} \bar{a} \bar{c} \bar{f}\, \exists^{\mathrm{fin}} e' \bar{d}' \bar{b}' i' \bar{h}' k' n'\, \forall w\, \exists e \in e'\, \bar{d} \in \bar{d}'\, \bar{b} \in \bar{b}'\, i \in i'\, \bar{h} \in \bar{h}'\, k \in k'\, n \in n'$$
$$((\phi(\bar{a}, \bar{b}, w) \Rightarrow (\psi(\bar{c}, \bar{d}, w) \Rightarrow \lambda(e, \bar{f}, w)) \Rightarrow ((\phi(\bar{g}, \bar{h}, w) \Rightarrow \psi(i, \bar{j}, w))$$
$$\Rightarrow (\phi(k, \bar{l}, w) \Rightarrow \lambda(m, n, w)))), \quad (10)$$

where we have omitted the arguments of the functions. Let $e' = \{m\}$, $e = m$,

$n' = \{\bar{f}\}$, $n = \bar{f}$, $i' = \{\bar{c}\}$, $i = \bar{c}$, $\bar{d}' = \{\bar{j}\}$, $\bar{d} = \bar{j}$. Then $\psi(\bar{c}, \bar{d}, w)$ and $\psi(i, \bar{j}, w)$ are the same formula ψ, and $\lambda(\bar{e}, \bar{f}, w)$ and $\lambda(m, n, w)$ are the same formula λ. However, there is no way to make $\phi(\bar{a}, \bar{b}, w)$, $\phi(\bar{g}, \bar{h}, w)$, and $\phi(k, \bar{l}, w)$ all be the same formula, so we let k' be the two-element set $k' = \{\bar{a}, \bar{g}\}$ and let $\bar{b}' = \{\bar{l}\}$, $\bar{b} = \bar{l}$, $\bar{h}' = \{\bar{l}\}$, $\bar{h} = \bar{l}$. Then the statement becomes the closure of a tautology of the form

$$((\phi_1 \Rightarrow (\psi \Rightarrow \lambda)) \Rightarrow ((\phi_2 \Rightarrow \psi) \Rightarrow (\phi_1 \Rightarrow \lambda))$$
$$\vee ((\phi_1 \Rightarrow (\psi \Rightarrow \lambda)) \Rightarrow ((\phi_2 \Rightarrow \psi) \Rightarrow (\phi_2 \Rightarrow \lambda)). \tag{11}$$

(To see that (11) is a tautology, observe that if ϕ_1 and ϕ_2 have opposite truth values, then one of the final consequents $\phi_1 \Rightarrow \lambda$ or $\phi_2 \Rightarrow \lambda$ is true and so (11) is true. But if ϕ_1 and ϕ_2 are equivalent, then each disjunct in (11) is equivalent to a tautology of the form (ii).) Similarly, the closure of (iii) has the reduction

$$\forall \bar{j}\bar{e}\bar{h}\bar{a}\bar{c}\ \exists^{\text{fin}}\bar{d}'\bar{b}'g'\bar{f}'i'\ \forall w\ \exists \bar{d} \in \bar{d}'\ \bar{b} \in \bar{b}'\ g \in g'\ \bar{f} \in \bar{f}'\ i \in i'$$
$$((\sim\psi(\bar{a}, \bar{b}, w) \Rightarrow \sim\phi(\bar{c}, \bar{d}, w)) \Rightarrow ((\sim\psi(\bar{e}, \bar{f}, w) \Rightarrow \phi(\bar{g}, \bar{h}, w)) \Rightarrow \psi(\bar{i}, \bar{j}, w))).$$

Let $\bar{d}' = \{\bar{h}\}$, $\bar{d} = \bar{h}$, $\bar{b}' = \{\bar{j}\}$, $\bar{b} = \bar{j}$, $g' = \{\bar{c}\}$, $g = \bar{c}$, $\bar{f}' = \{\bar{j}\}$, $\bar{f} = \bar{j}$, and let $i' = \{\bar{a}, \bar{e}\}$. Then the statement becomes the closure of a tautology of the form

$$((\sim\psi_1 \Rightarrow \sim\phi) \Rightarrow ((\sim\psi_2 \Rightarrow \phi) \Rightarrow \psi_1))$$
$$\vee ((\sim\psi_1 \Rightarrow \sim\phi) \Rightarrow ((\sim\psi_2 \Rightarrow \phi) \Rightarrow \psi_2)).$$

Finally, if t is a variable the closure of (iv) may be written as

$$\forall w\ \forall t\ (\forall y\ \forall^{\text{st}}a\ \exists^{\text{st}}b\ \phi(a, b, y, w) \ \Rightarrow\ \forall^{\text{st}}c\ \forall^{\text{st}}d\ \phi(c, d, t, w)),$$

whose reduction is

$$\forall c\ \forall^{\text{fin}}\bar{b}'\ \exists^{\text{fin}}a'\ \exists^{\text{fin}}d'\ \forall w\ \forall t\ \exists a \in a'\ \exists d \in d'$$
$$(\forall y\ \exists b \in \bar{b}'(a)\ \phi(a, b, y, w) \ \Rightarrow\ \phi(c, d, t, w)).$$

(If t is a constant, omit $\forall t$.) Let $a' = \{c\}$, $a = c$, $d' = \{\bar{b}'(c)\}$. Then the statement becomes equivalent to the closure of

$$\forall y\ \exists b \in \bar{b}'(c)\ \phi(c, b, y, w) \ \Rightarrow\ \exists b \in \bar{b}'(c)\ \phi(c, b, t, w)$$

which is of the form (iv) and so is a theorem of $\bar{\mathscr{A}}$. □

Theorem 4. *The reductions of the closures of* (v) *and* (vi) *are theorems of* $\bar{\mathscr{A}}$.

Proof. We give the proof for (vi). The case (v) follows simply by omitting the variable z.

We may write the closure of (vi) as

$$\forall w\ \forall x\ \exists y\ \exists z\ ((\Phi(y, w) \Rightarrow \Psi(y, w)) \Rightarrow (\Phi(z, w) \Rightarrow \Psi(x, w))). \tag{12}$$

The idea, of course, is to let y and z be x. We claim that the reduction of

$$\forall w\, \forall x\, ((\Phi(x, w) \Rightarrow \Psi(x, w)) \Rightarrow (\Phi(x, w) \Rightarrow \Psi(x, w)) \tag{13}$$

is a theorem of $\tilde{\mathscr{A}}$. To see this, notice that (13) is of the form

$$\forall t\, (\Gamma(t) \Rightarrow \Gamma(t)). \tag{14}$$

Let the reduction of $\Gamma(t)$ be $\forall^{\mathrm{st}} u\, \exists^{\mathrm{st}} v\, \gamma(u, v, t)$. Then the reduction of (14) may be written as

$$\forall c\, \forall \tilde{b}\, \exists^{\mathrm{fin}} a'\, \exists^{\mathrm{fin}} d'\, \forall t\, \exists a \in a'\, \exists d \in d'\, (\gamma(a, \tilde{b}(a), t) \Rightarrow \gamma(c, d, t)).$$

If we let $a' = \{c\}$, $a = c$, $d' = \{\tilde{b}(c)\}$, $d = \tilde{b}(c)$ this becomes the theorem

$$\forall c\, \forall \tilde{b}\, \forall t\, (\gamma(c, \tilde{b}(c), t) \Rightarrow \gamma(c, \tilde{b}(c), t)).$$

Therefore, letting $\Lambda(x, y, z, w)$ be the part of (12) after the quantifiers, we need only show in $\tilde{\mathscr{A}}$ that the reduction of $\forall w\, \forall x\, \Lambda(x, x, x, w)$ implies the reduction of $\forall w\, \forall x\, \exists y\, \exists z\, \Lambda(x, y, z, w)$. By taking ordered pairs $\langle w, x\rangle$ and $\langle y, z\rangle$ we may simplify the notation: we need only show in $\tilde{\mathscr{A}}$ that the reduction of $\forall x\, \Sigma(x, x)$ implies the reduction of $\forall x\, \exists y\, \Sigma(x, y)$. Equivalently, we need only show in $\tilde{\mathscr{A}}$ that the reduction of $\exists x\, \forall y\, \Delta(x, y)$ implies the reduction of $\exists x\, \Delta(x, x)$.

Let the partial reduction of $\Delta(x, y)$ be $\forall^{\mathrm{st}} u\, \exists^{\mathrm{st}} v\, \delta(u, v, x, y)$. Then the reduction of $\exists x\, \forall y\, \Delta(x, y)$ is

$$\exists \tilde{v}'\, \forall^{\mathrm{fin}} u'\, \exists x\, \forall u \in u'\, \forall y\, \exists v \in \tilde{v}'(u)\, \delta(u, v, x, y), \tag{15}$$

where $\tilde{v}'$ denotes a finite-set-valued function. Writing $\tau(u, v, x)$ for $\delta(u, v, x, x)$, we see that (15) implies

$$\exists \tilde{v}'\, \forall^{\mathrm{fin}} u'\, \exists x\, \forall u \in u'\, \exists v \in \tilde{v}'(u)\, \tau(u, v, x). \tag{16}$$

The reduction of $\exists x\, \Delta(x, x)$ is

$$\exists \tilde{v}\, \forall^{\mathrm{fin}} u'\, \exists x\, \forall u \in u'\, \tau(u, \tilde{v}(u), x). \tag{17}$$

We need to show that (16) implies (17). Suppose that (16) holds, and let $\tilde{v}'$ be such that

$$\forall^{\mathrm{fin}} u'\, \exists x\, \forall u \in u'\, \exists v \in \tilde{v}'(u)\, \tau(u, v, x). \tag{18}$$

Recall that u is restricted to range over a certain set X in $\hat{M}$. Let Ω be the Cartesian product over u of the finite sets $\tilde{v}'(u)$. Thus Ω is the set of all cross-sections of $\tilde{v}'$. Give Ω the product topology, where each $\tilde{v}'(u)$ has the discrete topology. By Tychonov's theorem, Ω is a compact space. By (18), for each finite u' there is a $\tilde{v}_{u'}$ in Ω such that $\exists x\, \forall u \in u'\, \tau(u, \tilde{v}_{u'}(u), x)$. The u' are a directed set under inclusion, so that $u' \mapsto \tilde{v}_{u'}$ is a net of points in Ω. Let $\tilde{v}$ be a limit point of this net. By definition of the product topology, for all finite u' there is a finite w' containing u' such that $\tilde{v}$ agrees with $\tilde{v}_{w'}$ on u', and so $\exists x\, \forall u \in u'\, \tau(u, \tilde{v}(u), x)$. This proves (17). □

This completes the syntactical interpretation of $\mathscr{A}^*$ within $\tilde{\mathscr{A}}$.

4. The unrestricted standardization principle

In $\mathscr{A}^*$ we adopted (S_0) rather than (S') as an axiom scheme, but the unrestricted standardization principle (S') is quite useful in nonstandard analysis. We claim that (S') is a theorem of $\mathscr{A}^*$. Let the reduction of $\Phi(x, y)$ be $\forall^{st}u\, \exists^{st}v\, \phi(u, v, x, y)$; then we need to prove

$$\forall^{st}x\, \exists^{st}y\, \forall^{st}u\, \exists^{st}v\, \phi(u, v, x, y) \Rightarrow \exists^{st}\bar{y}\, \forall^{st}x\, \forall^{st}u\, \exists^{st}v\, \phi(u, v, x, \bar{y}(x)). \quad (19)$$

We claim that it suffices to prove

$$\forall^{st}x\, \exists^{st}y\, \forall^{st}u\, \phi(u, x, y) \Rightarrow \exists^{st}\bar{y}\, \forall^{st}x\, \forall^{st}u\, \phi(u, x, \bar{y}(x)). \quad (19a)$$

This is because, using (S_0), we can rewrite (19) as

$$\forall^{st}x\, \exists^{st}y\, \exists^{st}\bar{v}\, \forall^{st}u\, \phi(u, \bar{v}(u), x, y) \Rightarrow \exists^{st}\bar{y}\, \forall^{st}x\, \exists^{st}\bar{v}\, \forall^{st}u\, \phi(u, \bar{v}(u), x, \bar{y}(x)). \quad (19b)$$

Applying (19a) with y replaced by the pair $(y, \bar{v})$, we have

$$\forall^{st}x\, \exists^{st}y\, \exists^{st}\bar{v}\, \forall^{st}u\, \phi(u, \bar{v}(u), x, y) \Rightarrow$$
$$\exists^{st}\bar{y}\, \exists^{st}\bar{\bar{v}}\, \forall^{st}x\, \forall^{st}u\, \phi(u, \bar{\bar{v}}(x)(u), x, \bar{y}(x)), \quad (19c)$$

but the consequent of (19c) implies the consequent of (19b) (let $\bar{v} = \bar{\bar{v}}(x)$).

Using (S_0) twice, we can rewrite (19a) as

$$\forall^{st}x_0\, \forall^{st}\bar{u}_0\, \exists^{st}y_0\, \phi(\bar{u}_0(y_0), x_0, y_0) \Rightarrow \forall^{st}\bar{x}\, \forall^{st}\bar{u}\, \exists^{st}\bar{y}\, \phi(\bar{u}(\bar{y}), \bar{x}(\bar{y}), \bar{y}(\bar{x}(\bar{y}))), \quad (19d)$$

and using (S_0) again we can rewrite the antecedent of (19d) as

$$\exists^{st}\bar{y}_0\, \forall^{st}x_0\, \forall^{st}\bar{u}_0\, \phi(\bar{u}_0(\bar{y}_0(x_0, \bar{u}_0)), x_0, \bar{y}_0(x_0, \bar{u}_0)).$$

Thus (19a) is equivalent to

$$\forall^{st}\bar{y}_0\, \forall^{st}\bar{x}\, \forall^{st}\bar{u}\, \exists^{st}\bar{y}\, \exists^{st}x_0\, \exists^{st}\bar{u}_0$$
$$(\phi(\bar{u}_0(\bar{y}_0(x_0, \bar{u}_0)), x_0, \bar{y}_0(x_0, \bar{u}_0)) \Rightarrow \phi(\bar{u}(\bar{y}), \bar{x}(\bar{y}), \bar{y}(\bar{x}(\bar{y})))). \quad (19e)$$

But (19e) is a consequence of

$$\forall^{st}\bar{y}_0\, \forall^{st}\bar{x}\, \forall^{st}\bar{u}\, \exists^{st}\bar{y}\, \exists^{st}x_0\, \exists^{st}\bar{u}_0$$
$$(\bar{u}_0(\bar{y}_0(x_0, \bar{u}_0)) = \bar{u}(\bar{y}) \wedge x_0 = \bar{x}(\bar{y}) \wedge \bar{y}_0(x_0, \bar{u}_0) = \bar{y}(\bar{x}(\bar{y}))). \quad (19f)$$

By (T) we may delete the superscripts 'st' from the quantifiers in (19f). (We could not do this earlier because ϕ might contain undisplayed free variables.) Now we can use the axiom of choice freely to interchange quantifiers, and we find that (19f) is equivalent to

$$\forall\bar{y}_0\, \exists\bar{y}\, \forall x\, \forall u\, \exists x_0\, \exists\bar{u}_0\, (\bar{u}_0(\bar{y}_0(x_0, \bar{u}_0)) = u \wedge x_0 = x \wedge \bar{y}_0(x_0, \bar{u}_0) = \bar{y}(x)), \quad (19g)$$

$$\forall\bar{y}_0\, \forall x\, \exists y\, \forall u\, \exists x_0\, \exists\bar{u}_0\, (\bar{u}_0(\bar{y}_0(x_0, \bar{u}_0)) = u \wedge x_0 = x \wedge \bar{y}_0(x_0, \bar{u}_0) = y), \quad (19h)$$

$$\forall\bar{y}_0\, \forall x\, \forall\bar{u}\, \exists y\, \exists x_0\, \exists\bar{u}_0\, (\bar{u}_0(\bar{y}_0(x_0, \bar{u}_0)) = \bar{u}(y) \wedge x_0 = x \wedge \bar{y}_0(x_0, \bar{u}_0) = y). \quad (19i)$$

But (19i) is trivial: let $y = \bar{y}_0(x, \bar{u})$, let $x_0 = x$, and let $\bar{u}_0 = \bar{u}$. This concludes the proof of the unrestricted standardization principle.

I am grateful to Professor L. Haddad for pointing out an erroneous argument in the original version of this section.

5. The saturation principle

In [3, Section 2] we demonstrated two useful general principles of nonstandard analysis which follow from the reduction algorithm. Here is a third, which we call the *saturation principle*. It is a theorem of $\mathscr{A}^*$ in which $\Phi(x, y)$ is any formula of $\mathscr{A}^*$:

$$\forall^{\mathrm{st}} x\, \exists y\, \Phi(x, y) \Rightarrow \exists \bar{y}\, \forall^{\mathrm{st}} x\, \Phi(x, \bar{y}(x)). \tag{20}$$

Proof. Let the partial reduction of $\Phi(x, y)$ be $\forall^{\mathrm{st}} u\, \exists^{\mathrm{st}} v\, \phi(u, v, x, y)$. Then we need to show that

$$\forall^{\mathrm{st}} x\, \exists y\, \forall^{\mathrm{st}} u\, \exists^{\mathrm{st}} v\, \phi(u, v, x, y) \tag{21}$$

implies

$$\exists \bar{y}\, \forall^{\mathrm{st}} x\, \forall^{\mathrm{st}} u\, \exists^{\mathrm{st}} v\, \phi(u, v, x, \bar{y}(x)). \tag{22}$$

Applying (S_0), (I), and (S_0) we see that (21) is equivalent to

$$\exists^{\mathrm{st}} \bar{v}\, \forall^{\mathrm{st}} x\, \forall^{\mathrm{st\,fin}} u'\, \exists y\, \forall u \in u'\, \phi(u, \bar{v}(x, u), x, y), \tag{23}$$

and applying (S_0) and (I) we see that (22) is equivalent to

$$\exists^{\mathrm{st}} \bar{v}\, \forall^{\mathrm{st\,fin}} x'\, \forall^{\mathrm{st\,fin}} u'\, \exists \bar{y}\, \forall x \in x'\, \forall u \in u'\, \phi(u, \bar{v}(x, u), x, \bar{y}(x)). \tag{24}$$

Let $\bar{v}$ be a standard function such that (23) holds for it, and let u' be any standard finite set. Then (23) becomes

$$\forall^{\mathrm{st}} x\, \exists y\, \forall u \in u'\, \phi(u, \bar{v}(x, u), x, y) \tag{25}$$

and we need to show that

$$\forall^{\mathrm{st\,fin}} x'\, \exists \bar{y}\, \forall x \in x'\, \forall u \in u'\, \phi(u, \bar{v}(x, u), x, \bar{y}(x)). \tag{26}$$

But since every element of a standard finite set is standard [3, Theorem 1.1], the implication (25) $\Rightarrow$ (26) is clear. □

6. Application to IST

Suppose that the first-order theory $\mathscr{A}$ we start with is ZFC itself. Then $\tilde{\mathscr{A}}$ will not be ZFC, but the theory of a model of ZFC. In doing ordinary mathematics it is important to stay with set theory itself, rather than introduce a model of set theory which may have unpleasant features. The simplest way around this difficulty is perhaps the following, which uses an idea of Lévy as suggested to me

by Powell. Let ZFC[V] be ZFC with a constant V and the following additional axioms:

$$V = R(\alpha) \quad \text{for some ordinal } \alpha, \tag{27}$$

$$\phi \Leftrightarrow \phi^V, \tag{28}$$

where ϕ is any formula of ZFC and ϕ^V is its relativization to V (see the Appendix to [3] for a definition of these notions). Then ZFC[V] is a conservative extension of ZFC; this is the reflection principle (see e.g., Theorem 8.7 of [3]). The reduction algorithm interprets any statement of IST as a statement of ZFC[V], and the argument given in this paper shows that any proof in IST may be regarded as an abbreviation of a proof in ZFC[V].

The theory ZFC[V] is essentially ZFC$^{\sim}$ with the additional axiom (27). The advantage of (27) is that the increase in logical type arising from the reduction algorithm is, in the case of limited formulas, only apparent: if $\bar{y}: V \to V$ and $x \in V$, $z \in V$ and the range of $\bar{y}$ on x is contained in z, then the restriction $\bar{y}_0$ of $\bar{y}$ to x is an ordinary function from x to z (i.e., is an element of V).

For example, the saturation principle (20) gives the following theorem of IST, in which $\Phi(x, y)$ is any formula of IST:

Theorem 5 (Saturation principle). *Let X and Y be standard sets, and suppose that for all standard x in X there is a y in Y such that we have $\Phi(x, y)$. Then there is a function $\bar{y}: X \to Y$ such that for all standard x in X we have $\Phi(x, \bar{y}(x))$.*

7. An example

In the literature it is customary to make sharp distinctions between standard theorems and nonstandard theorems, and between standard proofs and nonstandard proofs. Our results show that this is not a very basic distinction. Every nonstandard theorem of IST can be rewritten, using the reduction algorithm, as an equivalent standard theorem in ZFC[V], and every nonstandard proof in IST can be rewritten, using the algorithm developed here, as a standard proof in ZFC[V].

The advantage of stating certain theorems in nonstandard form is one of simplicity: the nonstandard version may have fewer quantifiers and be of simpler logical type. It is worth remarking that this is only possible when one adopts the full idealization principle (I) with free variables taking arbitrary internal values (corresponding semantically to an adequate ultralimit) rather than the restricted idealization principle of Robinson's notion of an enlargement in which all parameters must be standard (corresponding semantically to an adequate ultrapower).

An additional advantage of nonstandard proofs is that inferences which are trivial in nonstandard analysis may be difficult when reduced to standard form. As

an example, consider the nonstandard proof of the Tychonov theorem (see e.g., [3, Theorem 6.2]).

Let T be a set, $t \mapsto X_t$ be a function from T to compact topological spaces, let $\mathcal{U}_t$ be the set of all open subsets of X_t, and let Ω be the Cartesian product $\Omega = \prod_{t \in T} X_t$ with the product topology. We need to show that Ω is compact. By transfer, we may assume that T, $t \mapsto X_t$, $t \mapsto \mathcal{U}_t$, and Ω are standard, so we treat them as constants.

Let $\Phi(\omega)$ be the formula

$$\forall^{st} t \in T\, \exists^{st} \eta \in X_t\, \forall^{st} U \in \mathcal{U}_t\, (\omega(t) \in U \Rightarrow \eta \in U)$$

and let $\Psi(\omega)$ be the formula

$$\exists^{st} \bar{\eta} \in \Omega\, \forall^{st} t\, \forall^{st} U \in \mathcal{U}_t\, (\omega(t) \in U \Rightarrow \bar{\eta}(t) \in U).$$

Then the nonstandard proof goes as follows. Let $\omega \in \Omega$. Then $\Phi(\omega)$ holds because a point $\omega(t)$ in a standard compact space is near standard. Therefore $\Psi(\omega)$ holds by the standardization principle, but by definition of the product topology $\Psi(\omega)$ says that ω is near-standard in Ω, so that Ω is compact.

The reduction of $\forall \omega\, \Phi(\omega)$ is simply the hypothesis that each factor is compact, the reduction of $\forall \omega\, (\Phi(\omega) \Rightarrow \Psi(\omega))$ is somewhat complicated but trivial (being an instance of standardization), and the reduction of $\forall \omega\, \Psi(\omega)$ is the conclusion of the theorem! This illustrates that the reduction of (vi), to take care of modus ponens with a parameter, is the key to reducing a nonstandard proof to a standard one. (The example is of course somewhat circular because the Tychonov theorem for products of finite spaces was used in the proof of Theorem 4.)

What is really new in nonstandard analysis is not theorems or proofs, but concepts—external predicates, such as "x is infinitesimal" or "ω is near-standard". These correspond to nothing in conventional mathematics. Some of them express old mathematical intutitions which were long regarded as illegitimate; others express intriguing new concepts for which intuition can be developed.

For a striking example of a new and important result discovered by nonstandard analysis, and making highly nontrivial use of nonstandard concepts, the reader is referred to the paper [1] by Lawler on self-avoiding random walks.

References

[1] G.F. Lawler, A self-avoiding random walk, Duke Math. J. 47 (1980) 655–696.
[2] E. Mendelson, Introduction to Mathematical Logic (Van Nostrand, New York, 1964).
[3] E. Nelson, Internal set theory: a new approach to nonstandard analysis, Bull. Amer. Math. Soc. 83 (1977) 1165–1198.
[4] A. Robinson, Non-Standard Analysis, 2nd ed. (American Elsevier, New York, 1974; 1st ed., 1966).
[5] A. Robinson, Metamathematical problems, J. Symbolic Logic 38 (1973) 500–515.