THE BULLETIN OF SYMBOLIC LOGIC
Volume 4, Number 3, Sept. 1998

DEVELOPMENTS IN CONSTRUCTIVE NONSTANDARD ANALYSIS

ERIK PALMGREN

Abstract. We develop a constructive version of nonstandard analysis, extending Bishop's constructive analysis with infinitesimal methods. A full transfer principle and a strong idealisation principle are obtained by using a sheaf-theoretic construction due to I. Moerdijk. The construction is, in a precise sense, a reduced power with variable filter structure. We avoid the nonconstructive standard part map by the use of nonstandard hulls. This leads to an infinitesimal analysis which includes nonconstructive theorems such as the Heine–Borel theorem, the Cauchy–Peano existence theorem for ordinary differential equations and the exact intermediate-value theorem, while it at the same time provides constructive results for concrete statements. A nonstandard measure theory which is considerably simpler than that of Bishop and Cheng is developed within this context.

§1. Introduction.

> 'It has often been held that nonstandard analysis is highly nonconstructive, thus somewhat suspect, depending as it does upon the ultrapower construction to produce a model [...] On the other hand, nonstandard *praxis* is remarkably constructive; having the extended number set we can proceed with explicit calculations.' [S. Albeverio, J.-E. Fenstad, R. Høegh-Krohn and T. Lindstrøm. *Nonstandard Methods in Stochastic Analysis and Mathematical Physics.* Academic Press 1986, page 31.]

The very first model of nonstandard analysis, due to Schmieden and Laugwitz [27], was in fact completely constructive. It used, in effect, a reduced power construction modulo the *Fréchet* filter, i.e., the filter of cofinite sets of natural numbers. This led to a rather restricted transfer principle, and the set of hyperreals became merely a partially ordered ring. (For background on the early history of nonstandard analysis, see the foreword to Robinson's book [26] by W. A. J. Luxemburg.) Classical logic was used

Received August 15, 1997; revised June 2, 1998.
The research reported herein was supported by grants from the Swedish Research Councils for Engineering Sciences (TFR) and Natural Sciences (NFR).

1079-8986/98/0403-0001/$5.00

in the Schmieden–Laugwitz approach for deriving certain results. Subsequent developments of this approach within strict constructivism are found in Martin-Löf [16] and in papers by the author [22, 23].

Nelson [21] showed that substantial parts of nonstandard probability theory can be developed using only an 'elementary' fragment of his internal set theory. This elementary part includes the transfer and the idealisation principles, but avoids the general standardisation principle (and taking standard parts). Further work in this restricted framework, has been done by the Reeb school in nonstandard analysis; see the volume edited by Diener and Diener [9], and in particular the papers by Delfini and Lobry [8] and Cartier and Perrin [5]. Though these developments use classical logic, it is often not used in an essential way. Indeed, since we can make intuitionistic sense of both full transfer and idealisation, we believe that such 'elementary' nonstandard analysis lends itself to constructivisation. A constructively problematic device in classical nonstandard analysis is still the *standard part map*, which to any finite hyperreal $x \in {}^*\mathbb{R}$ assigns a standard real $x^\circ \in \mathbb{R}$ infinitely close to x. For a discussion of some non-constructive consequences of such a principle, see Wattenberg [29].

It is not at all obvious that the transfer principle can be constructivised, and if we require that the model has the standard interpretation of the logical symbols, it is outright impossible. Suppose that there exists a nonstandard model ${}^*\mathbb{N}$ of arithmetic and that the transfer principle

$$\mathbb{N} \models \Phi \iff {}^*\mathbb{N} \models \Phi, \tag{1}$$

holds for all arithmetical formulas Φ. Then we can constructively decide whether Φ or $\neg\Phi$ holds for Π^0_1-formulas Φ. Indeed, let $\Phi \equiv \forall n\, \Psi(n)$, where Ψ is a bounded formula. With such a formula $\Psi(n)$ we can express, for example, "$2n+4$ is a sum of two primes". Clearly $\forall m\, (\forall n < m\, \Psi(n) \vee \exists n < m\, \neg\Psi(n))$ holds in $\mathbb{N}$, so by transfer it holds in ${}^*\mathbb{N}$. Now taking $m \in {}^*\mathbb{N}$ to be infinite, we obtain ${}^*\mathbb{N} \models \Psi({}^*n)$, for all $n \in \mathbb{N}$, or ${}^*\mathbb{N} \models \exists n < m\, \neg\Psi(n)$. It follows by transfer in the other direction that $\mathbb{N} \models \Phi \vee \neg\Phi$. This means, in particular, that we can decide Goldbach's conjecture. So we have obtained a Brouwerian counterexample to such a transfer principle. In two earlier papers [22, 23] we investigated a weaker form of nonstandard model which built on the idea of Schmieden and Laugwitz and Martin-Löf's constructive interpretation of this model [16]. However we obtained only a restricted, but quite useful, kind of transfer for formulas of the form $\forall \bar{x}\, [\Phi(\bar{x}) \to \exists \bar{y}\, \Psi(\bar{x}, \bar{y})]$ where Φ and Ψ are conjunctions of atomic formulas. To extend the transfer principle further seemed difficult. The advantage of this model is that we have a direct and fixed interpretation of nonstandard objects.

A surprising solution to the transfer problem was given by Moerdijk [17]. He changed the notion of model on the nonstandard side and was able to regain, not only the full transfer principle, but also the full overspill principle (which, constructively, is not a consequence of transfer). The model consists of sheaves over a category of filters, and the semantics required is a general form of forcing. The model consists in a sense of reduced powers with varying or floating filter structure. It can be seen as a constructive counterpart to an ultrapower. The idea of using a nonstandard notion of model to obtain full transfer was considered already by Martin-Löf [16], but his notion of nonstandard truth is not compositionally defined. Martin-Löf's paper also contains the idea of using different filters, or rather inverse systems, but they are not combined into one model. In [24] we extended Moerdijk's model to all constructive sets and showed how to develop some basic nonstandard analysis. We also established an idealisation principle strong enough to show that the standard elements of a standard set have a hyperfinite enumeration.

The present paper is an exposition and a continuation of the work in [17, 24]. We emphasise that the development is done in compliance with Bishop's strict constructivism [1], and that it may indeed be formalised within Martin-Löf's constructive type theory [15], which will be the official metatheory in case of doubt. Thus it is in principle possible to extract algorithms from all the existence results we establish. A first application of the nonstandard model is to give intuitive nonstandard proofs of standard theorems of constructive analysis. We thus supplement our earlier paper [24] with further examples. A second more radical application is to develop an infinitesimal calculus inside the model based on hyperrational numbers. In this calculus we regain, somewhat mysteriously, the classical nonconstructive theorems: the Heine–Borel theorem, the Cauchy–Peano existence theorem and an exact version of the intermediate value theorem. The explanation is that the notion of existence is weaker than in constructive analysis, but nevertheless stronger than in classical analysis, since constructive content may be obtained using the transfer principle. For example, if we find a sequence of arbitrarily close approximations to zeros of a real function, in a compact interval, then we can use the overspill principle to find an 'ideal' zero up to an infinitesimal. This is possible even in the case the approximations are not converging to the same point, and there may be no effective way of extracting a convergent subsequence. In classical nonstandard analysis we could extract a real zero by applying the nonconstructive standard part map. This map is not available in the sheaf model [24]. Instead we work directly with the ideal objects, and may on this basis develop a calculus with constructive content which is intermediate between constructive and classical analysis. The sheaf model should also be suited for developing a constructive version of Nelson's radically elementary probability theory [21]. We present here an integration theory which is considerably simpler than the Bishop–Cheng theory [2], though it provides less information about the measurable functions.

An outline of the paper. Section 2 is mainly an exposition of the sheaf-theoretic model [24] and its basic properties. It is not necessary to be familiar with sheaf theory to understand this construction. In Section 3 we show that the model has some apparently nonconstructive features. For instance, König's lemma and a weak principle of limited omniscience are true in the model. Some nonstandard characterisations of standard notions are established in Section 4. In Section 5 we prove theorems of elementary calculus using the new nonstandard model, thus supplementing the examples given in [24]. In Section 6 we develop infinitesimal calculus from hyperrational numbers inside the model, in a spirit which is similar to Nelson's radically elementary probability theory. In the Appendix some classical aspects of the model are discussed.

§**2. The sheaf-theoretic model of nonstandard analysis.** The most intuitive classical construction of a basic model for nonstandard analysis is perhaps the ultrapower

$$\mathbb{R}^{\mathbb{N}}/\mathcal{U}.$$

Here $\mathbb{R}$ is the structure of real numbers with a sufficiently rich supply of predicates and functions, and $\mathcal{U}$ is a nonprincipal ultrafilter on $\mathbb{N}$. Since all interesting ultrafilters are noneffective this construction will not work for our purposes. The ultrapower is an instance of the general reduced-power construction $\mathcal{A}^I/\mathcal{F}$. But as soon as we reduce by a filter $\mathcal{F}$ which is not an ultrafilter, e.g., the Fréchet filter, the transfer principle becomes much more limited (but still useful, see [22]). We shall see that by allowing $\mathcal{F}$ to vary, when interpreting a statement, the full transfer principle may be regained in a completely constructive setting. This does not contradict the discussion in the introduction, since we will use a nonstandard interpretation of the logical symbols. Our method uses sheaf semantics over a category of filters.

A *filter base* $\mathcal{F} = (I, U, F)$ consists of a nonvoid index set I and a family $F = \{F_i\}_{i \in I}$ of subsets of an underlying set U, called *base sets*, satisfying the filtering condition: for all $i, j \in I$, there exists $k \in I$ such that $F_k \subseteq F_i \cap F_j$. The filter generated is then $\{ S \subseteq U : (\exists i \in I)\ F_i \subseteq S \}$. For constructive reasons, it will be better to work only with the bases of filters. So in the sequel we shall abuse the language and simply call them *filters*. Let $\mathcal{F} = (I, U, F)$ and $\mathcal{G} = (J, V, G)$ be filters. A *continuous map* α from $\mathcal{F}$ to $\mathcal{G}$, in symbols $\alpha\colon \mathcal{F} \to \mathcal{G}$, is a partial function $\alpha\colon U \to V$ which is totally defined on some base set of $\mathcal{F}$, and satisfies the *continuity* condition,

$$(\forall j \in J)\ (\exists i \in I)\ \alpha[F_i] \subseteq G_j.$$

Two such morphisms are equivalent if they agree and are both totally defined on some base set of $\mathcal{F}$. The composition $\beta \circ \alpha$ of two continuous maps $\alpha\colon \mathcal{F} \to \mathcal{G}$ and $\beta\colon \mathcal{G} \to \mathcal{H}$ is again a continuous map $\mathcal{F} \to \mathcal{H}$. The filters together with the continuous maps then form a category $\mathbb{B}$ with pullbacks and a terminal object 1, see [17] and [24]. For a detailed study of such categories, see Blass [3].

The reduced-power construction is slightly modified, since constructively, sets with equivalence relations are preferred over equivalence classes. Let $\mathcal{A} = (A; R, \dots)$ be a first-order structure. The reduced power of $\mathcal{A}$ with respect to $\mathcal{F}$, denoted ${}^*\mathcal{A}(\mathcal{F}) = \mathcal{A}^U/F$, consists of partial functions $\alpha\colon U \to A$ which are totally defined on some base set of $\mathcal{F}$. Two such objects are equivalent if they are both totally defined, and agree, on some base set of $\mathcal{F}$. Moreover, define

$$ {}^*\mathcal{A}(\mathcal{F}) \models R(\alpha_1, \dots, \alpha_n) $$

if, and only if,

$$ (\exists i \in I)\ (\forall u \in F_i)\ R(\alpha_1(u), \dots, \alpha_n(u)), $$

where $R(\alpha_1(u), \dots, \alpha_n(u))$ is considered false if not all $\alpha_i(u)$s are defined. If we consider the set A above as equipped with the trivial filter structure (the only base set is A), the objects of the reduced power ${}^*\mathcal{A}(\mathcal{F})$ are just continuous maps $\mathcal{F} \to A$. To motivate the model construction below we consider how the truth values of a basic relation R in two reduced powers of the same structure are related when their filters are related via a continuous map $\beta\colon \mathcal{G} \to \mathcal{F}$. Suppose ${}^*\mathcal{A}(\mathcal{F}) \models R(\alpha)$, i.e., for some i, $(\forall u \in F_i)\ R(\alpha(u))$. Then since β is continuous there exists j with $\beta[G_j] \subseteq F_i$. Hence $(\forall v \in G_j)\ R(\alpha(\beta(v)))$, that is ${}^*\mathcal{A}(\mathcal{G}) \models R(\alpha \circ \beta)$. Thus we have the monotonicity property for truth

$$ {}^*\mathcal{A}(\mathcal{F}) \models R(\alpha) \Longrightarrow {}^*\mathcal{A}(\mathcal{G}) \models R(\alpha \circ \beta). \tag{2} $$

Let $\mathcal{G}^{(k)} = (J^{(k)}, V^{(k)}, G^{(k)})$, $k = 1, \dots, n$, and $\mathcal{F} = (I, U, F)$ be filters. A finite set of continuous maps $\{\beta_k : \mathcal{G}^{(k)} \to \mathcal{F}\}_{k=1}^n$ is called a *cover* of $\mathcal{F}$ if for all $j_1 \in J^{(1)}, \dots, j_n \in J^{(n)}$ there exists $i \in I$ for which

$$ \beta_1[G^{(1)}_{j_1}] \cup \cdots \cup \beta_n[G^{(n)}_{j_n}] \supseteq F_i. \tag{3} $$

In case a single map makes up the cover, it is called a *covering map*. (Nontrivial covering maps are important in several of the results below.) For reduced powers (3) means that we have a kind of covering property for truth:

$$ {}^*\mathcal{A}(\mathcal{G}^{(1)}) \models R(\alpha \circ \beta_1), \dots, {}^*\mathcal{A}(\mathcal{G}^{(n)}) \models R(\alpha \circ \beta_n) \Longrightarrow {}^*\mathcal{A}(\mathcal{F}) \models R(\alpha). \tag{4} $$

The *nonstandard interpretation* of a set $S = (S, =_S)$ *at a filter* $\mathcal{F}$ is the reduced power ${}^*S(\mathcal{F})$ where S is considered as a set with equality. Then for each set S, ${}^*S(\mathcal{F}) = \mathrm{Hom}_{\mathbb{B}}(\mathcal{F}, S)$ defines a contravariant functor on the category $\mathbb{B}$, the action on morphisms being ${}^*S(\beta)(\alpha) = \alpha \circ \beta$. This functor is in fact a *sheaf* with respect to the Grothendieck topology on $\mathbb{B}$ induced by the cover relation above (see [24]). Hence the name *sheaf semantics*.

Since we are forcing with a category, and thus quantifying over it, we shall have to make the category into a set. Classically, this is done by introducing a (Grothendieck) universe of sets into set theory, i.e., V_κ, where κ is an inaccessible cardinal. The sets belonging to the universe are called *small*. The type universes of Martin-Löf [15] form an exact constructive analogue of set-theoretic universes. An additional complication in the constructive setting is the treatment of subsets and relations. Classically, each subset R of S may be described by a characteristic function $\chi\colon S \to \{0,1\}$. Intuitionistically, such characteristic functions describe only decidable subsets. Instead one replaces $\{0,1\}$ by a Heyting algebra Ω, and if this is rich enough, e.g., a *subobject classifier* [14], the function space $S \to \Omega$ becomes an intuitionistic version of the power set. By the propositions-as-types principle a universe U of types [15] can be considered as a set of propositions or truth-values. In fact letting $\Omega = U$, with logical consequence as the preorder, we get a weak (predicative) version of the subobject classifier, that classifies functions in U. Thus a subset R of S is called U-*small*, or just *small*, if it can be described by a characteristic function $S \to U$.

We assume that $\mathbb{B}$ consists of filters and morphisms where all components are small. The first-order language L we work with has a sort ${}^*S$ and equality symbol ${}^*{=}_S$ for each nonvoid small set S with small equality $=_S$; a relation symbol ${}^*R$ for each small relation R; a function symbol ${}^*f$ for each function f and a constant symbol ${}^*c$ for each constant c. Note that we overload the notation: ${}^*S$ denotes a sort, while ${}^*S(\mathcal{F})$ denotes an interpretation of this sort at the filter $\mathcal{F}$. If $R \subseteq S_1 \times \cdots \times S_n$, then the interpretation of ${}^*R$ at $\mathcal{F}$ is

$$
\begin{aligned}
{}^*R(\mathcal{F}) = \{\, (\alpha_1, \dots, \alpha_n) \in {}^*S_1(\mathcal{F}) \times \cdots \times {}^*S_n(\mathcal{F}) :\\
(\exists i \in I)\ (\forall u \in F_i)\ R(\alpha_1(u), \dots, \alpha_n(u)) \,\}.
\end{aligned}
$$

If $f : S_1 \times \cdots \times S_n \to S$ then the interpretation at $\mathcal{F}$ is an internal function from ${}^*S_1(\mathcal{F}) \times \cdots \times {}^*S_n(\mathcal{F})$ to ${}^*S(\mathcal{F})$,

$$
{}^*f(\mathcal{F})(\alpha_1, \dots, \alpha_n) = \lambda u. f(\alpha_1(u), \dots, \alpha_n(u)).
$$

The formulas are built up from predicates ${}^*R$ and falsity $\perp$ by use of logical symbols, $\wedge$, $\vee$, $\to$ and quantifiers over sorts ${}^*S$, $(\forall x \in {}^*S)$ and $(\exists x \in {}^*S)$. We often omit the sort when it can be inferred from the context. Negation is defined by $\neg\Phi \equiv \Phi \to \perp$. Formulas in the language L are called *internal*. Below we shall introduce new predicates St^S for standardness. The formulas of this extended language L' are called *external*. An internal formula Φ is true in the standard world if it is true as a formula stripped of its stars. We define the interpretation of internal formulas in the nonstandard model by forcing semantics. Then we extend to external formulas.

Definition 1. If $\alpha_1 \in {}^*S_1(\mathcal{F}), \dots, \alpha_n \in {}^*S_n(\mathcal{F})$, write $\bar\alpha = \alpha_1, \dots, \alpha_n$ and $\bar\alpha(u) = \alpha_1(u), \dots, \alpha_n(u)$, and $\bar\alpha \circ \beta = \alpha_1 \circ \beta, \dots, \alpha_n \circ \beta$, for $\beta\colon \mathcal{G} \to \mathcal{F}$. Define inductively,

(i) $\mathcal{F} \Vdash \perp$ if and only if $\mathcal{F}$ has an empty base set,
(ii) $\mathcal{F} \Vdash {}^*R(\bar\alpha)$ if and only if $(\exists i \in I)\ (\forall u \in F_i)\ R(\bar\alpha(u))$,

(iii) $\mathcal{F} \Vdash \Phi(\bar{\alpha}) \wedge \Psi(\bar{\alpha})$ if and only if $\mathcal{F} \Vdash \Phi(\bar{\alpha})$ and $\mathcal{F} \Vdash \Psi(\bar{\alpha})$,

(iv) $\mathcal{F} \Vdash \Phi(\bar{\alpha}) \vee \Psi(\bar{\alpha})$ if and only if for some cover $\{\beta_i : \mathcal{G}_i \to \mathcal{F}\}_{i=1}^n$ at least one of $\mathcal{G}_i \Vdash \Phi(\bar{\alpha} \circ \beta_i)$ or $\mathcal{G}_i \Vdash \Psi(\bar{\alpha} \circ \beta_i)$ holds for each $i = 1, \ldots, n$,

(v) $\mathcal{F} \Vdash \Phi(\bar{\alpha}) \to \Psi(\bar{\alpha})$ if and only if for all $\beta : \mathcal{G} \to \mathcal{F}$: $\mathcal{G} \Vdash \Phi(\bar{\alpha} \circ \beta)$ implies $\mathcal{G} \Vdash \Psi(\bar{\alpha} \circ \beta)$,

(vi) $\mathcal{F} \Vdash (\exists x \in {}^*S)\ \Phi(\bar{\alpha}, x)$ if and only if there is a cover $\{\beta_i : \mathcal{G}_i \to \mathcal{F}\}_{i=1}^n$ and a set of witnesses $\{\delta_i \in {}^*S(\mathcal{G}_i)\}_{i=1}^n$ such that $\mathcal{G}_i \Vdash \Phi(\bar{\alpha} \circ \beta_i, \delta_i)$ for all $i = 1, \ldots, n$,

(vii) $\mathcal{F} \Vdash (\forall x \in {}^*S)\ \Phi(\bar{\alpha}, x)$ if and only if for all $\beta : \mathcal{G} \to \mathcal{F}$ and each $\delta \in {}^*S(\mathcal{G})$, $\mathcal{G} \Vdash \Phi(\bar{\alpha} \circ \beta, \delta)$.

We read $\mathcal{F} \Vdash \Phi(\bar{\alpha})$ as $\mathcal{F}$ *forces* $\Phi(\bar{\alpha})$, or perhaps more suggestively, $\Phi(\bar{x})$ *is true of* $\bar{\alpha}$ *at* $\mathcal{F}$. Note that (i)–(iii) correspond exactly to truth in a reduced power. The following fundamental result is due to Moerdijk [17], and was adapted to a strictly constructive setting in [24].

Theorem 1. *Let* $\Phi(x_1^{*S_1}, \ldots, x_n^{*S_n})$ *be an internal formula, and suppose that* $\alpha_1 \in {}^*S_1(\mathcal{F}), \ldots, \alpha_n \in {}^*S_n(\mathcal{F})$. *Then for* $\bar{\alpha} = \alpha_1, \ldots, \alpha_n$,

$$\mathcal{F} \Vdash \Phi(\bar{\alpha}) \quad \textit{if and only if} \quad (\exists i \in I)\ (\forall u \in F_i)\ \Phi(\bar{\alpha}(u)).$$

Proof. By induction on formulas. The cases (i)–(iii) are trivial.

(iv, $\Longrightarrow$) Suppose $\mathcal{F} \Vdash \Phi(\bar{\alpha}) \vee \Psi(\bar{\alpha})$. Thus there is a cover $\{\beta_k : \mathcal{G}^{(k)} \to \mathcal{F}\}_{k=1}^n$ such that

$$\mathcal{G}^{(k)} \Vdash \Phi(\bar{\alpha} \circ \beta_k) \text{ or } \mathcal{G}^{(k)} \Vdash \Psi(\bar{\alpha} \circ \beta_k) \quad (k = 1, \ldots, n).$$

By using the inductive hypothesis on the disjuncts that happen to hold, we find $j_1, \ldots, j_n$ such that for $k = 1, \ldots, n$,

$$(\forall v \in G_{j_k}^{(k)})\ [\Phi(\bar{\alpha}(\beta_k(v))) \vee \Psi(\bar{\alpha}(\beta_k(v)))].$$

But $\{\beta_k\}$ is a cover, so $\beta[G_{j_1}^{(1)}] \cup \cdots \cup \beta[G_{j_n}^{(n)}]$ includes some F_i. Hence $(\forall u \in F_i)\ [\Phi(\bar{\alpha}(u)) \vee \Psi(\bar{\alpha}(u))]$.

(iv, $\Longleftarrow$) Now assume $(\forall u \in F_i)\ [\Phi(\bar{\alpha}(u)) \vee \Psi(\bar{\alpha}(u))]$. Define filters $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ by

$$u \in G_i^{(1)} \iff u \in F_i \wedge \Phi(\bar{\alpha}(u)),$$
$$u \in G_i^{(2)} \iff u \in F_i \wedge \Psi(\bar{\alpha}(u)).$$

We have trivially $\mathcal{G}^{(1)} \Vdash \Phi(\bar{\alpha})$ and $\mathcal{G}^{(2)} \Vdash \Psi(\bar{\alpha})$ by the definition of the new filters and the inductive hypothesis. By the assumption $\{\mathcal{G}^{(k)} \hookrightarrow \mathcal{F}\}_{k=1}^2$ is a cover. Hence $\mathcal{F} \Vdash \Phi(\bar{\alpha}) \vee \Psi(\bar{\alpha})$.

(v, $\Longleftarrow$) Let i be such that for all $u \in F_i$, $\Phi(\bar{\alpha}(u))$ implies $\Psi(\bar{\alpha}(u))$. Consider $\beta : \mathcal{G} \to \mathcal{F}$ with $\mathcal{G} \Vdash \Phi(\bar{\alpha} \circ \beta)$. By the induction hypothesis, $\Phi(\bar{\alpha}(\beta(v)))$ for all v on some fixed G_j. Since β is continuous, $\beta[G_{j_1}] \subseteq F_i$

for some j_1. Let $G_{j_2} \subseteq G_j \cap G_{j_1}$. Thus using the assumption we get for $v \in G_{j_2}$, $\Psi(\bar{\alpha}(\beta(v)))$. The inductive hypothesis gives $\mathcal{G} \Vdash \Psi(\bar{\alpha} \circ \beta)$.

(v, $\Longrightarrow$) Suppose $\mathcal{F} \Vdash \Phi(\bar{\alpha}(u)) \rightarrow \Psi(\bar{\alpha}(u))$. Define a filter $\mathcal{G}$ by

$$u \in G_i \iff u \in F_i \wedge \Phi(\bar{\alpha}(u)).$$

Then the inclusion $\imath\colon \mathcal{G} \hookrightarrow \mathcal{F}$ is obviously continuous, and we have trivially $\mathcal{G} \Vdash \Phi(\bar{\alpha} \circ \imath)$. Hence $\mathcal{G} \Vdash \Psi(\bar{\alpha} \circ \imath)$. But using the inductive hypothesis, we easily see that this means that there is some i such that for $u \in F_i$: $\Phi(\bar{\alpha}(u)) \Longrightarrow \Psi(\bar{\alpha}(u))$.

(vi, $\Longrightarrow$) This is analogous to (iv, $\Longrightarrow$).

(vi, $\Longleftarrow$) Let i be such that for all $u \in F_i$, $(\exists x \in S)\ \Phi(\bar{\alpha}(u), x)$. Define the filter $\mathcal{G}$ by

$$(u, x) \in G_i \iff u \in F_i \wedge \Phi(\bar{\alpha}(u), x).$$

Then the first projection $\pi_1\colon \mathcal{G} \rightarrow \mathcal{F}$ is a covering map. This follows from the assumption. Hence we have trivially by the induction hypothesis, $\mathcal{G} \Vdash \Phi(\bar{\alpha} \circ \pi_1, \pi_2)$. So by definition, $\mathcal{F} \Vdash (\exists x \in {}^*S)\ \Phi(\bar{\alpha}, x)$.

(vii, $\Longrightarrow$) Assume $\mathcal{F} \Vdash (\forall x \in {}^*S)\ \Phi(\bar{\alpha}, x)$. We have continuous projections $\pi_1\colon \mathcal{F} \times \hat{S} \rightarrow \mathcal{F}$ and $\pi_2\colon \mathcal{F} \times \hat{S} \rightarrow \hat{S}$. Then

$$\mathcal{F} \times \hat{S} \Vdash \Phi(\bar{\alpha} \circ \pi_1, \pi_2).$$

By the induction hypothesis, $\exists i\ (\forall u \in F_i)\ (\forall x \in S)\ \Phi(\bar{\alpha}(u), x)$ since $\mathcal{F} \times \hat{S}$ has base sets of the form $F_i \times S$.

(vii, $\Longleftarrow$) Let $i \in I$ be such that $(\forall u \in F_i)\ (\forall x \in S)\ \Phi(\bar{\alpha}(u), x)$. Let $\beta\colon \mathcal{G} \rightarrow \mathcal{F}$ and $\delta \in {}^*S(\mathcal{G})$. By continuity of β there is some j with $\beta[G_j] \subseteq F_i$. Let $G_k \subseteq G_j$ be a base where δ is defined. Hence $\Phi(\bar{\alpha}(\beta(v)), \delta(v))$ holds for all $v \in G_k$. The inductive hypothesis gives $\mathcal{G} \Vdash \Phi(\bar{\alpha} \circ \beta, \delta)$. Since β and δ were arbitrary, $\mathcal{F} \Vdash (\forall x \in {}^*S)\ \Phi(\bar{\alpha}, x)$. ⊣

REMARK 1. The need to force with nonproper filters becomes apparent, for instance, in the case of disjunction.

We say that a closed (not necessarily internal) formula Φ is *true in the model*, if $\mathcal{F} \Vdash \Phi$, for all $\mathcal{F}$. In symbols we write this as $\Vdash \Phi$. From Theorem 1, the *full transfer principle* follows.

COROLLARY 2 (Transfer). *For internal formulas Φ and proper filters $\mathcal{F}$,*

$$\Phi \textit{ is true} \quad \textit{if and only if} \quad \Vdash \Phi \quad \textit{if and only if} \quad \mathcal{F} \Vdash \Phi.$$

The forcing relation has the following properties which generalise (2) and (4), see [17] for details. For $\beta\colon \mathcal{G} \rightarrow \mathcal{F}$ and a cover $\{\beta_i : \mathcal{G}_i \rightarrow \mathcal{F}\}_{i=1}^n$,

(Monotonicity) $\mathcal{F} \Vdash \Phi(\bar{\alpha}) \Longrightarrow \mathcal{G} \Vdash \Phi(\bar{\alpha} \circ \beta)$,

(Local character) $\mathcal{G}_1 \Vdash \Phi(\bar{\alpha} \circ \beta_1), \dots, \mathcal{G}_n \Vdash \Phi(\bar{\alpha} \circ \beta_n) \Longrightarrow \mathcal{F} \Vdash \Phi(\bar{\alpha})$.

These are, in fact, general properties of any sheaf semantics.

The cost of the full transfer theorem is that the logical symbols are given a *nonstandard* interpretation. Fortunately, the model satisfies intuitionistic logic (and even classical logic if we reason only about internal formulas, but this requires a classical metatheory, see the Appendix). One problem is that we need some way of speaking about standard objects inside the model. This cannot be done as simply as in Robinson's approach, where being standard is equated with being a constant element. This is because of the mismatch of logics. Therefore we introduce for each sort $^*S$, a unary predicate, St^{*S}, for the standard elements of $^*S$. We extend the forcing relation to St^{*S} by the clause

(viii) $\mathcal{F} \Vdash \mathsf{St}^{*S}(\alpha)$ if and only if there is some cover $\{\beta_i : \mathcal{G}^{(i)} \to \mathcal{F}\}_{i=1}^n$, and some elements $s_1, \ldots, s_n \in S$, such that for each i, $\alpha \circ \beta_i$ is identically s_i on some $G_j^{(i)}$.

That is, an object is considered standard if it is *locally constant*. We use the abbreviations $(\forall^{\mathrm{st}} x \in {}^*S)\ldots$ for $(\forall x \in {}^*S)[\mathsf{St}(x) \to \cdots]$, and $(\exists^{\mathrm{st}} x \in {}^*S)\cdots$ for $(\exists x \in {}^*S)[\mathsf{St}(x) \wedge \cdots]$. Then we have the following basic results about standard quantifiers.

LEMMA 3. *Let $\Phi(\bar{x}, y^{*S})$ be an external formula. Then*

(a) $\mathcal{F} \Vdash (\forall^{\mathrm{st}} y \in {}^*S)\ \Phi(\bar{\alpha}, y)$ *if and only if for all* $t \in S$, $\mathcal{F} \Vdash \Phi(\bar{\alpha}, {}^*t)$,

(b) $\mathcal{F} \Vdash (\exists^{\mathrm{st}} y \in {}^*S)\ \Phi(\bar{\alpha}, y)$ *if and only if for some cover* $\{\beta_i : \mathcal{G}_i \to \mathcal{F}\}_{i=1}^n$ *and some* $t_1, \ldots, t_n \in S$

$$\mathcal{G}_1 \Vdash \Phi(\bar{\alpha} \circ \beta_1, {}^*t_1), \ldots, \mathcal{G}_n \Vdash \Phi(\bar{\alpha} \circ \beta_n, {}^*t_n).$$

PROOF. (a, $\Longrightarrow$) is direct. (a, $\Longleftarrow$) follows by using monotonicity and local character of the forcing relation.

(b, $\Longrightarrow$) follows by composition of covers. (b, $\Longleftarrow$) is immediate. $\dashv$

From this lemma follows easily that induction on standard numbers is valid inside the model (see [24]). The St-predicate defines an inner standard world as opposed to the outer standard world. If Φ is a closed internal formula, let Φ^{st} be the result of restricting all quantifiers to standard objects using the standard predicate. It is easy to see that Lemma 3 implies

COROLLARY 4. *If $\Phi(\bar{x}, \bar{y})$ is an internal formula then*

$$\forall \bar{x}\, \exists \bar{y}\ \Phi(\bar{x}, \bar{y}) \textit{ is true if, and only if, } \Vdash \forall^{\mathrm{st}} \bar{x}\, \exists^{\mathrm{st}} \bar{y}\ \Phi(\bar{x}, \bar{y}).$$

THEOREM 5 ([24]). *Suppose Φ is an internal formula which is built up using only $\vee$, $\wedge$, $\exists$ and $\forall$ from quantifier free formulas. Then*:

$$\Phi \textit{ true implies } \Vdash \Phi^{\mathrm{st}}.$$

In particular, this implication holds when Φ is an internal prenex formula.

From this follows that, for example, $\mathsf{St}^{\mathbb{R}}$ has the same algebraic properties as $\mathbb{R}$. However it is doubtful whether $\mathsf{St}^{\mathbb{R}}$ is Cauchy complete. The set of standard functions in $^*(\mathbb{N}^{\mathbb{N}}) = {}^*(\mathbb{N} \to \mathbb{N})$ is too small to allow even the axiom of unique choice, let alone the countable axiom of choice.

PROPOSITION 6. *The following principle does not hold in the model*:

$$(\forall f \in {}^*(\mathbb{N}^{\mathbb{N}}))\ [(\forall^{\mathrm{st}} m)\ (\exists^{\mathrm{st}} n) f(m) = n \rightarrow (\exists^{\mathrm{st}} g \in {}^*(\mathbb{N}^{\mathbb{N}}))\ (\forall^{\mathrm{st}} m)\ f(m) = g(m)].$$

PROOF. For clarity we write $\mathrm{Ap}(f, m)$ for the term $f(m)$. It suffices to find a filter $\mathcal{F}$ and a nonstandard function $\varphi \in {}^*(\mathbb{N} \to \mathbb{N})(\mathcal{F})$ such that

$$(\forall m \in \mathbb{N})\ \mathcal{F} \Vdash (\exists^{\mathrm{st}} y \in {}^*\mathbb{N})\ \mathrm{Ap}(\varphi, m) = y, \tag{5}$$

but for which

$$\mathcal{F} \Vdash (\exists^{\mathrm{st}} g \in {}^*(\mathbb{N} \to \mathbb{N}))(\forall^{\mathrm{st}} m)\, \mathrm{Ap}(\varphi, m) = g(m) \tag{6}$$

is false. Let $\mathcal{F} = (\{0\}, F, \mathbb{N})$ be the trivial filter on $\mathbb{N}$: $F_0 = \mathbb{N}$. Define $h\colon \mathbb{N}^2 \to \mathbb{N}$ by $h(u, m) = u \bmod (m+1)$. This yields a nonstandard function φ at $\mathcal{F}$ by letting $\varphi(u)(m) = h(u, m)$ for $u \in F_0$. Intuitively speaking, $\mathrm{Ap}(\varphi, m)$ takes on all the values $k = 0, 1, \dots, m$ simultaneously. For fixed $m \in \mathbb{N}$, we divide $\mathcal{F}$ in parts $\mathcal{G}^{(k)}$ according to these values. Let $\mathcal{G}^{(k)} = (\{0\}, G^{(k)}, \mathbb{N})$ be the trivial filter

$$G_0^{(k)} = \{\, u \in \mathbb{N} : u \bmod (m+1) = k \,\}.$$

Clearly the inclusions $\eta^{(k)}\colon \mathcal{G}^{(k)} \hookrightarrow \mathcal{F}$, $k = 0, 1, \dots, m$, give a cover of $\mathcal{F}$, and $\mathcal{G}^{(k)} \Vdash \mathrm{Ap}(\varphi \circ \eta^{(k)}, m) = k$ is trivially true. Thus (5) holds. Suppose that (6) holds. Thus there are $g_1, \dots, g_\ell\colon \mathbb{N} \to \mathbb{N}$ and a cover $\{\psi^{(k)}\colon \mathcal{H}^{(k)} \to \mathcal{F}\}_{k=1}^{\ell}$ such that for all $m \in \mathbb{N}$,

$$\mathcal{H}^{(k)} \Vdash \mathrm{Ap}(\varphi \circ \psi^{(k)}, m) = g_k(m).$$

Consider $m = \ell$. Thus for some $q_1, \dots, q_\ell$ we have

$$g_k(m) = \mathrm{Ap}(\varphi(\psi^{(k)}(u)), m) \quad (u \in H_{q_k}^{(k)}). \tag{7}$$

By the cover property

$$\psi^{(1)}[H_{q_1}^{(1)}] \cup \cdots \cup \psi^{(\ell)}[H_{q_\ell}^{(\ell)}] \supseteq F_0 = \mathbb{N}.$$

The k equations (7) and the cover property says that for all $v \in \mathbb{N}$,

$$\mathrm{Ap}(\varphi(v), m) = h(v, m) \in \{g_1(m), \dots, g_\ell(m)\}.$$

But this is impossible since $h(\cdot, m)$ assumes $m + 1$ different values. ⊣

Nevertheless we have the following useful substitute for countable choice.

THEOREM 7 (Monotone Choice). *For an external formula $\Phi(\bar{x}, m, n)$, the following is true in the model: Suppose Φ is monotone in n, i.e., $\Phi(\bar{x}, m, n)$ and $n \leq k$ implies $\Phi(\bar{x}, m, k)$. If for all standard $m \in {}^*\mathbb{N}$ there is a standard $n \in {}^*\mathbb{N}$ with $\Phi(\bar{x}, m, n)$, then there is a standard choice function $f \in {}^*(\mathbb{N} \to \mathbb{N})$ such that for all standard $m \in {}^*\mathbb{N}$, $\Phi(\bar{x}, m, f(m))$.*

PROOF. Suppose that

(8) $$\mathcal{F} \Vdash (\forall m, n, k \in {}^*\mathbb{N})\, [\Phi(\bar{\alpha}, m, n) \wedge n \leq k \Longrightarrow \Phi(\bar{\alpha}, m, k)]$$

and $\mathcal{F} \Vdash (\forall^{\mathrm{st}} m \in {}^*\mathbb{N})(\exists^{\mathrm{st}} n \in {}^*\mathbb{N})\, \Phi(\bar{\alpha}, m, n)$. Then by Lemma 3, we have for each $m \in \mathbb{N}$, a cover $\{\beta_k : \mathcal{G}^{(k)} \to \mathcal{F}\}_{k=1}^{p}$ with witnesses $n_1, \ldots, n_p \in \mathbb{N}$ such that

$$\mathcal{G}^{(k)} \Vdash \Phi(\bar{\alpha} \circ \beta_k, {}^*m, {}^*n_k) \quad (k = 1, \ldots, p).$$

Let $f(m) = \max(n_1, \ldots, n_p)$. By the monotonicity of forcing, the assumption (8) holds at each $\mathcal{G}^{(k)}$. Thus

$$\mathcal{G}^{(k)} \Vdash \Phi(\bar{\alpha} \circ \beta_k, {}^*m, {}^*f({}^*m)).$$

Since $\{\beta_k\}$ is a cover, we get

$$\mathcal{F} \Vdash \Phi(\bar{\alpha}, {}^*m, {}^*f({}^*m))$$

for all $m \in \mathbb{N}$. Thus

$$\mathcal{F} \Vdash (\exists^{\mathrm{st}} f)\, (\forall^{\mathrm{st}} m)\, \Phi(\bar{\alpha}, m, f(m)). \qquad \dashv$$

An example of a formula *monotone* in n is

$$\Phi(a, b, m, n) \equiv (\forall k \geq n)\, |a_k - b| < 2^{-m}.$$

Using the above choice principle we can extract a convergence modulus for a convergent sequence. Continuity and differentiability moduli may be extracted analogously by considering appropriate monotone formulas. As we shall see in Corollary 10 below, it possible to obtain a stronger choice theorem for internal formulas.

The third important result we have is the *idealisation principle*. It is a type theoretic version of Nelson's I-axiom.

THEOREM 8 (Idealisation). *Let $\Phi(\bar{x}, y, z)$ be an internal formula. The following result holds in the model: If for all standard $n \in {}^*\mathbb{N}$ and all standard $f \in {}^*(\mathbb{N} \to S)$ there exists $y \in {}^*T$ such that*

$$(\forall k < n)\, \Phi(\bar{x}, y, f(k)),$$

then there is $y \in {}^*T$ so that for all standard $z \in {}^*S$, $\Phi(\bar{x}, y, z)$.

PROOF. By Lemma 3 it is enough to suppose that for all n and all $f \in \mathbb{N} \to S$, $\mathcal{F} \Vdash (\exists y \in {}^*T)\, (\forall k < {}^*n)\, \Phi(\bar{\alpha}, y, {}^*f(k))$. Form the filter base $\mathcal{G}$ by defining

$$(u, y) \in G_{p,n,f} \iff u \in F_p \wedge (\forall k < n)\, \Phi(\bar{\alpha}(u), y, f(k)).$$

This can be seen to be a filter by considering concatenation of finite sequences. From the assumption it follows that $\pi_1 : \mathcal{G} \to \mathcal{F}$ is a cover. To prove the theorem it suffices to check $\mathcal{G} \Vdash \Phi(\bar{\alpha} \circ \pi_1, \pi_2, {}^*r)$ for every $r \in S$. But for fixed r this statement is, by Theorem 1, equivalent to

$$\exists p, n, f\ (\forall u \in F_p)\, \forall y\ [(\forall k < n)\Phi(\bar{\alpha}(u), y, f(k)) \Longrightarrow \Phi(\bar{\alpha}(u), y, r)].$$

This is trivially true putting $f(x) = r$ and $n = 1$. ⊣

Here are some definitions which are meaningful inside the model. A number x is *finite* if for some standard n, $|x| < n$, or formally $\exists^{\mathrm{st}} n\ (|x| < n)$; the number is *infinite* if $|x| > n$ for all standard n, i.e., formally $\forall^{\mathrm{st}} n\ (|x| > n)$. We use the abbreviation $\forall^{\mathrm{inf}} x\ \dots$ for $\forall x\ [\forall^{\mathrm{st}} n\ (|x| > n) \to \dots]$. Similarly, $\exists^{\mathrm{inf}} x\ \dots$ abbreviates $\exists x\ [\forall^{\mathrm{st}} n\ (|x| > n) \wedge \dots]$. For some properties of these notions, see [24].

Corollary 9 (Overspill Principle [17]). *Let $\Psi(\bar{x}, n)$ be internal. Then in the model: If for all standard $n \in {}^*\mathbb{N}$, $\Psi(\bar{x}, n)$, then there is some infinite $\nu \in {}^*\mathbb{N}$ such that $\Psi(\bar{x}, \nu)$.*

Proof. Apply the theorem to $\Phi(\bar{x}, y, n) \equiv y \geq n \wedge \Psi(\bar{x}, y)$. ⊣

Corollary 10. *Let $\Phi(\bar{x}, m, n)$ be an internal formula. Then the following holds in the model: If for each standard $m \in {}^*\mathbb{N}$, there is a standard $n \in {}^*\mathbb{N}$ with $\Phi(\bar{x}, m, n)$, then there exists a choice function $f \in {}^*(\mathbb{N} \to \mathbb{N})$ such that $f(m)$ is standard and $\Phi(\bar{x}, m, f(m))$ for all standard $m \in {}^*\mathbb{N}$.*

Proof. We may render the assumption as

$$\forall^{\mathrm{st}} m\ \exists^{\mathrm{st}} k\ (\exists n < k)\ \Phi(\bar{x}, m, n)$$

and apply monotone choice to obtain a standard function $g \in {}^*(\mathbb{N} \to \mathbb{N})$ such that for all standard m

$$(\exists n < g(m))\ \Phi(\bar{x}, m, n).$$

By induction on k we see that for all standard k

$$(\exists f \in {}^*(\mathbb{N} \to \mathbb{N}))\ (\forall m < k)\ f(m) < g(m) \wedge \Phi(\bar{x}, m, f(m)).$$

Since this statement is internal, we conclude by overspill that it is valid for some infinite k. Thus for each standard m, $f(m) < g(m)$ and hence $f(m)$ is standard. ⊣

The choice function obtained cannot be assumed to be standard, as we have seen in Proposition 6.

Corollary 11 (Hyperfinite Enumeration [24]). *Let S be a nonvoid set. Then in the model: there is a function $f \in {}^*(\mathbb{N} \to S)$ and a natural number ν such that for every standard $s \in {}^*S$ there is an index $i < \nu$ with $s = f(i)$.*

Proof. Apply the theorem to $\Phi(y, z) \equiv (\exists i < m)(z = g(i))$, where $y = (m, g) \in {}^*(\mathbb{N} \times (\mathbb{N} \to S))$. ⊣

Note that the overspill principle also implies the existence of infinite natural numbers. Classically, the *underspill principle* is just the contraposition of overspill, but constructively it seems to need a separate proof.

Theorem 12. *Let $\Psi(\bar{x}, n)$ be internal. Then the following holds in the model: If $\Psi(\bar{x}, \nu)$, for all infinite $\nu \in {}^*\mathbb{N}$, then for some standard $n \in {}^*\mathbb{N}$, $\Psi(\bar{x}, n)$.*

PROOF. Assume $\mathcal{H} \Vdash \forall^{\inf} \nu\, \Psi(\bar{\alpha}, \nu)$. We shall use this fact for a well-chosen infinite number. Let $\mathcal{F}$ be the Fréchet filter given by $F_n = \{ m \in \mathbb{N} : m \geq n \}$. The product $\mathcal{H} \times \mathcal{F}$ is then determined by the base sets $H_p \times F_n$. The projection $\pi_2 \colon \mathcal{H} \times \mathcal{F} \to \mathbb{N}$ satisfies $\mathcal{H} \times \mathcal{F} \Vdash \forall^{\mathrm{st}} n\, |\pi_2| > n$. The projection $\pi_1 \colon \mathcal{H} \times \mathcal{F} \to \mathcal{H}$ is continuous, so by assumption $\mathcal{H} \times \mathcal{F} \Vdash \Psi(\bar{\alpha} \circ \pi_1, \pi_2)$. But this means,

$$\exists p\, \exists n\, (\forall u \in H_p)(\forall m \in F_n)\, \Psi(\bar{\alpha}(u), m),$$

so in particular, putting $m = n$, we obtain $\mathcal{H} \Vdash \Psi(\bar{\alpha}, n)$. ⊣

An *external set* of the model is a sheaf $\mathsf{F} \colon \mathbb{B}^{\mathrm{op}} \to \mathrm{Sets}$, together with an equality sheaf relation $=_{\mathsf{F}}$ on F. As is well-known from topos theory [14] we may construct function space sheaves and subsheaves. These constructions can be made predicatively, whereas the powersheaves cannot. The external formulas, and the sheaf semantics, are naturally expanded to include also quantification over external sets. We show here how subsheaves may be constructed. Let $\Phi(x)$ be an external formula where x ranges over the sort F corresponding to the sheaf F. Define a presheaf $\mathsf{G} \colon \mathbb{B}^{\mathrm{op}} \to \mathrm{Sets}$ by letting $\mathsf{G}(\mathcal{H})$ be the set of $\alpha \in \mathsf{F}(\mathcal{H})$ such that $\mathcal{H} \Vdash \Phi(\alpha)$, and letting G act on morphisms as F does. It follows from the monotonicity and local character of forcing relation that G is a sheaf. The equality of G is that of F. The inclusion maps $i_{\mathcal{H}}$ from $\mathsf{G}(\mathcal{H})$ into $\mathsf{F}(\mathcal{H})$, $\mathcal{H} \in \mathbb{B}$, make up a natural transformation $i \colon \mathsf{G} \stackrel{\cdot}{\to} \mathsf{F}$ so that in the model holds:

$$(\forall x, y \in \mathsf{G})[x =_{\mathsf{G}} y \leftrightarrow i(x) =_{\mathsf{F}} i(y)]$$

and

$$(\forall u \in \mathsf{F})[\Phi(u) \leftrightarrow (\exists x \in \mathsf{G})\, u =_{\mathsf{F}} i(x)].$$

For instance, if $\mathsf{F} = {}^*S$ and $\Phi(x)$ is the formula $\mathsf{St}(x)$ then G is the subsheaf of ${}^*S$ consisting of standard elements; we write naturally S for this sheaf.

REMARKS 1. Here we have considered only the first-order part of the topos $\mathrm{Sh}(\mathbb{B})$, the reason being that it is well behaved with respect to transfer properties (cf. Corollary 4 and Theorem 5). In particular the function spaces in the various versions of choice principles are not the canonical ones expressed by the cartesian closure of the topos.

The new results of this section are the monotone choice theorem and the underspill principle. The idealisation theorem is an improved formulation of [24, Theorem 3.9].

§3. Some nonconstructive features of the model. We show that the overspill principle has some consequences for the inner standard world which are not accepted in strict constructivism. There is no contradiction in this, since there is no strong relation between the inner standard world and the metatheory. We may hope to use the nonconstructive features to simplify proofs about those concrete statements that are indeed subject to the transfer principle. This is a kind of elimination of nonconstructive ideals in the spirit of Hilbert's Programme. As examples we show König's lemma and the de Bruijn–Erdös theorem on graph colouring. Further investigation is necessary in order to judge whether such nonconstructive reasoning can be fruitful. For some interesting applications, see Coquand [7] and the references in there.

To distinguish degrees of nonconstructivity Brouwer introduced various principles of *omniscience* (Bishop's terminology). None of following limited principles is valid in the metatheory. Let (a_n) be an infinite sequence of natural numbers.

(LPO) Either $\forall n\ a_n = 0$ or $\exists n\ a_n \neq 0$.

(LLPO) If at most one a_n is nonzero, then $a_{2n} = 0$ for all n, or $a_{2n+1} = 0$ for all n.

Clearly, the principle LPO implies LLPO, but the converse is not true. For discussions about these principles and their relation to various classical theorems, see Bridges and Richman [4], and Richman [25]. Somewhat surprisingly we have a version of LLPO in the model.

THEOREM 13. *The following holds in the model*: *If* (a_n) *is a standard sequence of natural numbers, at most one nonzero, then* $a_{2n} = 0$ *for all standard* n, *or* $a_{2n+1} = 0$ *for all standard* n. *In short this is* LLPO$^{\mathrm{st}}$.

PROOF. We work within the model. Under the assumption on (a_n) we can prove by induction on standard numbers that

$$\forall^{\mathrm{st}} m\ [(\forall n < m)\ a_{2n} = 0 \vee (\forall n < m)\ a_{2n+1} = 0].$$

Thus by overspill, this holds for some infinite m. Hence $\forall^{\mathrm{st}} n\ (a_{2n} = 0)$ or $\forall^{\mathrm{st}} n\ (a_{2n+1} = 0)$. ⊣

However the stronger LPO does not hold in the model unless it holds in the metatheory.

PROPOSITION 14. $\Vdash$ LPO$^{\mathrm{st}}$ *implies that* LPO *holds in the metatheory.*

PROOF. Suppose

$$\Vdash \forall^{\mathrm{st}} n\ {}^*a_n = 0 \vee \exists^{\mathrm{st}} n\ {}^*a_n \neq 0.$$

Hence there are $n_1, \dots, n_p \in \mathbb{N}$ and a cover $\{\mathcal{F}^{(i)} \to 1\}_{i=1}^{p}$ such that for all i,

$$\mathcal{F}^{(i)} \Vdash \forall^{\mathrm{st}} n\ {}^*a_n = 0 \quad \text{or} \quad \mathcal{F}^{(i)} \Vdash {}^*a_{n_i} \neq 0.$$

Now if $a_{n_1} = a_{n_2} = \cdots = a_{n_p} = 0$ then $\mathcal{F}^{(i)} \Vdash \forall^{\mathrm{st}} n\ {}^*a_n = 0$, for all $i = 1, \dots, p$. Hence by local character, $\Vdash \forall^{\mathrm{st}} n\ {}^*a_n = 0$. Transfer then gives $a_n = 0$ for all n. Otherwise $a_{n_i} \neq 0$ for some $i = 1, \dots, p$. Hence LPO is proved. ⊣

EXAMPLE 1. A graph is *n-colourable* if its nodes can be coloured using n colours such that no adjacent nodes have the same colour. The de Bruijn–Erdös theorem on graph colouring states that an infinite graph $G \subseteq \mathbb{N} \times \mathbb{N}$

may be n-coloured, if all of its finite subgraphs may be n-coloured. This follows almost immediately from the overspill principle. We work within the model. The assumption is rendered

$$(9)\quad \forall^{\mathrm{st}} k\ (\exists c \in {}^*(\mathbb{N} \to \mathbb{N}))\ [(\forall i \leq k)\ (c(i) \leq {}^*n) \wedge (\forall i, j \leq k)\ (G(i, j) \to c(i) \neq c(j))].$$

Since the formula within square brackets is internal, we find according to the overspill principle an infinite k and a function $c \in {}^*(\mathbb{N} \to \mathbb{N})$ such that the formula holds. All standard i, j are clearly less than k. Hence for all standard i, j: if $G(i, j)$ then $c(i) \neq c(j)$, and $c(i) \leq {}^*n$ for all standard i. That is we have an ${}^*n$-colouring of G.

EXAMPLE 2. A weak, yet nonconstructive, version of König's lemma is: each infinite, decidable binary tree has an infinite branch. We prove this in the model. Assume a 1-1 coding of binary strings x as natural numbers: $|x|$ denotes the length of the string and $x \subseteq y$ means that x is an initial substring of y. We assume a coding method such that $w \leq 4^{|w|}$. Let $x_{[n]}$ denote the truncation of x to n bits. We may then represent a binary tree by a function f on natural numbers, which takes the value 0 if the argument binary string describes a path within the tree. We also require that

$$\forall^{\mathrm{st}} xy\ [f(y) = 0 \wedge x \subseteq y \Longrightarrow f(x) = 0];$$

call the formula expressing this $\Phi(f)$. The formal statement of *weak König lemma* is then

$$\Phi(f) \wedge \forall^{\mathrm{st}} n\ \exists^{\mathrm{st}} x\ [|x| = n \wedge f(x) = 0] \Longrightarrow \exists g\ \forall^{\mathrm{st}} n[|g(n)| = n \wedge f(g(n)) = 0 \wedge g(n) = g(n+1)_{[n]}].$$

According to the assumption of this statement we have for each standard n,

$$\forall x\ \forall y\ [|x|, |y| \leq n \wedge f(y) = 0 \wedge x \subseteq y \Longrightarrow f(x) = 0],$$
$$\exists w\ [|w| = n \wedge f(w) = 0].$$

The overspill principle thus yields some infinite string w with $f(w) = 0$, and moreover writing $v = |w|$,

$$(10)\qquad (\forall |x| \leq v)(\forall |y| \leq v)\ [f(y) = 0 \wedge x \subseteq y \Longrightarrow f(x) = 0].$$

Let $g(n) = w_{[n]}$. We have for standard n that $n < v$, so by (10)

$$f(g(n)) = f(w_{[n]}) = 0,$$

and, since $w_{[n]} \leq 4^n$, $w_{[n]}$ must be standard.

Brouwer suggested his Fan Theorem as an intuitionistic substitute for the König lemma, one that is nevertheless not accepted in Bishop's strict constructivism. We show that a version of this "theorem" is valid in the following form

$$\Phi(f) \wedge \forall g\, [\forall^{\mathrm{st}} n\, |g(n)| = n \wedge g(n) = g(n+1)_{[n]} \Longrightarrow \exists^{\mathrm{st}} m\, f(g(m)) > 0] \\ \Longrightarrow \exists^{\mathrm{st}} n\, \forall w\, [|w| = n \Longrightarrow f(w) > 0].$$

By the underspill principle we need only to show that $f(w) > 0$, whenever w is of infinite length. Suppose that w is such a string, and define $g(n) = w_{[n]}$. From the assumption in the theorem it follows that there is some standard m, $f(g(m)) > 0$. Obviously $w \supseteq w_{[m]}$, so since f is a tree, also $f(w) > 0$.

§**4. Nonstandard characterisations.** Using the overspill and underspill principles we may prove nonstandard characterisations of standard notions inside the model. By the extended transfer principle, Corollary 4, these characterisations will also be characterisations of the corresponding notions in the metatheory. It is also possible to give direct characterisations of metatheoretical notions, see [24]. The Schmieden–Laugwitz approach is confined to characterising topological notions of a 'sequential' nature, and this does not give satisfactory characterisations of 'uniform' notions in a constructive setting [22, 23].

Recall that Bishop's constructive reals $\mathbb{R}$ are sequences of rational numbers $x = (x_n)_n$ satisfying the regularity condition

$$|x_m - x_n| \leq_{\mathbb{Q}} \frac{1}{m} + \frac{1}{n} \quad (m, n \in \mathbb{Z}_+).$$

Such a real is *positive* if for some $n \in \mathbb{Z}_+$, $1/n <_{\mathbb{Q}} x_n$. By $x <_{\mathbb{R}} y$ we mean of course that $y - x$ is positive. We refer to [1] or [2] for further background. In the following we let $\mathbb{F}$ denote either of the constructive fields $\mathbb{Q}$ or $\mathbb{R}$.

Definition 2. For a metric space (X, d) with metric $d\colon X \times X \to \mathbb{F}$ we define the following nonstandard notions.

(a) The points x and y in ${}^*X$ are *infinitely close* if ${}^*d(x, y) < 2^{-k}$ for all standard $k \in {}^*\mathbb{N}$. In symbols we write this as $x \simeq_d y$, or just $x \simeq y$ when d is evident from the context.
(b) A point $x \in {}^*X$ is *finite* if ${}^*d(x, y) < n$ for some standard $y \in {}^*X$ and some standard $n \in {}^*\mathbb{N}$.

Theorem 15. *Let (X, d) and (Y, e) be metric spaces whose metrics take values in $\mathbb{F}$. Let M be a subset of X. Then the following conditions on a function $f \in {}^*(M \to Y)$ in the model are equivalent*:

(a) $\forall^{\mathrm{st}} m\, \exists^{\mathrm{st}} n\, (\forall x, y \in {}^*M)\, [d(x, y) < 2^{-n} \Longrightarrow e(f(x), f(y)) < 2^{-m}]$,
(b) $(\forall x, y \in {}^*M)\, [x \simeq_d y \Longrightarrow f(x) \simeq_e f(y)]$,
(c) $(\forall x \in {}^*M)\, \forall^{\mathrm{st}} m\, \exists^{\mathrm{st}} n\, (\forall y \in {}^*M)\, [d(x, y) < 2^{-n} \Longrightarrow e(f(x), f(y)) < 2^{-m}]$.

PROOF. We first prove the most useful equivalence, that between (a) and (b). The direction from (a) to (b) is straightforward. Conversely, suppose (b). For a fixed standard m, we have for all infinite n

$$(\forall x, y \in {}^*M)\,[d(x,y) < 2^{-n} \Longrightarrow e(f(x), f(y)) < 2^{-m}].$$

Since the statement is internal, we may use underspill and thus find some standard n for which it is true.

The equivalence between (a) and (c) follows by a straightforward application of underspill. ⊣

An internal function f satisfying (a) above is called *S-uniformly continuous.* (The prefix "S" is commonly used in nonstandard analysis to indicate that a notion uses a comparison of hyperreals to standard reals.) This notion characterises uniform continuity in the metatheory.

COROLLARY 16. *Let (X,d) and (Y,e) be metric spaces whose metrics take values in $\mathbb{F}$. Let $f : M \to Y$, where M is a subset of X. Then the following are equivalent*:

(a) *f is uniformly continuous on M*,
(b) *in the model*: *${}^*f$ is S-uniformly continuous on ${}^*M$.*

PROOF. This follows using Corollary 4. ⊣

We introduce two notions of derivative. For S-uniformly continuous functions $f, g \in {}^*(I \to \mathbb{F})$ we say that g is a (*strong*) *S-derivative* of f if for each standard m there is a standard n so that

$$(11)\quad (\forall x, y \in {}^*I)\,\forall p \left[2^{-p} < |x-y| < 2^{-n} \Longrightarrow \left|\frac{f(x)-f(y)}{x-y} - g(x)\right| < 2^{-m}\right].$$

And we say that g is a *weak S-derivative*, or briefly a *W-derivative*, of f means the same thing except that in (11) p is restricted to standard numbers.

A comment on partial functions is in order here. Since we are working in a first-order language, these must be made into total functions by introducing extra parameters. The inverse value of a real x bounded away from zero by a positive rational is a function of this bound also. We have

$$|x| > 2^{-n} \Longrightarrow x \cdot \imath(x,n) =_{\mathbb{R}} 1.$$

Naturally $\imath(x,m) =_{\mathbb{R}} \imath(x,n)$ for $|x| > \max(2^{-m}, 2^{-n})$. Usually n is suppressed and we write x^{-1} or $1/x$ for $\imath(x,n)$.

THEOREM 17. *Let I be a subset of $\mathbb{F}$. Then, in the model, let f, $g \in {}^*(I \to \mathbb{F})$. The following two statements are equivalent*:

(a) *g is an S-derivative of f*,
(b) $(\forall x, y \in {}^*I)\,\forall p\,[2^{-p} < |x-y| \simeq 0 \Longrightarrow (f(x)-f(y))/(x-y) \simeq g(x)]$.

Moreover (c) *and* (d) *below are equivalent*:

(c) *g is a W-derivative of f,*
(d) $\exists^{\inf} p\ (\forall x, y \in {}^*I)\ [2^{-p} < |x - y| \simeq 0 \Longrightarrow (f(x) - f(y))/(x - y) \simeq g(x)]$.

Here $z/(x - y)$ is, more precisely, $z \cdot \iota(x - y, p)$.

Proof. The equivalence between (a) and (b) is entirely analogous to that of Theorem 15.

(c) $\Longrightarrow$ (d): From (c) we obtain by monotone choice a standard function $b\colon {}^*(\mathbb{N} \to \mathbb{N})$ such that for all standard ℓ and all $p, m \leq \ell$

$$(12)\quad (\forall x, y \in {}^*I)\ \left[2^{-p} < |x - y| < 2^{-b(m)} \to \left|\frac{f(x) - f(y)}{x - y} - g(x)\right| < 2^{-m}\right].$$

By overspill there is an infinite ℓ for which this holds. Let $x, y \in {}^*I$ be such that $2^{-\ell} < |x - y| \simeq 0$. Then by (12) we have

$$\frac{f(x) - f(y)}{x - y} \simeq g(x),$$

since $b(m)$ is standard for each standard m.

(d) $\Longrightarrow$ (c): Let p be infinite as in (d). For a fixed standard m, we have for all infinite n

$$(13)\quad (\forall x, y \in {}^*I)\ \left[2^{-p} < |x - y| < 2^{-n} \Longrightarrow \left|\frac{f(x) - f(y)}{x - y} - g(x)\right| < 2^{-m}\right].$$

Thus by underspill there is a standard n such that this holds. Obviously p may be replaced by any standard natural number. $\dashv$

Any infinitesimal $\varepsilon > 2^{-p}$ in (d) above is called an *admissible differentiation mesh for f.*

Corollary 18. *Let $\tilde{I}$ be an interval of $\mathbb{F}$, and let $f, g\colon \tilde{I} \to \mathbb{F}$ be uniformly continuous. Then the following statements are equivalent*:

(a) *g is the derivative of f,*
(b) *in the model: ${}^*g$ is the S-derivative of ${}^*f$,*
(c) *in the model: ${}^*g$ is the W-derivative of ${}^*f$.*

Proof. The equivalence between (a) and (b) follows from Theorem 17 and Corollary 4. Clearly, (b) implies (c). To show that (c) implies (a) we use Theorem 17, and note that the formula in Theorem 17 (c) is monotone in n. Hence we may proceed similarly as in the proof of the monotone choice theorem and for each $m \in \mathbb{N}$ obtain an $n \in \mathbb{N}$ such that

$$\Vdash (\forall x, y \in {}^*I)\, \forall^{\mathrm{st}} p$$
$$\left[2^{-p} < |x - y| < 2^{-{}^*n} \Longrightarrow \left| \frac{{}^*f(x) - {}^*f(y)}{x - y} - {}^*g(x) \right| < 2^{-{}^*m} \right].$$

Then (a) follows by Corollary 4. ⊣

Different notions of S-limit and S-convergence are distinguished. Let (a_n) be a sequence of points in the metric space $({}^*X, {}^*d)$. We say that b is a *short S-limit* of the sequence, if for all standard k there is a standard ℓ such that for all *standard* $n \geq \ell$, ${}^*d(a_n, b) < 2^{-k}$. On the other hand b is a *long S-limit* of the sequence, if for all standard k there is a standard ℓ such that for *all* $n \geq \ell$, ${}^*d(a_n, b) < 2^{-k}$. The sequence is a *short Cauchy sequence* if for all standard k there is some standard ℓ, such that for all *standard* $m, n \geq \ell$, ${}^*d(a_m, a_n) < 2^{-k}$. The sequence is a *long Cauchy sequence* if for all standard k there is a standard ℓ such that for *all* $m, n \geq \ell$, ${}^*d(a_m, a_n) < 2^{-k}$.

THEOREM 19. *Let (X, d) be a metric space whose metric takes values in $\mathbb{F}$. In the model, consider a sequence of points $a \in {}^*(\mathbb{N} \to X)$ and a point $b \in {}^*X$. Then*:

(i) *b is a long S-limit of (a_n) if, and only if,* $\forall^{\mathrm{inf}} n\, [a_n \simeq b]$,
(ii) *b is a short S-limit of (a_n) if, and only if,* $\exists^{\mathrm{inf}} \nu\, (\forall^{\mathrm{inf}} n \leq \nu)\, [a_n \simeq b]$,
(iii) *(a_n) is a long Cauchy sequence if, and only if,* $\forall^{\mathrm{inf}} m, n\, [a_m \simeq a_n]$,
(iv) *(a_n) is a short Cauchy sequence if, and only if,* $\exists^{\mathrm{inf}} \nu\, (\forall^{\mathrm{inf}} m, n \leq \nu)\, [a_m \simeq a_n]$.

PROOF. Part (i) and (iii) are straightforward by underspill. Part (ii) and (iv) are analogous to Theorem 17. ⊣

However the 'long' and 'short' notions coincide for a real-world sequence ${}^*a$.

COROLLARY 20. *Let (X, e) be a metric space and let (a_n) be a sequence in X and b a point in X. Then* (i)–(iii) *below are equivalent, as well as* (iv)–(vi).

(i) $b = \lim_n a_n$.
(ii) *In the model*: *${}^*b$ is a long S-limit of $({}^*a_n)$.*
(iii) *In the model*: *${}^*b$ is a short S-limit of $({}^*a_n)$.*
(iv) *(a_n) is a Cauchy sequence.*
(v) *In the model*: *$({}^*a_n)$ is a long Cauchy sequence.*
(vi) *In the model*: *$({}^*a_n)$ is a short Cauchy sequence.*

PROOF. Use the idea from monotone choice as in Corollary 18. ⊣

REMARK 2. In a model where transfer holds internally, such as a classical model, the strong and weak S-derivatives coincide for standard functions. Analogously short and long Cauchy sequence are the same for standard sequences.

§5. Nonstandard proofs of standard theorems. The first use of the model is a conservative one. We use it to establish standard results. In this section we employ the usual 'strong' and 'long' versions of the nonstandard characterisations.

THEOREM 21. *Let f and g be differentiable real functions on the compact interval $\tilde{I}$, with derivatives f' and g' respectively. Then*:

(a) *$h = f + g$ is differentiable and its derivative is $h' = f' + g'$.*
(b) *$h = fg$ is differentiable and its derivative is $h' = fg' + f'g$.*
(c) *$h = 1/f$ is differentiable and its derivative is $h' = -f'/f^2$, provided that for some n: $|f(x)| > 2^{-n}$ whenever $x \in \tilde{I}$.*

PROOF. We use Corollary 18 and transfer. Parts (a) and (b) are straightforward and are left to the reader. Part (c): Working inside the model assume $x, y \in \tilde{I}$ are such that $|x - y| > 2^{-p}$ and $x \simeq y$. By transfer we have $|{}^*f(x)|$, $|{}^*f(y)| > 2^{-{}^*n}$. Then

$$\begin{aligned}
\frac{{}^*h(x) - {}^*h(y)}{x - y} &= \frac{1}{x - y}\left[\frac{1}{{}^*f(x)} - \frac{1}{{}^*f(y)}\right] \\
&= \frac{1}{x - y} \cdot \frac{{}^*f(y) - {}^*f(x)}{{}^*f(x){}^*f(y)} \\
&= -\frac{{}^*f(x) - {}^*f(y)}{x - y} \cdot \frac{1}{{}^*f(x){}^*f(y)} \\
&\simeq -\frac{{}^*f'(x)}{{}^*f(x){}^*f(x)}.
\end{aligned}$$

In the last member we have used Corollary 18 and uniform continuity of x^{-1} on $[2^{-n}, \infty)$. Above $z/(x - y)$ means $z \cdot \imath(x - y, p)$, $z/{}^*f(x)$ means $z \cdot \imath({}^*f(x), {}^*n)$ and $z/({}^*f(u){}^*f(v))$ means $z \cdot \imath({}^*f(u){}^*f(v), 2n)$. ⊣

This proof is close to the classical nonstandard proof, and thus avoids the somewhat intricate estimates of the standard constructive proof (cf. Bishop and Bridges [2, p. 45]).

THEOREM 22 (Chain rule). *Let $\tilde{I}$ and $\tilde{J}$ be compact intervals and let $f : \tilde{J} \to \mathbb{R}$ and $g : \tilde{I} \to \tilde{J}$ be differentiable functions. Then the composition $h = f \circ g$ is differentiable and its derivative is $h' = (f' \circ g) \cdot g'$.*

PROOF. Let m_f be the continuity modulus of f. We use the characterisation of Corollary 18 and work inside the model. Suppose x, $y \in \tilde{I}$ is such that $0 \simeq |x - y| > 2^{-p}$. Let $\delta = x - y$ and $\varepsilon = {}^*g(x) - {}^*g(y)$. By uniform continuity $\varepsilon \simeq 0$. We need to make a constructive version of the classical case distinction between $\varepsilon = 0$ and $\varepsilon \neq 0$. Since $\delta > 0$ we find some n such that $2^{-n} < \min(\delta^2, {}^*m_f(\delta^2))$. By the constructive substitute for linear order we now have two possibilities for ε,

$$2^{-n} < |\varepsilon| \quad \text{or} \quad |\varepsilon| < \min(\delta^2, {}^*m_f(\delta^2)).$$

In the first case, we have by the characterisation of derivative, and the uniform continuity of $\cdot$ on compact intervals,

$$\begin{aligned}\frac{{}^*f({}^*g(x)) - {}^*f({}^*g(y))}{x-y} &= \frac{{}^*f({}^*g(x)) - {}^*f({}^*g(y))}{{}^*g(x) - {}^*g(y)} \cdot \frac{{}^*g(x) - {}^*g(y)}{x-y} \\ &\simeq {}^*f'({}^*g(x)) \cdot {}^*g'(x).\end{aligned}$$

Here $z/(x-y)$ means $z \cdot \iota(x-y, p)$ and $z/({}^*g(x) - {}^*g(y))$ is short for $z \cdot \iota({}^*g(x) - {}^*g(y), n)$.

Consider the second case, $|\varepsilon| < \min(\delta^2, {}^*m_f(\delta^2))$. On the one hand we have

$$\left|\frac{{}^*f({}^*g(x)) - {}^*f({}^*g(y))}{x-y}\right| \le \frac{|x-y|^2}{|x-y|} = |x-y| \simeq 0.$$

On the other hand

$$\begin{aligned}|{}^*f'({}^*g(x)) \cdot {}^*g'(x)| &\simeq |{}^*f'({}^*g(x))| \cdot \left|\frac{{}^*g(x) - {}^*g(y)}{x-y}\right| \\ &\le |{}^*f'({}^*g(x))| \cdot \frac{|x-y|^2}{|x-y|} \\ &= |{}^*f'({}^*g(x))| \cdot |x-y| \simeq 0. \qquad \dashv\end{aligned}$$

Integration. A finite sequence $P = (x_0, \dots, x_n)$ is called a *partition* of the interval $[a, b]$ if

$$x_0 = a \le x_1 \le \cdots \le x_n = b.$$

The *mesh* of P is defined by

$$\mathrm{mesh}(P) = \min(x_1 - x_0, \dots, x_n - x_{n-1}).$$

If f is uniformly continuous on $[a, b]$ with continuity modulus m_f and $\varepsilon > 0$, then for any subdivision $P = (x_0, \dots, x_n)$ with $\mathrm{mesh}(P) \le m_f(\varepsilon)$ and any numbers $t_i \in [x_i, x_{i+1}]$ $(i = 0, \dots, n-1)$ we have the basic estimate

$$\left|\int_a^b f(x)dx - \sum_{i=0}^{n-1} f(t_i)(x_{i+1} - x_i)\right| \le \varepsilon(b-a). \tag{14}$$

A nonstandard proof of the fundamental theorem of calculus can be given as in [22]. Here we prove a change-of-variables theorem. We use the observation that if $\Vdash {}^*u \simeq {}^*v$ for some $u, v \in \mathbb{R}$, then $u = v$.

THEOREM 23 (Change of Variables). *Let $\varphi\colon [A, B] \to [a, b]$ be a strictly increasing, differentiable function with $A < B$ $\varphi(A) = a$ and $\varphi(B) = b$. Let $f\colon [a, b] \to \mathbb{R}$ be a uniformly continuous function. Then*

$$\int_a^b f(x)\,dx = \int_A^B f(\varphi(y)) \cdot \varphi'(y)\,dy.$$

Proof. Let m_φ and m_f be the continuity moduli of φ and f respectively. We work inside the model. For all infinite ω we have

$$(15)\qquad {}^*\left(\int_A^B f(\varphi(y))\cdot\varphi'(y)\,dy\right) \simeq \sum_{k=0}^{\omega-1} {}^*f({}^*\varphi(A_k))\cdot{}^*\varphi'(A_k)\Delta,$$

where $A_k = {}^*A + k\Delta$ and $\Delta = ({}^*B - {}^*A)/\omega$. Let $\varepsilon > 0$ be infinitesimal. Choose an infinite ω such that

$$(16)\qquad \Delta = A_{k+1} - A_k \leq {}^*m_\varphi({}^*m_f(\varepsilon)).$$

Using Corollary 18 we find that for some infinitesimal δ_k

$${}^*\varphi'(A_k) = \frac{{}^*\varphi(A_{k+1}) - {}^*\varphi(A_k)}{\Delta} + \delta_k.$$

Write $a_k = {}^*\varphi(A_k)$. The right hand side of (15) may thus be written

$$\begin{aligned}
&\sum_{k=0}^{\omega-1} {}^*f(a_k)\cdot\left[\frac{a_{k+1}-a_k}{\Delta} + \delta_k\right]\cdot\Delta \\
&\qquad = \sum_{k=0}^{\omega-1} {}^*f(a_k)\cdot(a_{k+1}-a_k) + \sum_{k=0}^{\omega-1} {}^*f(a_k)\delta_k\Delta \\
&\qquad \simeq \sum_{k=0}^{\omega-1} {}^*f(a_k)\cdot(a_{k+1}-a_k).
\end{aligned}$$

The second member follows since

$$\left|\sum_{k=0}^{\omega-1} {}^*f(a_k)\delta_k\Delta\right| \leq {}^*M\Delta\sum_{k=0}^{\omega-1}|\delta_k| \leq {}^*M\Delta\omega(\delta_j + \omega^{-1}) \simeq 0,$$

where ${}^*M \geq |{}^*f(x)|$ for all $x \in [{}^*a, {}^*b]$. Now $a_0 = a,\ a_\omega = b$ and $a_k < a_{k+1}$ since ${}^*\varphi$ is strictly increasing. So we have a partition, which due to (16), has mesh less than ${}^*m_f(\varepsilon)$. Hence by (14),

$$\left|{}^*\left(\int_a^b f(x)\,dx\right) - \sum_{k=0}^{\omega-1} {}^*f(a_k)(a_{k+1}-a_k)\right| \leq \varepsilon(b-a) \simeq 0.$$

This proves the theorem. ⊣

Several variables. Let $\tilde{I}_1, \ldots, \tilde{I}_m \subset \mathbb{R}$ be compact intervals, and put $I = \tilde{I}_1 \times \cdots \times \tilde{I}_m$. Let f and g be uniformly continuous functions $I \to \mathbb{R}$. Then g is the *ith partial derivative* of f in case for every k there is some n such that for all p if $2^{-p} < |x_i - u| \leq 2^{-n}$, then

$$\left|\frac{f(\mathbf{x}) - f(\mathbf{y})}{x_i - u} - g(\mathbf{x})\right| \leq 2^{-k},$$

whenever $\mathbf{x} = (x_1, \ldots, x_m)$ and $\mathbf{y} = (x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_m)$.

THEOREM 24. *The function g is the ith partial derivative of f if and only if the following implication is valid in the model*:

$$2^{-p} < |x_i - u| \simeq 0 \Longrightarrow \frac{{}^*f(\boldsymbol{x}) - {}^*f(\boldsymbol{y})}{x_i - u} \simeq {}^*g(\boldsymbol{x}),$$

whenever $\boldsymbol{x} = (x_1, \dots, x_m)$ *and* $\boldsymbol{y} = (x_1, \dots, x_{i-1}, u, x_{i+1}, \dots, x_m)$.

PROOF. Analogous to Corollary 18. ⊣

In constructive analysis the general mean value theorem is only approximative.

LEMMA 25. *Let $f : I \to \mathbb{R}$ be a differentiable function on the compact interval I. Let $a \neq b$ in I. Then for each $\varepsilon > 0$ there exists $c \in (a, b)$ for which*

$$\left| f'(c) - \frac{f(a) - f(b)}{a - b} \right| < \varepsilon.$$

LEMMA 26. *Let $f : I = \tilde{I}_1 \times \cdots \times \tilde{I}_m \to \mathbb{R}$ be a uniformly continuous function, which has uniformly continuous partial derivatives of first and second order. Then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

PROOF. We work inside the model. Let $(a_1, \dots, a_m) \in I$. Write

$$g(x, y) = {}^*f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_{j-1}, y, a_{j+1}, \dots, a_m),$$

and $\partial g/\partial x$ for $\partial f/\partial x_i$ and so on. Let $a = a_i$ and $b = a_j$. Pick infinitesimals h and k such that $|h|, |k| > 2^{-n}$, for some infinite n, and so that $a + h \in \tilde{I}_i$ and $b + k \in \tilde{I}_j$. Consider the expression

$$\frac{1}{hk}[g(a + h, b + k) - g(a + h, b) - g(a, b + k) + g(a, b)]. \tag{17}$$

The mean-value theorem applied to $\varphi(t) = g(t, b+k) - g(t, b)$ with $\varepsilon = |hk|$ gives some $c \in (a, a + h)$ and some δ with $|\delta| < \varepsilon$, such that (17) equals

$$\begin{aligned}\frac{1}{k}\left[\frac{\partial g}{\partial x}(c, b + k) - \frac{\partial g}{\partial x}(c, b) + \delta\right] &\simeq \frac{1}{k}\left[\frac{\partial g}{\partial x}(c, b + k) - \frac{\partial g}{\partial x}(c, b)\right] \\ &\simeq \frac{\partial^2 g}{\partial y \partial x}(c, d) \simeq \frac{\partial^2 g}{\partial y \partial x}(a, b),\end{aligned}$$

where $d \in (b, b + k)$. The second member follows by applying the mean-value theorem with ε infinitesimal. The last member follows by uniform continuity. Analogously if we apply the mean-value theorem to $\psi(t) = g(a + h, t) - g(a, t)$ we find that (17) equals

$$\frac{1}{h}\left[\frac{\partial g}{\partial y}(a + h, d') - \frac{\partial g}{\partial y}(a, d')\right] \simeq \frac{\partial^2 g}{\partial x \partial y}(c', d') \simeq \frac{\partial^2 g}{\partial x \partial y}(a, b).$$ ⊣

Some further examples of standard theorems that have quite natural nonstandard proofs are the implicit function theorem and Fubini's theorem.

Euler's product formula for sine. Kanovei [12] showed that Euler's original argument for the product formula can be vindicated using nonstandard analysis. A nonstandard proof in the spirit of Euler, but simpler, was however given earlier by Luxemburg [13]. We shall here outline a constructive proof based on Kanovei's account of this simplified proof. Again, only a slightly subtler case distinction is the main point in the constructivisation. The product formulas of Euler are

$$\sin z = z\left(1-\frac{z^2}{\pi^2}\right)\left(1-\frac{z^2}{4\pi^2}\right)\left(1-\frac{z^2}{9\pi^2}\right)\dots \tag{18}$$

and

$$\sinh z = z\left(1+\frac{z^2}{\pi^2}\right)\left(1+\frac{z^2}{4\pi^2}\right)\left(1+\frac{z^2}{9\pi^2}\right)\dots. \tag{19}$$

Using the substitution $z = iw$ these formulas are seen to be equivalent, so it is sufficient to derive (19). By the nonstandard characterisation of limit we need only to show

THEOREM 27. *For every finite $z \in {}^*\mathbb{C}$ and for every infinite $v \in {}^*\mathbb{N}$:*

$$\sinh z \simeq z \prod_{k=1}^{v}\left(1+\frac{z^2}{k^2\pi^2}\right).$$

PROOF. Let z be finite. We have for all infinite natural numbers ω

$$\sinh z = \frac{e^z - e^{-z}}{2} \simeq \frac{1}{2}\left(1+\frac{z}{\omega}\right)^{\omega} - \frac{1}{2}\left(1-\frac{z}{\omega}\right)^{\omega}, \tag{20}$$

using the characterisation of limit. Denote the righthand side of (20) by $\sinh_\omega z$. The crucial step involves the algebraic formula ($m \in {}^*\mathbb{N}$)

$$a^{2m+1} - b^{2m+1} = (a-b)\prod_{k=1}^{m}\left(a^2+b^2-2ab\cos\frac{2k\pi}{2m+1}\right).$$

Letting $m = v,\ \omega = 2v+1,\ a = 1+z/\omega$ and $b = 1-z/\omega$ we obtain from this, using elementary trigonometric identities in each factor,

$$\sinh_\omega z = \frac{z}{\omega}\left(\prod_{k=1}^{v} 4\sin^2\frac{k\pi}{\omega}\right)\prod_{k=1}^{v}\left(1+\frac{z^2}{k^2\pi^2}h_k\right), \tag{21}$$

where $h_k = (k\pi/\omega)^2/\tan^2(k\pi/\omega)$. Now $\sinh_\omega z$ is a polynomial where the coefficient for z is 1, by the binomial theorem. Hence on the righthand side of (21) the coefficient of z is also 1, so

$$\frac{1}{\omega}\prod_{k=1}^{v} 4\sin^2\frac{k\pi}{\omega} = 1.$$

Thus we have established:

$$\sinh_\omega z = z \prod_{k=1}^{\nu} \left(1 + \frac{z^2}{k^2\pi^2} h_k\right) \tag{22}$$

(actually, without any restrictions on ν and z). We use two typical nonstandard lemmas to derive the standard product formula from (22). There will be two cases to consider.

Case 1. We prove first that for infinite ν

$$z \prod_{k=1}^{\nu} \left(1 + \frac{z^2}{k^2\pi^2}\right) \simeq 0 \quad (z \simeq j\pi i \text{ and } j \text{ finite}). \tag{23}$$

In estimating infinite sums and products the following result, known as *Robinson's lemma* [23], is often useful.

Lemma 28. *For every sequence* $a \in {}^*(\mathbb{N} \to \mathbb{C})$, *with* $a_n \simeq 0$ *for all standard* $n \in {}^*\mathbb{N}$, *there is an infinite natural number* η *such that*

$$a_n \simeq 0 \quad (n \leq \eta).$$

Suppose $z = j\pi i + \delta$ where δ infinitesimal, and j finite. Thus $1 + z^2/j^2\pi^2 \simeq 0$, so for each finite $n \geq j$:

$$z \prod_{k=1}^{n} \left(1 + \frac{z^2}{k^2\pi^2}\right) \simeq 0.$$

By Robinson's lemma this holds for some infinite $n \leq \nu$. But for infinite n

$$\prod_{k=n+1}^{\nu} \left(1 + \frac{|z|^2}{k^2\pi^2}\right) \leq \exp\left(\sum_{k=n+1}^{\nu} \frac{|z|^2}{k^2\pi^2}\right) \simeq 1. \tag{24}$$

Thus (23) holds for any finite j and any infinite ν. For Case 2 we need

Lemma 29. *Let* $\nu \in {}^*\mathbb{N}$ *and let* $p, u \in {}^*(\mathbb{N} \to \mathbb{C})$ *be sequences. Suppose that* $\sum_{k=1}^{\nu} |p_k| \simeq 0$ *and that for some standard* $\varepsilon > 0$, $|u_k| > \varepsilon$ *holds for all* $k = 1, \ldots, \nu$. *Then*

$$\frac{\prod_{k=1}^{\nu}(u_k + p_k)}{\prod_{k=1}^{\nu} u_k} \simeq 1.$$

Proof. Let $v_k = p_k/u_k$. By the assumptions

$$\sum_{k=1}^{\nu} |v_k| \leq \frac{1}{\varepsilon} \sum_{k=1}^{\nu} |p_k| \simeq 0.$$

Hence

$$\left|\frac{\prod_{k=1}^{\nu}(u_k+p_k)}{\prod_{k=1}^{\nu}u_k}-1\right| \leq \prod_{k=1}^{\nu}(1+|v_k|)-1$$
$$\leq \exp\left(\sum_{k=1}^{\nu}|v_k|\right)-1$$
$$\simeq 0. \qquad \dashv$$

Case 2. Suppose there is a standard $\delta > 0$ such that $|z - j\pi i| > \delta$ for all $j \in {}^*\mathbb{Z}$. Let $u_k = 1 + z^2/k^2\pi^2$ and $p_k = (h_k - 1)z^2/k^2\pi^2$. Let $N \geq |z| + 1$ be standard, and put $\varepsilon = \min(1/2, \delta^2/N^2\pi^2)$. Thus $|u_k| > \varepsilon$ for all k. We have $0 < h_k < 1$ for $1 \leq k \leq \nu$, and for finite k, we have

$$h_k \simeq \frac{(k\pi/\omega)^2}{\tan^2(k\pi/\omega)} \simeq 1.$$

Hence by Robinson's lemma there is some infinite n

$$\sum_{k=1}^{\nu}|p_k| = \sum_{k=1}^{n}|p_k| + \sum_{k=n+1}^{\nu}|p_k| \simeq \sum_{k=n+1}^{\nu}|p_k| \leq \sum_{k=n+1}^{\nu}\frac{|z|^2}{k^2\pi^2} \simeq 0. \tag{25}$$

By a similar estimate as in (23) it follows that $\prod_{k=1}^{\nu} u_k$ is finite. Hence we can apply Lemma 29

$$z\prod_{k=1}^{\nu}\left(1+\frac{z^2}{k^2\pi^2}\right) = z\left(\prod_{k=1}^{\nu}u_k\right)\left(\prod_{k=1}^{\nu}\frac{u_k+p_k}{u_k}\right) \simeq z\prod_{k=1}^{\nu}u_k.$$

We have thus for each standard $n \in {}^*\mathbb{N}$:

$$(\forall j \in {}^*\mathbb{Z}\ |z - j\pi i| > 2^{-n}) \Longrightarrow \left|\sinh z - \prod_{k=1}^{\nu}u_k\right| < 2^{-n}.$$

By overspill this holds for some infinite $n = \eta$ as well.

Using transfer, we have that for any $z \in {}^*\mathbb{C}$ and any positive hyperreal $\varepsilon < 1/2$: $|z - j\pi i| < \varepsilon$ for some $j \in {}^*\mathbb{Z}$, or $|z - j\pi i| > \varepsilon/2$ for all $j \in {}^*\mathbb{Z}$. Taking $\varepsilon = 2^{-\eta+1}$, we see that Case 1 or Case 2 applies, so either way,

$$\sinh z \simeq \prod_{k=1}^{\nu}u_k.$$

This proves Theorem 27. $\dashv$

§6. Infinitesimal calculus based on hyperrational numbers. Nelson [21] showed that it is possible to develop nonstandard probability theory using only a fragment of his internal set theory. In this development the hyperfinite probability measures are the central objects of study and the standard part map is used very sparingly. Also Harthong [11] has shown how to develop elements of analysis from only hyperfinite concepts. Since one of the main problems with constructivisation of more advanced nonstandard analysis is the lack of a standard part map, Nelson's work should be a good starting point. Classically, $\mathbb{R}$ and its *nonstandard hull* $\mathrm{Fin}({}^*\mathbb{R})/\simeq$ are isomorphic via the standard part map. By consistently working with the latter structure one can often avoid taking standard parts. We shall here go a bit further and use $\mathrm{Fin}({}^*\mathbb{Q})/\simeq$, which is also classically isomorphic to $\mathbb{R}$ (cf. Stroyan and Luxemburg [28]). This has the advantage that in many situations we may use the decidable order on hyperrational numbers. We show that the construction $\mathrm{Fin}({}^*\mathbb{Q})/\simeq$ works as a substitute for $\mathbb{R}$ also in the present sheaf model.

Definition 3. Suppose that (S, d) is a metric space, where S is a subset of ${}^*X$ and $d \in {}^*(X \times X \to \mathbb{Q})$. Then the metric space is *W-complete* if every short Cauchy sequence $a \in {}^*(\mathbb{N} \to X)$, with $a_n \in S$ for all standard n, has a short S-limit in S.

The following theorem gives the essential property of the nonstandard hull.

Theorem 30. *Let (X, d) be a metric space where the metric takes values in $\mathbb{Q}$. Then in the model: $({}^*X, {}^*d)$ is a W-complete metric space, where ${}^*X$ is equipped with the equality $x \simeq_d y \iff d(x, y) \simeq 0$.*

Proof. Working in the model, suppose that $a : {}^*(\mathbb{N} \to X)$ is a short Cauchy sequence. This assumption can be written

$$\forall^{\mathrm{st}} k\ \exists^{\mathrm{st}} \ell\ \forall^{\mathrm{st}} m, n\ [{}^*d(a_{\ell+m}, a_{\ell+n}) < 2^{-k}].$$

Thus by monotone choice we obtain some standard function $c\colon {}^*(\mathbb{N} \to \mathbb{N})$ such that

$$\forall^{\mathrm{st}} m, n, k\ [{}^*d(a_{c(k)+m}, a_{c(k)+n}) < 2^{-k}].$$

Now using overspill it follows that for some infinite v

$$(\forall m, n, k < v)\ [{}^*d(a_{c(k)+m}, a_{c(k)+n}) < 2^{-k}].$$

Let m and k be standard and put $n = v - c(k)$. Then ${}^*d(a_{c(k)+m}, a_v) < 2^{-k}$ for all standard k and m. So $a_v \in {}^*X$ is the limit. ⊣

The collection of finite points in ${}^*X$, denoted $\mathrm{Fin}({}^*X)$, together with equivalence relation $\simeq_d$ is the nonstandard hull of X. It is a W-complete metric space.

We shall use $\mathrm{Fin}({}^*\mathbb{Q})$ as a substitute for the real numbers, where $\mathbb{Q}$ is equipped with the absolute value as the norm. The relation $\simeq$ plays the role of equality between reals. Order relations $\ll$ and $\lesssim$ corresponding the customary order relations on real numbers $<$ and $\leq$, respectively, are defined follows.

DEFINITION 4. For $x, y \in {}^*\mathbb{Q}$

(i) x is *appreciable* if and only if $|x| > 2^{-n}$, for some standard n,
(ii) x is *appreciably smaller than* y, in symbols $x \ll y$, if and only if $y - x$ is positive and appreciable,
(iii) x is *not appreciably greater than* y, $x \lesssim y$, if and only if for all standard n, $x - 2^{-n} < y$.

For the remainder of this section we will work inside the model. The following characterisations are easy consequences of overspill and underspill.

PROPOSITION 31. *Let* $x, y \in {}^*\mathbb{Q}$. *Then*:

(i) $x \ll y$ *if and only if for all positive infinitesimal* $\delta \in {}^*\mathbb{Q}$, $x + \delta < y$,
(ii) $x \lesssim y$ *if and only if* $\neg(y \ll x)$ *if and only if for some positive infinitesimal* $\delta \in {}^*\mathbb{Q}$, $x - \delta < y$.

In contrast to these relations, note that the usual order relation $<$, $\leq$ and the equality $=$ on ${}^*\mathbb{Q}$ are all decidable in the model. However we have the following surprising consequence of overspill.

PROPOSITION 32. *For all* $x, y \in {}^*\mathbb{Q}$: $x \lesssim y$ *or* $y \lesssim x$.

PROOF. Apply overspill to

$$\forall^{\mathrm{st}} m\ [(\forall n < m)\ x - 2^{-n} < y \vee (\forall n < m)\ y - 2^{-n} < x]. \qquad \dashv$$

Note that Proposition 32 does not hold for constructive reals! The following results parallel those for constructive reals. However in some cases they apply also to infinite numbers.

PROPOSITION 33. *Let* $a, b, c, d \in {}^*\mathbb{Q}$.

(i) *If* a *and* b *are finite, then so are* $-a$, $a + b$, ab *and* $\max(a, b)$.
(ii) *If* $a \gg 0$ *or* $a \ll 0$, *then* a^{-1} *is finite.*
(iii) *If* $a \simeq b$, *then* $-a \simeq -b$.
(iv) *If* $a \simeq b$, $c \simeq d$, *then* $a + c \simeq b + d$ *and* $\max(a, c) \simeq \max(b, d)$.
(v) *If* $a \simeq b$ *and* $c \simeq d$ *are finite, then* $ac \simeq bd$.
(vi) *If* $a \simeq b$ *and* $a \gg 0$ *or* $a \ll 0$, *then* $a^{-1} \simeq b^{-1}$.
(vii) *If* $a \ll b$, *then* $-b \ll -a$.
(viii) *If* $a \lesssim b$ *and* $c \lesssim d$, *then* $a + c \lesssim b + d$.
(ix) *If* $a \lesssim b$ *and* $c \ll d$, *then* $a + c \ll b + d$.
(x) *If* $a \gg 0$ *and* $b \gg 0$, *then* $ab \gg 0$.
(xi) *If* $a \gtrsim 0$ *and* $b \gtrsim 0$ *are finite, then* $ab \gtrsim 0$.
(xii) *If* $a \ll b$, *then* $a \ll c$ *or* $c \ll b$.

PROOF. Left to the reader. ⊣

Note that the requirement of finiteness is necessary in (v) and (xi). Consider $0 \simeq -\nu^{-1}$ and $\nu \simeq \nu$ for some infinite $\nu \gg 0$. We have $-\nu^{-1}\nu = -1$. The next proposition is not valid for ordinary constructive reals.

PROPOSITION 34. *For $a, b \in {}^*\mathbb{Q}$, $ab \simeq 0$ implies $a \simeq 0$ or $b \simeq 0$.*

PROOF. This is again a simple overspill argument. Let $ab \simeq 0$. Thus for each standard $n > 0$, $|ab| < n^{-2}$. Hence $|a| < n^{-1}$ or $|b| < n^{-1}$ for each standard $n > 0$. By overspill there is thus an infinite n such that this holds. ⊣

We note the following simple consequence of the nonstandard characterisation of continuity. It gives a version of Brouwer's theorem that all functions on real numbers are continuous.

PROPOSITION 35. *Let $f : {}^*(\mathbb{Q} \to \mathbb{Q})$ take finite rationals to finite rationals, and suppose that for all finite $x, y \in {}^*\mathbb{Q}$ with $x \simeq y$ we have $f(x) \simeq f(y)$. Then f is S-uniformly continuous on every finite interval $[-n, n]$.*

The intermediate-value theorem turns out to hold in its exact form, again unlike what is the case for constructive real numbers.

THEOREM 36. *Let $f : {}^*(\mathbb{Q} \to \mathbb{Q})$ be S-uniformly continuous on $[a, b]$. Suppose that $a \ll c < d \ll b$ and $f(c) \lesssim 0 \lesssim f(d)$. Then $f(x) \simeq 0$, for some $c \leq x \leq d$.*

PROOF. Let ω be the continuity modulus for f so that

$$\forall^{\mathrm{st}} n\, \forall x, y\, [a \leq x, y \leq b \wedge |x - y| < 2^{-\omega(n)} \Longrightarrow |f(x) - f(y)| < 2^{-n}].$$

By overspill there is some infinite $n = \nu$ for which this holds. Choose an (infinite) $\mu \in {}^*\mathbb{N}$ such that $(d - c)/\mu < 2^{-\omega(\nu)}$. Let $\Delta = (d - c)/\mu$ and define an infinite uniform subdivision $t_k = c + k\Delta$ $(k = 0, \ldots, \mu)$ of $[c, d]$. By the assumption there is some infinitesimal $\varepsilon > 0$ such that $f(t_0) < \varepsilon$ and $f(t_\mu) > -\varepsilon$. In case $f(t_0) \geq 0$ or $f(t_\mu) \leq 0$ we are done. Suppose thus that $f(t_0) < 0$ and $f(t_\mu) > 0$. Since $<$ is decidable on ${}^*\mathbb{Q}$ we can search for $k \leq \mu$ such that $f(t_k) \leq 0 < f(t_{k+1})$. But $|t_{k+1} - t_k| = \Delta < 2^{-\omega(\nu)}$, so $|f(t_{k+1}) - f(t_k)| < 2^{-\nu}$ and hence $f(t_k) \simeq 0$. ⊣

Also Rolle's theorem holds in the exact form.

THEOREM 37. *Let $f : {}^*(\mathbb{Q} \to \mathbb{Q})$ be a function W-differentiable on $a \ll b$ such that $f(a) \simeq f(b)$. Then $f'(x) \simeq 0$ for some $a \leq x \leq b$.*

PROOF. Let $\varepsilon > 0$ be an admissible differentiation mesh for f. Choose n such that $\Delta = (b - a)/n$ is infinitesimal and greater than ε. Then $x_k = a + k\Delta$ determines an infinite subdivision $x_0, \ldots, x_n$ of $[a, b]$. Since f is differentiable with derivative f', we have some infinitesimals η_k $(k = 0, \dots, n-1)$ such that

$$f(x_k) - f(x_{k+1}) = f'(x_k)(x_{k+1} - x_k) + \eta_k(x_{k+1} - x_k).$$

Thus we may write

$$\begin{aligned} f(a) - f(b) &= \sum_{k=0}^{n-1}(f(x_k) - f(x_{k+1})) \\ &= \sum_{k=0}^{n-1}[f'(x_k)(x_{k+1} - x_k) + \eta_k(x_{k+1} - x_k)] \\ &= \Delta\sum_{k=0}^{n-1}(f'(x_k) + \eta_k). \end{aligned}$$

Then

$$\sum_{k=0}^{n-1}(f'(x_k) + \xi_k) = 0,$$

where each $\xi_k = \eta_k - (f(a) - f(b))/n\Delta = \eta_k - (f(a) - f(b))/(b-a)$ is infinitesimal since $b - a$ is appreciable. In this sum, clearly not all terms $f'(x_k) + \xi_k$ can be positive (or negative). Hence there is some k where they change sign. Suppose

$$f'(x_k) + \xi_k \leq 0 \leq f'(x_{k+1}) + \xi_{k+1}. \tag{26}$$

Since f' is S-continuous, the difference between the left and right hand side must be infinitesimal. Hence $f'(x_k) \simeq f'(x_k) + \xi_k \simeq 0$. The argument is analogous when $\leq$ is replaced by $\geq$ in (26). ⊣

Next we give a version of the Heine-Borel theorem. Let $y \in [\![a, b]\!]$ denote $a \lesssim y \lesssim b$, and let $y \in (\!(a, b)\!)$ denote $a \ll y \ll b$.

THEOREM 38. *Let a, b: $^*(\mathbb{N} \to \mathbb{Q})$ be sequences of hyper-rationals. If for each $y \in [\![c, d]\!]$ there is some standard n such that $y \in (\!(a_n, b_n)\!)$, then for some standard ℓ, every $y \in [\![c, d]\!]$ lies in some $(\!(a_n, b_n)\!)$, $n < \ell$.*

PROOF. Using Proposition 31, the assumption may be written

$$\forall^{\text{inf}} p\, [c - 2^{-p} < y < d + 2^{-p} \Longrightarrow \exists^{\text{st}} n\, \exists^{\text{st}} k\, (a_n + 2^{-k} < y < b_n - 2^{-k})].$$

Thus for all infinite m,

$$\begin{aligned} (\forall p \geq m)\, [c - 2^{-p} &< y < d + 2^{-p} \\ &\Longrightarrow (\exists n, k < m)(a_n + 2^{-k} < y < b_n - 2^{-k})]. \end{aligned} \tag{27}$$

By underspill (27) holds for some standard m. Thus (27) holds for all infinite p and this proves the theorem. ⊣

Note that the theorem is more general than the common formulation since the interval covered need not be finite.

We will give a version of the Cauchy–Peano existence theorem which avoids the standard-part map by using W-differentiability instead of S-differentiability. For a classical nonstandard proof see [10]. Let $f : {}^*([0,1] \times \mathbb{Q} \to \mathbb{Q})$ be S-uniformly continuous and bounded by a standard M. Let $x_0 \in {}^*\mathbb{Q}$. For any infinitesimal $\varepsilon > 0$ define for $n \in {}^*N$, $n \leq 1/\varepsilon$,

$$\begin{cases} t_n = n\varepsilon, \\ x_n = x_{n-1} + \varepsilon f(t_{n-1}, x_{n-1}) \quad (n > 0). \end{cases}$$

From this a step function $\psi_\varepsilon : {}^*(\mathbb{Q} \to \mathbb{Q})$ is defined by

$$\psi_\varepsilon(t) = \begin{cases} x_n & \text{if } t_n \leq t < t_{n+1}, \\ x_0 & \text{otherwise.} \end{cases}$$

Notice that here we use the decidability of $\leq$ on ${}^*\mathbb{Q}$. Then ψ_ε *is S-uniformly continuous on* ${}^*[0,1]$: Let t, $t' \in {}^*[0,1]$ be infinitely close. Thus for some m, n we have $t_m \leq t < t_{m+1}$ and $t_n \leq t' < t_{n+1}$. We may suppose $m \leq n$. Then

$$\begin{aligned} |\psi_\varepsilon(t) - \psi_\varepsilon(t')| &= |x_m - x_n| \\ &\leq \sum_{i=0}^{n-m-1} |x_{m+i} - x_{m+i+1}| \\ &\leq \sum_{i=0}^{n-m-1} \varepsilon |f(t_{m+i}, x_{m+i})| \\ &\leq \varepsilon(n-m)M \leq (\varepsilon + |t - t'|)M \simeq 0. \end{aligned}$$

Moreover $\psi_\varepsilon(t)$ *is W-differentiable with derivative* $f(t, \psi_\varepsilon(t))$ *on* ${}^*[0,1]$*:* Let $\xi > 0$ be infinitesimal such that

$$\frac{\varepsilon}{\xi} \simeq 0.$$

Let t, $t' \in {}^*[0,1]$ be such that $\xi < |t' - t| \simeq 0$. Then $t_m \leq t < t_{m+1}$ and $t_n \leq t' < t_{n+1}$, for some m, n. It may be assumed that $m \leq n$.

$$\begin{aligned} \frac{\psi_\varepsilon(t') - \psi_\varepsilon(t)}{t' - t} &= \frac{x_n - x_m}{t' - t} \\ &= \frac{1}{t' - t} \sum_{i=m}^{n-1} (x_{i+1} - x_i) \\ &= \frac{\varepsilon}{t' - t} \sum_{i=m}^{n-1} f(t_i, x_i). \end{aligned}$$

By continuity, $f(t_i, \psi_\varepsilon(t_i)) = f(t, \psi_\varepsilon(t)) + \eta_i$, where $\eta_i \simeq 0$. Hence

$$\begin{aligned}\frac{\psi_\varepsilon(t') - \psi_\varepsilon(t)}{t' - t} &= \frac{\varepsilon}{t' - t} \sum_{i=m}^{n-1} (f(t, \psi_\varepsilon(t)) + \eta_i) \\ &= \frac{\varepsilon(n-m)}{t'-t} f(t, \psi_\varepsilon(t)) + \frac{\varepsilon}{t'-t} \sum_{i=m}^{n-1} \eta_i.\end{aligned}$$

We have by the assumption on ξ,

$$\left| \frac{\varepsilon(n-m)}{t'-t} - 1 \right| = \frac{|\varepsilon n - t' + t - \varepsilon m|}{|t'-t|} \leq \frac{2\varepsilon}{|t'-t|} \leq \frac{2\varepsilon}{\xi} \simeq 0.$$

Since

$$\left| \frac{\varepsilon}{t'-t} \sum_{i=m}^{n-1} \eta_i \right| \leq \left| \frac{\varepsilon(n-m)}{t'-t} \right| \max_{m \leq i < n} |\eta_i| \simeq 0,$$

and f is bounded by a standard number, we obtain

$$\frac{\psi_\varepsilon(t') - \psi_\varepsilon(t)}{t'-t} \simeq f(t, \psi_\varepsilon(t)).$$

We summarise this as a

THEOREM 39. *Let $f : {}^*([0,1] \times \mathbb{Q} \to \mathbb{Q})$ be an S-uniformly continuous function which is bounded by some standard M. Let $x_0 \in {}^*\mathbb{Q}$. Then there is a W-differentiable function $\psi : {}^*([0,1] \to \mathbb{Q})$ such that*

$$\begin{cases} \psi(0) = x_0, \\ \psi'(t) \simeq f(t, \psi(t)) \quad (t \in {}^*[0,1]). \end{cases}$$

Moreover if x_0 is finite, then $\psi(t)$ is finite for all $t \in {}^*[0,1]$.

A theory of integration. We adapt Nelson's theory of L^1 random variables to hyperfinite step functions, cf. [21]. A *step function* $f : {}^*(\mathbb{Q} \to \mathbb{Q})$ is given by a hyperfinite sum of the form

$$f = \sum_{k=1}^{\nu} c_k \chi_{[a_k, b_k)},$$

where $\chi_{[a,b)}$ is the characteristic function of $[a,b)$. Since inequalities are decidable on ${}^*\mathbb{Q}$ we may assume that all the $[a_k, b_k)$ are disjoint and nonvoid. The *integral* of f is then by definition

$$I(f) = \sum_{k=1}^{\nu} c_k (b_k - a_k).$$

Defining $||f|| = I(|f|)$ we obtain the L^1-*norm* on step functions.

PROPOSITION 40. *Let* $f : {}^*([a,b) \to \mathbb{Q})$ *be an S-uniformly continuous function. Then there is a step function* $g : {}^*([a,b) \to \mathbb{Q})$ *such that* $f(x) \simeq g(x)$ *for all* $x \in [a,b)$.

PROOF. Let v be infinite and let $t_k = k(b-a)/v + a$. Define

$$g = \sum_{k=0}^{v-1} f(t_k)\chi_{[t_k,t_{k+1})}.$$

For each $t \in [a,b)$ there is some k such that $t_k \le t < t_{k+1}$. Since $t_{k+1} - t_k$ is infinitesimal, $g(t) = f(t_k) \simeq f(t)$. ⊣

Note that if f is a step function then the *truncation*

$$f^{(t)} = f\chi_{\{x:|f(x)|\le t\}}$$

is also a step function.

DEFINITION 5. A step function $f : {}^*([0,1) \to \mathbb{Q})$ is *pre-L^1* if for all infinite $t > 0$, $\|f - f^{(t)}\| \simeq 0$.

Note that since $\|f - f^{(s)}\| \le \|f - f^{(t)}\|$ for $s \ge t$, f is pre-L^1 if and only if $\lim_n \|f - f^{(n)}\| = 0$. Any step function bounded by a standard number is pre-L^1.

PROPOSITION 41. *Let* $f, g : {}^*([0,1) \to \mathbb{Q})$ *be step functions.*

(i) f *is pre-*$L^1 \implies \|f\|$ *is finite.*
(ii) f *is pre-*$L^1 \iff |f|$ *is pre-*L^1.
(iii) $|f| \le g$ *and* g *is pre-*$L^1 \implies f$ *is pre-*L^1.
(vi) $\alpha \in {}^*\mathbb{Q}$ *finite and* f *is pre-*$L^1 \implies \alpha f$ *is pre-*L^1.
(v) f, g *are pre-*$L^1 \implies f + g$ *is pre-*L^1.

PROOF. To prove (i) note that for any m,

$$\|f\| \le \|f^{(m)}\| + \|f - f^{(m)}\| \le m + \|f - f^{(m)}\|.$$

Thus $\|f\|$ is less than any infinite number, and is therefore finite by underspill. The proofs of (ii)–(v) are left to the reader. ⊣

THEOREM 42. *Let* $f \in {}^*([0,1) \to \mathbb{Q})$ *be a pre-*L^1 *step function with* $\|f\| \gg 0$. *Then for some* $0 \le x < 1$, $|f(x)|$ *is finite and* $|f(x)| \gg 0$.

PROOF. Let k be standard such that $\|f\| > 2^{-k}$. Then for any m

$$2^{-k} < \|f\| \le \|f - f^{(m)}\| + \|f^{(m)}\|.$$

Since f is pre-L^1, we have $\|f - f^{(m)}\| \simeq 0$ for all infinite m. Thus $\|f^{(m)}\| > 2^{-k-1}$ for all infinite m. By underspill there is some standard m such that $\|f^{(m)}\| > 2^{-k-1}$. There must be some $x \in [0,1)$ such that $|f^{(m)}(x)| > 2^{-k-1}$. Consequently, $m \ge |f(x)| > 2^{-k-1}$. ⊣

Definition 6. A subset M of the hyperrational numbers is called *measurable* if it is a hyperfinite union of the form $M = \bigcup_{1\le k\le \nu}[a_k, b_k)$, where the intervals $[a_k, b_k)$ are disjoint and nonvoid. The *measure of M*, denoted $\mu(M)$, is the sum $\sum_{k=1}^{\nu}(b_k - a_k)$.

Note that the measurable subsets of any interval $[a, b)$ constitute a boolean algebra closed under hyperfinite unions and intersections.

Definition 7. A property $P(x)$ of hyperrational numbers is said to hold *almost everywhere on* $[a, b)$ if for every $\varepsilon \gg 0$ there is a measurable set M with measure less than ε such that $P(x)$ for all $x \in [a, b) \setminus M$.

The proofs of Theorems 43, 44 and 45 below have been obtained by straightforward modifications of the corresponding arguments in Nelson [21].

Theorem 43. *Let $f : {}^*([0, 1) \to \mathbb{Q})$ be a step function. Then the following are equivalent*:

(i) $f(x) \simeq 0$ *almost everywhere on* $[0, 1)$,

(ii) *for every* $\varepsilon \gg 0$,

$$\mu(\{\, x \in [0, 1) : |f(x)| \ge \varepsilon \,\}) \simeq 0,$$

(iii) *there exists an infinitesimal* $\varepsilon > 0$ *such that*

$$\mu(\{\, x \in [0, 1) : |f(x)| > \varepsilon \,\}) \simeq 0.$$

Proof.

(i) $\Longrightarrow$ (ii): Fix $\varepsilon \gg 0$ and let $\delta \gg 0$. Hence there is a measurable $M \subseteq [0, 1)$ such that $\mu(M) \le \delta$ and $f(x) \simeq 0$ for $x \in [0, 1) \setminus M$. Since M is decidable, $|f(x)| \ge \varepsilon$ implies $x \in M$. Thus, since $\delta \gg 0$ was arbitrary,

$$\mu(\{\, x \in [0, 1) : |f(x)| \ge \varepsilon \,\}) \simeq 0.$$

(ii) $\Longrightarrow$ (iii): For each standard n,

$$\mu(\{\, x : |f(x)| \ge 2^{-n} \,\}) < 2^{-n}.$$

Overspill gives an infinite n such that

$$\mu(\{\, x : |f(x)| \ge 2^{-n} \,\}) < 2^{-n} \simeq 0.$$

(iii) $\Longrightarrow$ (i): Let ε be a positive infinitesimal such that

$$\mu(\{\, x : |f(x)| > \varepsilon \,\}) \simeq 0.$$

Then $|f(x)| \le \varepsilon$ almost everywhere. ⊣

Theorem 44. *Let $f : {}^*([0, 1) \to \mathbb{Q})$ be a step function. Then f is pre-L^1 if and only if $\|f\|$ is finite and for all measurable M with infinitesimal measure,* $I(|f|\chi_M) \simeq 0$.

PROOF.
($\Longrightarrow$) Let f be pre-L^1 and $\mu(M) \simeq 0$. By overspill we find an infinite $a > 0$ such that $a\mu(M) \simeq 0$. Hence

$$\begin{aligned} I(|f|\chi_M) &\leq I(|f^{(a)}|\chi_M) + I(|f - f^{(a)}|\chi_M) \\ &\leq a\mu(M) + I(|f - f^{(a)}|) \simeq 0. \end{aligned}$$

($\Longleftarrow$) Let $a > 0$ be infinite and put

$$M = \{ x \in [0,1) : |f(x)| > a \}.$$

We have $\mu(M) \leq I(|f|)/a \simeq 0$. By assumption

$$0 \simeq I(|f|\chi_M) = \|f - f^{(a)}\|. \qquad \dashv$$

THEOREM 45. *Let f : *$([0,1) \to \mathbb{Q})$ be a pre-L^1 step function. If $f(x) \simeq 0$ almost everywhere on* $[0,1)$, *then* $I(f) \simeq 0$.

PROOF. According to Theorem 43 there is a positive infinitesimal α such that $\mu\{ x : |f(x)| > \alpha \} \simeq 0$. But

$$|f(x)| \leq |f(x)|\chi_{\{ x : |f(x)| > \alpha \}} + \alpha,$$

so by Theorem 44, $I(|f|) \simeq 0$. Hence $I(f) \simeq 0$. $\dashv$

THEOREM 46. *The collection of pre-L^1 step functions on* $[0,1)$ *is W-complete with respect to the L^1-norm.*

PROOF. Let (f_n) be a short Cauchy sequence with respect to the norm $\|\cdot\|$ such that f_n is pre-L^1 for each standard n. Analogously to Theorem 30, we find an infinite v and a standard modulus c such that for all standard k and m:

$$\|f_{c(k)+m} - f_v\| \leq 2^{-k}.$$

We need to check that the short S-limit f_v is pre-L^1. Clearly $\|f_v\|$ is finite. We have $|f_v| \leq |f_n| + |f_v - f_n|$ so in particular we have for a measurable set M $|f_v\chi_M| \leq |f_n\chi_M| + |f_v - f_n|$. Thus

$$I(|f_v\chi_M|) \leq I(|f_n\chi_M|) + I(|f_v - f_n|).$$

If $\mu(M) \simeq 0$ then the first term on the right is infinitesimal, for standard n, according to Theorem 44. By chosing n sufficiently large, the second term can be made arbitrarily small. Hence $I(|f_v\chi_M|) \simeq 0$. Thus f_v is pre-L^1. $\dashv$

A step function f : *$(\mathbb{Q} \to \mathbb{Q})$ is called *simple* if it has the form

$$\sum_{k=1}^{n} a_k \chi_{I_k},$$

where n is standard, $a_1, \ldots, a_n$ are finite and $I_k = \bigcup_{i=1}^{\nu_k} [b_{k,i}, c_{k,i})$ $(k = 1, \ldots, n)$. Any simple function is pre-L^1 and the simple functions are dense in the collection of pre-L^1 functions:

THEOREM 47. *Let $\varepsilon \gg 0$ and let $f : {}^*([0,1) \to \mathbb{Q})$ be a pre-L^1 step function. Then there is a simple step function $g : {}^*([0,1) \to \mathbb{Q})$ such that $\|f - g\| < \varepsilon$.*

PROOF. Since f is pre-L^1 there is a standard $m > 0$ such that $\|f - f^{(m)}\| < \varepsilon/2$. The function $h = f^{(m)}$ may be written as

$$\sum_{i=1}^{\nu} a_i \chi_{[b_i, c_i)},$$

where $|a_i| < m + 1$ and all $[b_i, c_i)$ are disjoint. Construct a partition

$$-m - 1 = t_1 < t_2 < \cdots < t_n = m + 1$$

where n standard and $t_{k+1} - t_k < \varepsilon/2$. Form the sets

$$I_k = \bigcup \{\, [b_i, c_i) : t_k \leq a_i < t_{k+1} \,\}, \quad k = 1, \ldots, n,$$

and let g be the simple function

$$\sum_{k=1}^{n} t_k \chi_{I_k}.$$

Then $\|h - g\| < \varepsilon/2$, so $\|f - g\| < \varepsilon$. ⊣

We now impose a stricter condition on step functions to obtain the true L^1-functions.

DEFINITION 8. A step function $f : {}^*([0,1) \to \mathbb{Q})$ is L^1 if f is a short limit of a standard sequence (g_n) of step functions in ${}^*([0,1) \to \mathbb{Q})$ with respect to the L^1-norm.

We may relate this definition to the Bishop–Cheng theory of integration. To them an L^1-function is a Cauchy sequence of some basic functions. (In the present setting, we consider the set of polygonal functions with rational vertices as basic functions.) Such a sequence can then be characterised in the model using Corollary 20. Thus for any sequence (f_n) of basic functions representing a Bishop–Cheng L^1-function we see that, in the model, f_ν represents an L^1-function in the above sense for any infinite ν. However we cannot be sure to extract a real world Bishop–Cheng L^1-function from a function satisfying our definition.

THEOREM 48. *Any L^1 step function $f : {}^*([0,1) \to \mathbb{Q})$ is pre-L^1.*

PROOF. Let g be a standard step function on $[0, 1)$ such that $\|f - g\| < 1$. Then $\|f\| \leq \|f - g\| + \|g\| < 1 + \|g\|$, so $\|f\|$ is finite. By Theorem

44, it suffices to show that $I(|f|\chi_M) \simeq 0$ for each M with measure 0. Let $\mu(M) \simeq 0$, and consider a standard $\varepsilon > 0$. Trivially

$$|f|\chi_M \leq |f-g|\chi_M + |g|\chi_M,$$

for any g. Let h be a standard step function such that $\|f-h\| < \varepsilon$, and suppose it is bounded by the standard number K.

$$\begin{aligned} I(|f|\chi_M) &\leq I(|f-h|\chi_M) + I(|h|\chi_M) \\ &\leq \|f-h\| + I(K\chi_M) \\ &\leq \varepsilon + K\mu(M) \simeq \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, $I(|f|\chi_M) \simeq 0$. ⊣

THEOREM 49. *The collection of L^1-functions is W-complete with respect to the L^1-norm.*

PROOF. A simple application of overspill. ⊣

It is easy to show that the L^1-functions are closed under the vector operations and absolute value. As an example of a pre-L^1 step function which is not L^1 consider the hyperfinite Rademacher-type function

$$f = \sum_{k=0}^{\nu-1} \chi_{\left[\frac{2k}{2\nu}, \frac{2k+1}{2\nu}\right)},$$

where ν is an infinite natural number.

REMARK 3. In the approach to nonstandard analysis presented in this section, we have a difficulty which is similar to the interpretation of the theorems of a classical system. A theorem does not mean in the real world what it appears to mean in the theory. However, we are in a better position since we have a large class of statements that can be interpreted directly via the transfer principles (Theorem 1 and Corollary 4). Note that

$$\Vdash (\exists x \in {}^*I)\, {}^*f(x) \simeq 0 \tag{28}$$

is equivalent to

$$\Vdash \forall^{\mathrm{st}}\varepsilon > 0\, \exists x \in {}^*I\ |f(x)| < \varepsilon,$$

by the overspill principle. So, in the real world, this means only that for every $\varepsilon > 0$ there exists some $x \in I$ so that $|f(x)| < \varepsilon$, and there is no guarantee that the approximations converge. The direct interpretation of (28) through sheaf semantics does not appear to give any additional information.

Appendix. Classical aspects of the model. Working within classical set theory, Moerdijk [17] showed that the full transfer principle is valid inside the model:

$$\Vdash \forall^{\mathrm{st}} \bar{x}\ [\Phi(\bar{x}) \leftrightarrow \Phi^{\mathrm{st}}(\bar{x})],$$

where Φ is internal. But he also showed (constructively) that the model is nonclassical, in that

$$\Vdash \neg(\forall n \in {}^*\mathbb{N})\ [\mathsf{St}(n) \vee \neg\mathsf{St}(n)]. \tag{29}$$

Note that the formula within brackets is not internal, whereas for internal formulas $\Phi(\bar{x})$, we have using classical set theory

$$\Vdash \forall \bar{x}\ [\Phi(\bar{x}) \vee \neg\Phi(\bar{x})].$$

This follows trivially from transfer.

Below we investigate the forcing relation $\mathcal{F} \Vdash \Phi$, when $\mathcal{F}$ is an ultrafilter; we employ classical set theory.

DEFINITION 9. A filter $\mathcal{F} = (I, U, F)$ is called an *ultrafilter* if

(i) $\mathcal{F}$ is proper, i.e., each F_i is nonvoid,
(ii) for every subset S of U there exists some i for which $F_i \subseteq S$ or $F_i \subseteq U \setminus S$.

THEOREM 50.

(i) *Any morphism from a proper filter to an ultrafilter is a covering map.*
(ii) *Any cover of an ultrafilter contains a single map covering the ultrafilter.*

PROOF.

(i): Let $\mathcal{F}$ be a proper filter, $\mathcal{G} = (J, V, G)$ be an ultrafilter and let $\varphi\colon \mathcal{F} \to \mathcal{G}$ be a morphism. We show that for each i there exists k such that $\varphi[F_i] \supseteq G_k$. Let i be arbitrary. Since $\mathcal{G}$ is an ultrafilter, either $\varphi[F_i] \supseteq G_k$ or $V \setminus \varphi[F_i] \supseteq G_k$, for some k. Suppose the latter disjunct holds. Because φ is continuous, there is some i_1 such that $\varphi[F_{i_1}] \subseteq G_k$. Take $F_{i_2} \subseteq F_i \cap F_{i_1}$. Thus $\varphi[F_{i_2}] \subseteq G_k \subseteq V \setminus \varphi[F_i]$. This is impossible since F_{i_2} is nonvoid. Hence $\varphi[F_i] \supseteq G_k$.

(ii): We prove the contrapositive statement. Let $\varphi_k\colon \mathcal{F}^{(k)} \to \mathcal{G}$, $k = 1, \dots, n$, be morphisms and let $\mathcal{G} = (J, V, G)$ be an ultrafilter. Suppose that no map φ_k is a cover of $\mathcal{G}$ by itself. Hence for each $k = 1, \dots, n$, there exists i_k such that for all j, $\varphi_k[F^k_{i_k}] \not\supseteq G_j$. Since $\mathcal{G}$ is an ultrafilter, we have for each j, $V \setminus \varphi_k[F^k_{i_k}] \supseteq G_j$. This is the same as saying that $\langle \varphi_k : \mathcal{F}^{(k)} \to \mathcal{G}\rangle_{k=1}^n$ is not a cover. ⊣

From this theorem follows that the forcing clauses may be simplified in the case of ultrafilters.

COROLLARY 51. *Let $\mathcal{F}$ be an ultrafilter, and let Φ, Ψ be external formulas.*

(i) $\mathcal{F} \Vdash \mathsf{St}(\gamma)$ *if and only if γ is constant on some F_i.*

(ii) $\mathcal{F} \Vdash \Phi(\bar{\alpha}) \vee \Psi(\bar{\alpha})$ *if and only if* $\mathcal{F} \Vdash \Phi(\bar{\alpha})$ *or* $\mathcal{F} \Vdash \Psi(\bar{\alpha})$.
(iii) $\mathcal{F} \Vdash \Phi(\bar{\alpha}) \rightarrow \Psi(\bar{\alpha})$ *if and only if* $\mathcal{F} \Vdash \Phi(\bar{\alpha})$ *implies* $\mathcal{F} \Vdash \Psi(\bar{\alpha})$.
(iv) $\mathcal{F} \Vdash (\exists^{\mathrm{st}} y \in {}^*S)\, \Phi(\bar{\alpha}, y)$ *if and only if for some* $s \in S$, $\mathcal{F} \Vdash \Phi(\bar{\alpha}, {}^*s)$.

Proof. Parts (i), (ii) and (iv) follows from Theorem 50.(i) and the local character of the forcing relation. Part (iii) follows from Theorem 50.(ii) and by noting that if a filter $\mathcal{G}$ is void, then trivially $\mathcal{G} \Vdash \Theta$ for any formula Θ. $\dashv$

However, we do not have, in general, that for ultrafilters $\mathcal{F}$:

$$\mathcal{F} \Vdash (\forall y \in {}^*S)\, \Phi(\bar{\alpha}, y) \quad \text{if and only if for all } \gamma \in {}^*S(\mathcal{F}),\ \mathcal{F} \Vdash \Phi(\bar{\alpha}, \gamma),$$

since this would contradict (29). This means that $\mathcal{F} \Vdash$ is different from truth in ultrapowers based on $\mathcal{F}$.

Acknowledgement. I am grateful to Per Martin-Löf for introducing me to the problem of constructivising nonstandard analysis, and for numerous stimulating conversations on this topic.

REFERENCES

[1] E. Bishop, ***Foundations of constructive analysis***, McGraw-Hill, 1967.

[2] E. Bishop and D. S. Bridges, ***Constructive analysis***, Springer-Verlag, Berlin, 1985.

[3] A. Blass, *Two closed categories of filters*, ***Fundamenta Mathematicae***, vol. 94 (1977), pp. 129–143.

[4] D. S. Bridges and F. Richman, ***Varieties of constructive mathematics***, Cambridge University Press, Cambridge, 1987.

[5] P. Cartier and Y. Perrin, *Integration over finite sets*, in [9].

[6] R. Chuaqui and P. Suppes, *Free-variable axiomatic foundations of infinitesimal analysis: A fragment with finitary consistency proof*, this Journal, vol. 60 (1995), pp. 122–159.

[7] T. Coquand, *A formal space of ultra filters*, *Annals of Pure and Applied Logic*, to appear.

[8] P. Delfini and C. Lobry, *The vibrating string*, in [9].

[9] F. Diener and M. Diener (editors), ***Nonstandard analysis in practice***, Springer-Verlag, Berlin, 1995.

[10] F. Diener and G. Reeb, ***Analyse non standard***, Hermann, Paris, 1989.

[11] J. Harthong, *Éléments pour une théorie du continu*, ***Astérisque***, vol. 109–110 (1983), pp. 235–244.

[12] V. G. Kanovei, *The correctness of Euler's method for the factorization of the sine function into an infinite product*, ***Russian Mathematical Surveys***, vol. 43 (1988), pp. 65–94.

[13] W. A. J. Luxemburg, *What is nonstandard analysis?*, ***American Mathematical Monthly***, vol. 80 (1973), pp. 38–67.

[14] S. Mac Lane and I. Moerdijk, ***Sheaves in geometry and logic***, Springer-Verlag, Berlin, 1992.

[15] P. Martin-Löf, ***Intuitionistic type theory***, Bibliopolis, Naples, 1984.

[16] ———, *Mathematics of infinity*, ***COLOG-88 Computer logic*** (P. Martin-Löf and G. E. Mints, editors), Lecture Notes in Computer Science, vol. 417, Springer-Verlag, Berlin, 1990, pp. 146–197.

[17] I. Moerdijk, *A model for intuitionistic nonstandard arithmetic*, ***Annals of Pure and Applied Logic***, vol. 73 (1995), pp. 37–51.

[18] I. MOERDIJK and E. PALMGREN, *Minimal models of Heyting arithmetic*, this JOURNAL, vol. 62 (1997), pp. 1448–1460.

[19] I. MOERDIJK and G. E. REYES, ***Smooth infinitesimal analysis***, Springer-Verlag, 1991.

[20] J. MYCIELSKI, *Analysis without actual infinity*, this JOURNAL, vol. 46 (1981), pp. 625–633.

[21] E. NELSON, ***Radically elementary probability theory***, Princeton University Press, 1987.

[22] E. PALMGREN, *A constructive approach to nonstandard analysis*, ***Annals of Pure and Applied Logic***, vol. 73 (1995), pp. 297–325.

[23] ———, *Constructive nonstandard analysis*, ***Méthodes et analyse non standard*** (A. Pétry, editor), Cahiers du Centre de Logique, vol. 9, Academia, Louvain-la-Neuve, 1996, pp. 369–397.

[24] ———, *Sheaf-theoretic foundation for nonstandard analysis*, ***Annals of Pure and Applied Logic***, vol. 85 (1997), pp. 69–86.

[25] F. RICHMAN, *Polynomials and linear transformations*, ***Linear Algebra and its Applications***, vol. 131 (1990), pp. 131–137.

[26] A. ROBINSON, ***Non-standard analysis***, Princeton University Press, 1996, reprint of the 1973 edition with a foreword by W. A. J. Luxemburg.

[27] C. SCHMIEDEN and D. LAUGWITZ, *Eine Erweiterung der Infinitesimalrechnung*, ***Mathematisches Zeitschrift***, vol. 69 (1958), pp. 1–39.

[28] K. D. STROYAN and W. A. J. LUXEMBURG, ***Introduction to the theory of infinitesimals***, Academic Press, New York, 1976.

[29] F. WATTENBERG, *Nonstandard analysis and constructivism?*, ***Studia Logica***, vol. 47 (1988), pp. 303–309.

DEPARTMENT OF MATHEMATICS
UPPSALA UNIVERSITY
P.O. BOX 480
S-751 06 UPPSALA, SWEDEN
E-mail: palmgren@math.uu.se