Cahiers du Centre de logique
Volume 9

Constructive Nonstandard Analysis

by

E. Palmgren
Uppsala University

1. Introduction

Nonstandard analysis is generally perceived to be highly nonconstructive. But, in fact, notions like infinitesimals and infinite numbers are both meaningful and useful in constructive analysis, even in the strict form of Bishop [3]. A few years before the advent of Robinson's nonstandard analysis, Schmieden and Laugwitz [17] introduced a more constructive approach to infinitesimal calculus. Their approach relied on classical logic, however. It was further developed by D. Laugwitz in a series of papers and books. Martin-Löf [11] suggested a strictly constructive foundation for nonstandard analysis, which was derived from his analysis of Brouwer's choice sequences. The basic idea was essentially the same as Schmieden and Laugwitz', namely to use certain constructive reduced powers instead of nonconstructive ultra powers in building, e.g., the hyperreals. In [14] we introduced a formal system intended for developing nonstandard analysis along these lines. The system was an intuitionistic higher type arithmetic, with a predicate for "standard" analogous to the internal set theory of Nel-

Supported by a grant from the Swedish Research Council for Engineering Sciences (TFR).

son [12]. In this system, transfer principles strong enough to capture most conditional (in)equalities were obtained (see Section 3.1 below). Consequently it preserves many algebraic properties and estimates of classical nonstandard analysis. A nonstandard proof of the fundamental theorem of calculus was given. We also demonstrated some nonstandard characterisations of standard notions, in particular the limit concept. However, one important device of classical nonstandard analysis seems irrevocably lost, namely the standard part map (see Section 4.2 below).

The present paper develops the matter of [14], and it may be read independently of the earlier publication. In Section 2 we remark on the problem of formalising Bishop's constructive analysis. The nonstandard formal system and some of its properties are reviewed in Section 3. In Section 4 we prove various results in nonstandard analysis: some nonstandard characterisations of standard notions, a nonstandard result generalising Toeplitz' theorem on regular summation. S-continuity and S-differentiability are also discussed. Section 5 contains a nonstandard existence proof for solutions to ordinary differential equations, which uses Euler–Cauchy's method.

2. Formalising analysis in $\mathbf{HA}^\omega$ and related theories

It has been claimed by Goodman and Myhill [7], without much elaboration, that most of Bishop's constructive analysis can be formalised within an intuitionistic theory **S** of arithmetic in higher types. This theory is, more precisely, $\mathbf{HA}^\omega$ extended with two choice schemata (AC and RDC, see below). It is closely related to a theory suggested by Bishop himself [2], and has an easy translation into Martin-Löf's type theory. Our reason for not considering the more expressive type theory, at this stage, is its syntactical complexity; the simpler theory is more amenable to nonstandard extensions. In this section, we show how to cope with the limitations of **S**. Readers not interested in the more subtle points of formalisation are recommended to skip the discussion about representation of functions.

The mathematical objects of **S** are elements of the types generated thus: 0 is the basic type of natural numbers; if σ, τ are types, then the function space $\sigma \to \tau$ (also written $\sigma\tau$) and the cartesian product $\sigma \times \tau$ are types. The elements of these types are induced by the usual combinators (those of Gödel's T and pairing and projections). The combinators make the following constructions possible

- an operation f of type $\sigma \to \tau$ can be defined by abstraction on a variable x of type σ in an expression a of type τ: $f(x) = a(x)$;
- definition by *primitive recursion* (in higher types)

$$\begin{aligned} f(0, \vec{x}) &= g(\vec{x}) \\ f(n+1, \vec{x}) &= h(n, \vec{x}, f(n, \vec{x})), \end{aligned}$$

  where g and h are given operations.

Variables ranging over type σ are written x^σ. Note that we use the word 'operation' for a function which is not required to respect any equivalence relation except the (nonextensional) equality of the system (see discussion in Bishop [3]). Formally there are no sets in **S**. But there are classes, in the sense of set theory; the class $\{x : A(x)\}$ is defined by the extent of the formula $A(x)$. What class a particular formula defines depends on the variable displayed. We often write $x \in A$ for $A(x)$. In the sequel we shall call these classes *sets*, keeping in mind that we cannot in general quantify over the collection of all sets. In constructive mathematics each set comes with an equivalence relation. A *set of σ-objects* is represented by two formulas: $A(x)$ and $x =_A y$, where $=_A$ is an equivalence relation on $\{x : A\}$ and x and y varies over the type σ.

Let $(A, =_A)$ and $(B, =_B)$ be sets of σ- and τ-objects, respectively. A function can be given as a functional relation R_f between the sets, i.e.

$$(\forall x^\sigma \in A)\,(\exists y^\tau \in B)\,R_f(x, y),$$

$$(\forall x^\sigma, u^\sigma \in A)(\forall y^\tau, v^\tau \in B)[x =_A u \wedge R_f(x, y) \wedge R_f(u, v) \Rightarrow y =_B v].$$

The function is said to be *relationally presented* (r.p.). In **S** it would be more natural to represent a function by means of an operation. A function is *operationally presented* (o.p.), if there is an operation $f : \sigma \to \tau$ such that

$$(\forall x^\sigma \in A)\,f(x) \in B,$$

and f respects the equalities. For a fixed domain and codomain, we can quantify over all o.p. functions, whereas this is *not* the case for r.p. functions. This is the advantage of o.p. functions. If the domain is *decidable,* an r.p. function can be converted into an o.p. function, by means of the *constructive axiom of choice* (AC):

$$\forall x^\sigma\, \exists y^\tau\, C(x, y) \;\Rightarrow\; \exists z^{\sigma\tau}\, \forall x^\sigma\, C(x, z(x)).$$

This is possible for other important cases too, see remarks in Section 2.2. Finally **S** has the *relativised dependent choice* (RDC) schema:

$$\begin{aligned}&\forall u^\sigma\, \big[B(u^\sigma) \wedge \forall x^\sigma\, (B(x) \Rightarrow \exists y^\sigma\, B(y) \wedge C(x, y))\\ &\qquad \Rightarrow\; \exists f^{0\sigma}\, [f(0) = u \wedge \forall n\, B(f(n)) \wedge C(f(n), f(n+1))]\big].\end{aligned}$$

This intuitively valid axiom is useful, and perhaps even necessary, in certain arguments in constructive topology, e.g. Baire's category theorem. Note that, if $B(x)$ is decidable, then RDC is in fact provable from AC. It seems to be an open problem whether this is the case for arbitrary $B(x)$.

2.1. Real numbers

We follow Bishop's treatment. The rational numbers are unproblematic from the constructive point of view, and can be coded as natural numbers. A real number $a = (a_n)$ is a *regular sequence* of rational numbers, i.e. $a : 0 \to 0$ such that

$$\forall n \;\; a_n \in Q \;\; \wedge \;\; (\forall m, n > 0)\, |a_m - a_n| \leqslant_Q \frac{1}{m} + \frac{1}{n}\cdot \tag{1}$$

Denote this predicate by $\mathbb{R}(a)$. Two real numbers a and b are *equal,* $a =_\mathbb{R} b$, if

$$(\forall n > 0)\, |a_n - b_n| \leqslant \frac{1}{2n}\cdot$$

Thus $(\mathbb{R}, =_\mathbb{R})$ specifies the set of real numbers.

Clearly, closed intervals can be quantified over. The same is true for the collection of compact sets (and thus countable unions and intersections of compact sets) in a fixed space. To see this, we need only to observe, with Friedman [6], that a compact set is the closure of a denumerable, totally bounded set.

We now consider the problem of representing functions on real numbers. For the arithmetical operations $+$, $-$, $\cdot$ and the operations max and $|\cdot|$ the operational presentations pose no difficulty. The reciprocal x^{-1} seems unfortunately impossible to realise as an o.p. function defined on the whole of $\{x \in \mathbb{R} : |x| > 0\}$. The construction of x^{-1} (see [4]) depends on knowing $|x| > 0$. However it can be given an operational presentation on each set $\{x \in \mathbb{R} : |x| \geqslant c\}$, where c is a positive rational number. The reciprocal is thus an o.p. function of two arguments. A higher type example is the Riemann integral as an operation: it takes a function and its continuity modulus as arguments.

2.2. Functions defined by limits

In the relational mode of presentation, a pointwise limit $f(x) = \lim_n f_n(x)$ is defined simply by writing down the formula $R_f(x, y)$ for *y is the limit of $f_n(x)$ as $n \to \infty$*. On the other hand, defining limits in the operational mode of presentation seems to require more information about convergence.

Theorem 1. *Let $S \subseteq \mathbb{R}^d$ and suppose that (f_n) is a uniform Cauchy sequence of o.p. functions on S, each uniformly continuous. Then (f_n) converges uniformly to a uniformly continuous o.p. function $f : S \to \mathbb{R}$.*

PROOF. — We only show how to explicitly construct f as an *operation*, and leave it to the reader to check that this construction works.

Since (f_n) is a Cauchy sequence, there is an increasing $M : 0 \to 0$, with $M(k) \geqslant k$, such that for positive k and all $m, n \geqslant M(k)$

$$(\forall x \in S)\ |f_m(x) - f_n(x)| \leqslant k^{-1}.$$

Now define $f : (0 \to 0)^d \to (0 \to 0)$ as follows. Let $(y_k^{(n)})_k \equiv f_n(x)$ and put $f(x) = (z_k)$ where

$$z_k = y_{M(4k)}^{(4k)}.$$ □

We note the following facts on quantifying over functions. Let X be a compact subspace of $\mathbb{R}^d$, and suppose that $f : X \to \mathbb{R}$ is a continuous r.p. function. Then, by the Stone-Weierstrass theorem, there are polynomial functions $f_n : X \to \mathbb{R}$ with $||f_n - f||_\infty < n^{-1}$. These can clearly be given operational presentations, and form a Cauchy sequence. By Theorem 1 the sequence converges to an o.p. function, which is equal to f on X. This means that we can quantify over all continuous functions on a compact subspace of $\mathbb{R}^d$. Moreover, since any locally compact subspace X of $\mathbb{R}^d$ can be written as a countable union of compact subspaces $X_n \subseteq \mathbb{R}^d$, a continuous r.p. function $f : X \to \mathbb{R}$ can be expressed by a family of continuous o.p. functions $f_n : X_n \to \mathbb{R}$, coinciding on intersecting domains. Thus we can quantify over all continuous $f : X \to \mathbb{R}$. The same is in fact true for Lebesgue integrable functions (see [4, Ch. 6]).

3. A constructive foundation for nonstandard analysis

We review the theory we introduced in [14], more accurately, we consider a slight extension of that theory with the *external* RDC schema. The theory **iS** is obtained by expanding the language of **S** with a constant $\Omega : 0$, for an infinite number, and a predicate $St^\sigma(x)$ on each type σ standing for "x is standard". Let $\forall^{\mathrm{st}} x^\sigma A$, $\exists^{\mathrm{st}} x^\sigma A$ denote $\forall x^\sigma (St(x) \to A)$ and $\exists x^\sigma (St(x) \wedge A)$ respectively. For a formula B, B^{st} means that all quantifiers are restricted in this way. The axiom schemas of **S**, now expressed in the expanded language, are modified so that certain variables are restricted to standard objects: the induction variable in the induction schema, the variables x, y, z in AC, and u, x, y, f, n in RDC (refering to the statements in the previous section). These are schemas of **iS**, called *external induction*, *external* AC and *external* RDC, respectively. The usual defining equations for combinators, $0 \neq S(u)$, and equality axioms are also axioms of **iS**. Moreover we have the axioms: $St(c)$, for each **S**-constant c; and for application: $\forall^{\mathrm{st}} x^{\sigma\tau} \forall^{\mathrm{st}} y^\sigma St(x(y))$; and the crucial *limit axioms*:

$$\forall x^\sigma \exists^{\mathrm{st}} y^{0\sigma} [x = y(\Omega)],$$

$$\forall^{\mathrm{st}} x^{0\sigma} y^{0\sigma} [x(\Omega) = y(\Omega) \Leftrightarrow \exists^{\mathrm{st}} k \forall^{\mathrm{st}} n \geqslant k \, (x(n) = y(n))].$$

These state that the intended model is a reduced power of the standard type structure, modulo the Fréchet filter. In this model $St(x)$ is interpreted as *x is eventually constant*. The generic infinity, Ω, is interpreted as the identity function. (Note the affinity to Schmieden and Laugwitz' Ω-notation [17].)

We also consider definitional extensions of **iS** of the following form. For a formula $A(x_1^{\sigma_1}, \dots, x_m^{\sigma_m})$, we add a new predicate P and the definition

$$\begin{aligned}&\forall^{\mathrm{st}} y_1^{0\sigma_1} \dots y_m^{0\sigma_m} \, [P(y_1(\Omega), \dots, y_m(\Omega)) \\ &\Leftrightarrow \;\; \exists^{\mathrm{st}} k \, \forall^{\mathrm{st}} n \geqslant k \;\; A(y_1(n), \dots, y_m(n))].\end{aligned}$$

We call this an *extension by a basic predicate*. It is easy to show, using the limit axioms, that

$$\forall^{\mathrm{st}} \vec{y} \, \big[P(y_1(\Omega), \dots, y_m(\Omega)) \Leftrightarrow \exists^{\mathrm{st}} k \, \forall^{\mathrm{st}} n \geqslant k \;\; P(y_1(n), \dots, y_m(n))\big].$$

Of course we may add arbitrarily many basic predicates simultaneously, if we wish. In the sequel we consider x *is a real number*, $\mathbb{R}(x)$, and the relations $=$, $\leqslant$, $<$ on real numbers as new basic predicates. E.g. we take (overloading notation)

$$\forall^{\mathrm{st}} r^{0(00)} \, \big[\mathbb{R}(r(\Omega)) \Leftrightarrow \exists^{\mathrm{st}} k \, \forall^{\mathrm{st}} n \geqslant k \;\; A^{\mathrm{st}}(r(n))\big],$$

where A is the formula (1), page 72, defining the real numbers in the standard theory. This extension procedure is very important since it provides nonstandard versions of standard notions, like the nonstandard real numbers above. We refer to [14] for further examples.

3.1. Nonstandard principles

We recall the following definitions from [14], needed to state our results. A formula A free from Ω-symbols is called *internal* if it does not contain the standard predicate St; the formula is *almost internal* if the St-predicate occurs only in subformulas

$$\forall i^0 [St(i) \wedge i < t \Rightarrow B] \tag{2}$$

of A, where the free variables of t are either free in A or bound by quantifiers where the range is restricted to standard objects. A variable occuring in such a t is called *confining*. The idea is that such subformulas are really conjunctions of variable finite length when t is standard. A formula is *subgeometric* if it is formed from atomic formulas using only $\wedge$ and $\exists$; the formula is *almost subgeometric* if it in addition can contain universal quantifications of the form (2), subject to the same conditions on t. The class of *constructive Horn formulas* is the least class $\mathcal{CH}$ such that

- $\mathcal{CH}$ contains the atomic formulas,
- $\mathcal{CH}$ is closed under conjunction, existential and universal quantification,
- if A is subgeometric and $B \in \mathcal{CH}$, then $(A \Rightarrow B) \in \mathcal{CH}$.

One can prove that every Horn formula is classically equivalent to a constructive Horn formula, and conversely. If we allow almost subgeometric A in the last clause, we call the resulting class of formulas *almost constructive Horn*. Almost internal formulas, which are almost subgeometric are called *amenable*, those which are almost constructive Horn are called *light*. Any amenable formula $A(\vec{x},\vec{y})$, where $\vec{y}$ are nonconfining variables, is equivalent, for *standard* parameters $\vec{x}$ and arbitrary $\vec{y}$, to a normal form

$$\exists\vec{z}\,(\forall^{\mathrm{st}} i_1 < t_1(\vec{x}))\ldots(\forall^{\mathrm{st}} i_m < t_m(\vec{x}, i_1, \ldots, i_{m-1}))\;\; B(\vec{x},\vec{y},\vec{z},\vec{i}),$$

where B is a conjunction of atomic formulas. It is an easy exercise to show this using external AC.

Theorem 2. *Let $A(\vec{x}, n^0)$ be a formula where n is not confining.*

(a) (The Łos principle.) *If A is an amenable formula, then*

$$\forall^{\mathrm{st}}\vec{x}\,\big[\exists^{\mathrm{st}} k\,(\forall^{\mathrm{st}} n \geqslant k)\;\; A^{\mathrm{st}}(\vec{x},n) \Leftrightarrow A(\vec{x},\Omega)\big].$$

(b) (The lifting principle.) *If A is a light formula, then*

$$\forall^{\mathrm{st}}\vec{x}\,\big[\exists^{\mathrm{st}} k\,(\forall^{\mathrm{st}} n \geqslant k)\;\; A^{\mathrm{st}}(\vec{x},n) \Rightarrow A(\vec{x},\Omega)\big].$$

PROOF. — By induction on the formulas; for details see [14]. □

Remarks. The theorem corresponds to results about reduced powers in model theory. In its first order form (b) is familiar from universal algebra; many properties of ordered rings can be expressed by Horn formulas. Classically, it is possible to strengthen (a) by using a result of Palyutin [15, Section 1].

The following principle is close to Keisler's [9] elementary form of transfer ("the solution axiom"). Note, however, that we do not allow negative conditions in the antecedent of the transferred formula.

Theorem 3: The transfer principle

Let $A(\vec{x}) \equiv \forall \vec{y}\,[B(\vec{x},\vec{y}) \to C(\vec{x},\vec{y})]$, *be a light formula with* B *and* C *almost subgeometric. Then*

$$\forall^{\mathrm{st}} \vec{x}\,[A^{\mathrm{st}}(\vec{x}) \Longleftrightarrow A(\vec{x})].$$

PROOF. — By the Łos principle and the lifting principle. □

One may view the saturation properties of nonstandard universes as expressing that there are always enough 'limit points' to satisfy any consistent limit process, provided it is expressed in proper language. In the constructive setting, this language is as far as we can see limited to amenable formulas. Nevertheless, the saturation principles available proves useful cases such as Lemma 12 below or the result in Section 4.3.

Theorem 4: Saturation

Let $A(\vec{u}, v, \vec{w}, i)$ *be an amenable formula, where* $\vec{u}$ *are non-confining variables, and which satisfies the chain condition*

$$\forall^{\mathrm{st}} i\, \forall \vec{u}, v, \vec{w}\, \big[A(\vec{u}, v, \vec{w}, i+1) \;\Rightarrow\; A(\vec{u}, v, \vec{w}, i)\big].$$

(a) *If* v *is also non-confining in* A,

$$\forall \vec{u}\, \forall^{\mathrm{st}} \vec{w}\, \big[\forall^{\mathrm{st}} i\, \exists v\;\; A(\vec{u}, v, \vec{w}, i) \;\Rightarrow\; \exists v\, \forall^{\mathrm{st}} i\;\; A(\vec{u}, v, \vec{w}, i)\big].$$

(b) *Let* $\hat{A}$ *be obtained by removing all restrictions to* St *in* A. *Then*

$$\forall \vec{u}\, \forall^{\mathrm{st}} \vec{w}\, \big[\forall^{\mathrm{st}} i\, \exists^{\mathrm{st}} v\;\; A(\vec{u}, v, \vec{w}, i) \;\Rightarrow\; \exists v\, \forall^{\mathrm{st}} i\; \hat{A}(\vec{u}, v, \vec{w}, i)\big].$$

PROOF. — The proof relies on Theorem 2, and in case (a) the technique is somewhat analogous to that used for countably incomplete ultrapowers. For details we refer to [14]. □

Corollary 5: overspill

Let $A(\vec{u}, \vec{w}, n)$ *be an amenable formula, where the variables* $\vec{u}$ *are non-confining. Then:*

$$\forall \vec{u}\, \forall^{\mathrm{st}} \vec{w}\, \big[\forall^{\mathrm{st}} n\, A(\vec{u},\vec{w},n) \;\Rightarrow\; (\exists \nu)\, \nu \textit{ infinite} \,\wedge\, (\forall \mu \leqslant \nu)\, \hat{A}(\vec{u},\vec{w},\mu)\big].$$

□

Remarks. A more restricted form of the overspill principle is due to P. Martin-Löf. In [14], the above and further results were proven in the weaker theory $\mathbf{iHA}^{\omega}$. There is a classical strengthening of Theorem 4 (a): A can be any internal formula.

3.2. Relating S and iS

It is possible to give an interpretation $(\cdot)^{\dagger}$ of **iS** into **S**, by formalising the intuitive model of **iS**. In [14] this was carried out for a pair of weaker systems. The interpretation and its consequences are easily extended to the present systems by

- checking that external RDC can be $(\cdot)^{\dagger}$-interpreted by RDC,
- checking that **S** has the explicit definability property. This follows by extending the modified realisability interpretation to RDC.

The most important consequences are the following two theorems.

Theorem 6. *For closed formulas A in* **S**, *and B in* **iS**:

(a) $\mathbf{iS} \vdash A^{\mathrm{st}}$ *if, and only if,* $\mathbf{S} \vdash A$,

(b) $\mathbf{S} \vdash B^{\dagger}$ *if, and only if,* $\mathbf{iS} \vdash B$. □

The first part says that **iS** is a conservative extension of **S**. The second states that **iS** completely axiomatises the model given inside **S**. The theory **iS** has the explicit definability property in two distinct forms.

Theorem 7. *Let $A(x)$ be a formula with only x free.*

(a) *If* $\mathbf{iS} \vdash \exists x A(x)$, *then for some closed* $t : \sigma$, $\mathbf{iS} \vdash A(t)$.

(b) *If* $\mathbf{iS} \vdash \exists^{\mathrm{st}} x A(x)$, *then for some closed* internal $t : \sigma$, $\mathbf{iS} \vdash A(t)$. □

4. Nonstandard analysis

In this section we work within the theory **iS**. We start with some definitions.

A number a is said to be *finite* if for some standard k, $|a| < k$; and *infinite* if for every standard k, $|a| > k$; it is *infinitesimal* if for every positive standard ε, $|a| < \varepsilon$. On a metric space (X, d), we write $a \simeq_d b$ (*a is infinitely close to b*), if $d(a, b)$ is infinitesimal; the subscript d is omitted if the metric is obvious from the context. Let $A(x, y_1, \ldots, y_n)$ be an amenable formula where all free variables are indicated. The *set*

$$\{x : A(x, t_1, \ldots, t_n)\}$$

is called *standard* if $t_1, \ldots, t_n$ are standard, and *internal* otherwise. Note that in the latter case, $t_i = s_i(\Omega)$ for some standard s_i. Hence any internal set can be written in a normal form $\{x : A_\Omega(x, \vec{s})\}$ where $\vec{s}$ are standard parameters. These notions are not as restricted as they may appear, since we may always introduce new basic predicates. All other sets will be called *external*. Some examples: define $x \in [a, b]$ iff $a \leqslant x$, $x \leqslant b$ and $x \in \mathbb{R}$. Then $[0, 1]$ is standard, but contains infinitesimals, whereas $[0, \Omega]$ is internal; the set of standard points in $[0, 1]$ is an external set. A function $f : X \to Y$ is *standard*, if X and Y are standard sets, and f is a standard operation. The notion of internal function is more delicate. Let X_Ω, Y_Ω be internal sets. An *internal function* $f : X_\Omega \to Y_\Omega$ is an operation $f = g_\Omega$ where (g_n) is a standard sequence such that

$$(\exists^{\mathrm{st}} k)(\forall^{\mathrm{st}} n \geqslant k)\ g_n : X_n \to Y_n. \tag{3}$$

For every pair of internal sets X_Ω, Y_Ω the statement (3) defines a new basic predicate $\mathrm{Int}(g_\Omega; X_\Omega, Y_\Omega)$. Note that if $x_\Omega \in X_\Omega$, then $f_\Omega(x_\Omega) \in Y_\Omega$. Internal functions are called *normal* by Schmieden and Laugwitz [17].

4.1. Nonstandard characterisations

The standard notions that most naturally lend themselves to nonstandard characterisations are those which explicitly involve sequential limits (see [14]). In the case of classical, separable metric spaces there are sequential versions of many topological notions, e.g. closure, compactness, continuity, uniform continuity. With a few exceptions (e.g. closure) these are *not* the appropriate constructive notions (see [3]). This is one reason that we cannot expect the constructive nonstandard analysis to be as neat as the classical.

The most important characterisation is, of course, that of the limit.

Theorem 8. *Let $(a_n)_n$ be a standard sequence of points in the standard metric space (X, d), and let $b \in X$ be standard. Then*

$$\lim_n a_n = b \ \Leftrightarrow\ a_\Omega \simeq_d b \ \Leftrightarrow\ (\forall \text{ infinite } \nu)\, a_\nu \simeq_d b.$$

Proof. — Since a and b are standard we can use Theorem 2, with $<$ as a new basic predicate, to conclude

$$\begin{aligned} &\big[(\forall^{\mathrm{st}}\varepsilon > 0)(\exists^{\mathrm{st}}k)(\forall^{\mathrm{st}}n \geqslant k)\ d(a_n, b) < \varepsilon\big] \\ \Leftrightarrow\ &\big[(\forall^{\mathrm{st}}\varepsilon > 0)\, d(a_\Omega, b) < \varepsilon\big]. \end{aligned} \tag{4}$$

This establishes the first equivalence. To prove the second equivalence, we use the lifting principle (Theorem 2) on the lefthand side of (4), and get $(\forall^{\mathrm{st}}\varepsilon > 0)(\exists^{\mathrm{st}}k)(\forall n \geqslant k)\ d(a_n, b) < \varepsilon$. Thus letting $n = \nu$ be infinite proves the case. □

Constructively, we have only the following halves of the wellknown nonstandard characterisations of uniform continuity and uniform convergence. The converses can be proven by classical logic.

Proposition 9. *Let (X, d) and (Y, e) be standard metric spaces. If $f : X \to Y$ is a standard function, uniformly continuous on X, then f is* monad preserving, *i.e. $f(u) \simeq_e f(v)$ for $u \simeq_d v$ in X.*

Proof. — Left to the reader (see [14]). □

Proposition 10. *Let (Y, d) be a standard metric space, and let (g_n) be a standard sequence of functions $g_n : S \to Y$ converging uniformly to the standard function $g : S \to Y$. Then for all $x \in S$ and all infinite ν, $g_\nu(x) \simeq_d g(x)$.*

Proof. — The assumption can be rendered formally as

$$(\forall^{\mathrm{st}}\varepsilon > 0)(\exists^{\mathrm{st}}k)(\forall^{\mathrm{st}}n \geqslant k)\,(\forall^{\mathrm{st}}x \in S)\quad d(g_n(x), g(x)) < \varepsilon.$$

By the lifting principle, we can remove the last two $^{\mathrm{st}}$'s and let n be any infinite ν. □

Example 1. The polynomial functions $g_n(x) = (1 + xn^{-1})^n$ converge uniformly to e^x on any compact standard interval I. Thus for infinite ν and all $x \in I$,

$$(1 + \frac{x}{\nu})^\nu \simeq e^x.$$

4.2. Near standard points

A real number x is said to be *near standard,* if there exists a standard real number y with $x \simeq y$. Unlike classical nonstandard analysis, not every finite real number is near standard. We shall give a characterisation of near standard points using the order relation. A real number z is called *determined* if for all standard reals x, y

$$x < y \quad \text{implies} \quad x < z \text{ or } z < y.$$

Every standard real number is determined. $(-1)^\Omega$ is not determined (consider $x = 0, y = 1/2$). Ω is determined, but not finite.

Proposition 11. *A real number is near standard if, and only if, it is finite and determined.*

PROOF. ($\Rightarrow$) Suppose that $z \simeq u$, where u is standard. Clearly, z is finite. Let $x < y$ be standard. Hence $x < u$ or $u < y$. If $x < u$, then $x < z$ since trivially $|z - u| < (u - x)/2$. Similarly, $z < y$, if $u < y$.

($\Leftarrow$) Suppose z is finite and determined. Hence there are standard real numbers a_0, b_0 with $a_0 < z < b_0$. Now, $(2a_0 + b_0)/3 < (a_0 + 2b_0)/3$, so there are two possible cases: $(2a_0 + b_0)/3 < z$ or $z < (a_0 + 2b_0)/3$. In the first case we let $a_1 = (2a_0 + b_0)/3$ and $b_1 = b_0$. In the second case, let $a_1 = a_0$ and $b_1 = (a_0 + 2b_0)/3$. Continuing in this way, we get a sequence of intervals (a_n, b_n), containing z, whose diameter shrinks by $2/3$ at each step. Thus $z \simeq \lim_n a_n = \lim_n b_n$. □

Remarks. We note that, formally, external RDC is used in the above argument to obtain the sequences (a_n) and (b_n).

> One could well imagine that a strengthening $\simeq'$ of $\simeq$ would make it feasible to drop the condition of determinism, so as to obtain a *classical standard part map*. If this $\simeq'$ is reasonable, then continuous functions should still respect it. However this would lead to a proof of the *exact* intermediate value theorem, see [18] or [14, Section 5.3]. But this is known to be constructively impossible (see e.g. [1]).

The lattice points $\ell + pd$, $p \in \mathbb{Z}$, on the real line, can be regarded as a 'discrete' continuum if we let d be infinitesimal. More precisely we have the

Lemma 12. *For any positive infinitesimal d, and any standard real numbers ℓ and $a \leqslant x \leqslant b$, with $a < b$, there exists an integer p such that*

$$x \simeq \ell + pd$$

and $a \leqslant \ell + pd \leqslant b$. Moreover p can be assumed to be infinite.

PROOF. — The standard statement

$$\begin{aligned}(\forall^{\mathrm{st}} n > 0)\,(\exists^{\mathrm{st}} \varepsilon > 0)\,(\forall^{\mathrm{st}} d \in (0,\varepsilon))\,(\forall^{\mathrm{st}} x \in [a,b])\,(\exists^{\mathrm{st}} p \in \mathbb{Z})\\ \big[|x - \ell - pd| \leqslant n^{-1} \;\wedge\; a \leqslant \ell + pd \leqslant b \wedge |p| \geqslant n\big]\end{aligned} \tag{5}$$

is not too difficult to establish. We leave this to the reader.

We can lift (5) to

$$(\forall^{\mathrm{st}} n > 0)\,(\exists^{\mathrm{st}} \varepsilon > 0)\,(\forall d \in (0,\varepsilon))\,(\forall x \in [a,b])\,(\exists p \in \mathbb{Z})\;\; A(x,d,p,n)$$

where $A(x,d,p,n)$ is the formula within square brackets in (5). Let $d > 0$ be infinitesimal; then for any standard $\varepsilon > 0$, $d \in (0,\varepsilon)$, so for any standard $x \in [a,b]$

$$\forall^{\mathrm{st}} n\; \exists p \in \mathbb{Z} \quad A(x,d,p,n+1).$$

By the saturation principle (Theorem 4) there exists $p \in \mathbb{Z}$ such that

$$\forall^{\mathrm{st}} n\; \big[|x - \ell - pd| \leqslant (n+1)^{-1} \;\wedge\; a \leqslant \ell + pd \leqslant b \;\wedge\; |p| \geqslant n+1\big].$$

Clearly then $x \simeq \ell + pd$ and p is certainly infinite. □

The usage of saturation was not really essential in this lemma; one could equally well have constructed x directly. A more significant application is found in Section 4.3.

The lemma below will be needed in Section 5. It is easily proven using integrating factors. But we prove it more directly to illustrate Lemma 12.

Lemma 13. *Let $f : [a,b] \to \mathbb{R}$ be a standard, non-negative continuous function and let $L > 0$ be standard so that*

$$f(x) \leqslant L \int_a^x f(t)dt \qquad (\text{for standard } x \in [a,b]).$$

Then f vanishes identically on $[a,b]$.

PROOF. — Let μ be a continuity modulus for f. Suppose $\varepsilon > 0$ is standard. By the usual estimate of the Riemann integral, we have for standard $N > 0$, with $(b-a)N^{-1} \leqslant \max(\varepsilon, \mu(\varepsilon))$ and standard $n = 0, \ldots, N$

$$f(x_n) \;\leqslant\; L \int_a^{x_n} f(t)dt \;\leqslant\; L\Big(\varepsilon + \frac{b-a}{N} \sum_{k=0}^{n-1} f(x_k)\Big),$$

where $x_k = a + k(b-a)/N$. Let $\Delta = L(b-a)/N$. By external induction on n it follows that for $n = 0, \ldots, N$

$$f(x_n) \leqslant L\varepsilon(1+\Delta)^n. \tag{6}$$

Letting $\varepsilon > 0$ be infinitesimal, there is, by lifting, an infinite N such that (6) holds for all $n \leqslant N$. Then $\varepsilon \simeq 0$, so by Example 1, page 81,

$$f(x_n) \;\leqslant\; L\varepsilon(1+\Delta)^N \;\simeq\; L\varepsilon e^{L(b-a)} \;\simeq\; 0,$$

for all $n \leqslant N$. By Lemma 12, to any $x \in [a,b]$ there is some $n \leqslant N$ with $x \simeq x_n$. Since f is continuous,

$$f(x) \simeq f(x_n) \simeq 0. \qquad \square$$

4.3. A nonstandard summation method

We prove a nonstandard theorem about regular summation which generalises Toeplitz' method. The result is a constructive reformulation of a theorem due to Robinson [16]. The main tool is *Robinson's sequential lemma.*

Lemma 14. *If* (a_n) *is an internal sequence of real numbers with* $a_n \simeq 0$ *for all standard* n*, then there exists an infinite* μ *such that* $a_\nu \simeq 0$ *when* $\nu \leqslant \mu$.

PROOF. — Consider the amenable formula

$$A(a, m) \;\equiv\; (\forall^{\mathrm{st}} n \leqslant m)\, |a_n| \leqslant 2^{-m}.$$

Clearly $A(a, m)$ holds for all standard m.
Hence by overspill, $(\forall n \leqslant \mu)|a_n| \leqslant 2^{-\mu}$ for some infinite μ. □

Theorem 15. *Let* (a_n) *be an internal sequence of real numbers with the property that for some standard* B, $\sum_{n=0}^{\nu} |a_n| < B$ *holds for all* ν. *Then the following two conditions are equivalent.*

(a) *For every standard convergent sequence* (s_n) *with limit* L*, there exists an infinite* η *such that*

$$L \simeq \sum_{n=0}^{\mu} a_n s_n \qquad (\mu \geqslant \eta).$$

(b) *For all standard* k, $a_k \simeq 0$*; and there exists an infinite* η *with*

$$\sum_{n=0}^{\mu} a_n \simeq 1 \qquad (\mu \geqslant \eta).$$

PROOF. — (a) ⇒ (b): Let (s_n) be the standard sequence which is 1 at the standard index k and 0 otherwise. Its limit is 0 and so by (a) for some infinite μ:

$$\sum_{n=0}^{\mu} a_n s_n \simeq 0.$$

But the left hand side is a_k, so $a_k \simeq 0$. This proves the first part of (b); as for the second part we take $s_n \equiv 1$ instead.

(b) $\Rightarrow$ (a): Suppose (s_n) is a standard sequence converging to L. From the first part of the assumption (b) follows that $\sum_{n=0}^{k} |a_n| \simeq 0$ for all standard k. The sequence s is bounded by a standard number, C, say. Thus for standard k

$$\Big|\sum_{n=0}^{k} s_n a_n\Big| \leqslant \sum_{n=0}^{k} |s_n||a_n| \leqslant C \sum_{n=0}^{k} |a_n|.$$

Hence $\sum_{n=0}^{k} s_n a_n \simeq 0$ for all standard k. By Robinson's sequential lemma there is an infinite κ such that

$$\sum_{n=0}^{\kappa-1} a_n \simeq 0 \qquad \sum_{n=0}^{\kappa-1} s_n a_n \simeq 0. \tag{7}$$

By the second part of (b) there is an infinite τ with

$$\sum_{n=0}^{\mu} a_n \simeq 1 \qquad (\mu \geqslant \tau). \tag{8}$$

Let $\eta = \max(\tau, \kappa)$. Then for $\mu \geqslant \eta$

$$\Big|\sum_{n=\kappa}^{\mu} a_n s_n - \sum_{n=\kappa}^{\mu} a_n L\Big| \leqslant \sum_{n=\kappa}^{\mu} |a_n||s_n - L| \leqslant \delta \sum_{n=\kappa}^{\mu} |a_n|, \tag{9}$$

where $\delta = \max\{|s_n - L| : \kappa \leqslant n \leqslant \mu\}$. δ is infinitesimal, since $L = \lim_n s_n$. But $\sum_{n=0}^{\mu} |a_n| < B$ by the overall assumption, so the lefthand side of (9) is infinitesimal. Using (7) and (8) we get

$$\sum_{n=0}^{\mu} a_n s_n \simeq \sum_{n=\kappa}^{\mu} a_n s_n \qquad \sum_{n=0}^{\mu} a_n \simeq \sum_{n=\kappa}^{\mu} a_n \simeq 1.$$

Thus by (9)

$$\sum_{n=0}^{\mu} a_n s_n \simeq \sum_{n=\kappa}^{\mu} a_n s_n \simeq L \sum_{n=\kappa}^{\mu} a_n \simeq L \sum_{n=0}^{\mu} a_n \simeq L.$$

□

Let us now see how this result generalises Toeplitz' theorem.

Corollary 16. *Let $B = (b_{k,n})$ be a standard double sequence such that*

(a) $\forall^{\mathrm{st}} n \ \lim_k b_{k,n} = 0$,

(b) $\exists^{\mathrm{st}} C \ \forall^{\mathrm{st}} k \ \sum_{n=0}^{\infty} |b_{k,n}| < C$,

(c) $\lim_k \sum_{n=0}^{\infty} b_{k,n} = 1$.

For every bounded standard sequence (s_n) define

$$t_k = \sum_{n=0}^{\infty} b_{k,n} s_n .$$

If (s_n) is convergent, then (t_k) converges to the same limit.

PROOF. — Define an internal sequence by letting $a_n = b_{\Omega,n}$. By a suitable lifting of (b) we see that this sequence satisfies the overall assumption of Theorem 15. Clearly (a) implies that $a_n \simeq 0$, for all standard n. From (c) follows

$$(\forall^{\mathrm{st}} N)\,(\exists^{\mathrm{st}} \ell)\,(\forall^{\mathrm{st}} k \geqslant \ell)\,(\exists^{\mathrm{st}} m)\,(\forall^{\mathrm{st}} p \geqslant m)\ \Big|\sum_{n=0}^{p} b_{k,n} - 1\Big| \ < \ 2^{-N}.$$

By the lifting theorem, the restrictions to standard objects can be removed in the three last quantifiers. Let $k = \Omega$. Thus

$$(\forall^{\mathrm{st}} N)\,(\exists \mu_N)\,(\forall \nu \geqslant \mu_N)\ \Big|\sum_{n=0}^{\nu} a_n - 1\Big| \ < \ 2^{-N}. \tag{10}$$

By saturation, there is an infinite μ greater than all μ_N (N standard). Hence $\sum_{n=0}^{\nu} a_n \simeq 1$, for all $\nu \geqslant \mu$, thus satifying condition (b) of Theorem 15. For a standard sequence (s_n) with limit L, we get by the theorem $L \simeq \sum_{n=0}^{\mu} a_n s_n$ for all $\mu \geqslant \eta$, where η is some infinite number. Suppose $\mu = M(\Omega)$. Thus

$$\lim_k \sum_{n=0}^{M(k)} b_{k,n} s_n = L.$$

Since M may grow arbitrarily fast we have shown that $\lim_k t_k = L$. $\square$

Remark. Laugwitz [10, Ch.5] obtains, using his semiconstructive nonstandard analysis, a result similar to Theorem 15 together with interesting applications to divergent series.

4.4. S-continuity and S-differentiability

We extend the notion of continuity and differentiability to internal functions. First we introduce versions of the inequality relations compatible with $\simeq$:

$$\begin{aligned} x \precsim y \quad &\Leftrightarrow_{\mathrm{def}} \quad (\forall^{\mathrm{st}} n > 0)\; x - n^{-1} < y, \\ x \prec y \quad &\Leftrightarrow_{\mathrm{def}} \quad (\exists^{\mathrm{st}} n > 0)\; x + n^{-1} < y. \end{aligned}$$

It is easily checked that $\precsim$ and $\leqslant$ coincide for standard arguments. The same holds for $\prec$ and $<$. Also, $\precsim$ is a partial order w.r.t. the equivalence relation $\simeq$.

Let (X, d) and (Y, e) be internal metric spaces. An internal function $f : X \to Y$ is *uniformly S-continuous on* $U \subseteq X$ if there exists a standard modulus m such that for standard $\varepsilon > 0$ and any $x, y \in U$

$$d(x, y) \precsim m(\varepsilon) \;\Rightarrow\; e(f(x), f(y)) \precsim \varepsilon.$$

Note that a standard continuous function is continuous also in the "S" sense. We follow the usual convention in constructive analysis, and drop the prefix "uniformly" if U is compact.

Example 2. Let $\delta > 0$ be infinitesimal. Then $f(x) = (x + \delta)^2$ defines an S-continuous function $[0, 1] \to \mathbb{R}$.

Let $I \subseteq \mathbb{R}$ and let $(X, \|\cdot\|)$ be a normed space. An internal function $f : I \to X$ is *uniformly S-differentiable on* I if there exist a uniformly S-continuous function $g : I \to X$ (an *S-derivative of* f) and a standard modulus d such that for all standard $\varepsilon > 0$ and all $x, y \in I$

$$|x - y| \precsim d(\varepsilon) \;\Rightarrow\; \|f(x) - f(y) - g(x)(x - y)\| \precsim \varepsilon |x - y|.$$

Piecewise linear functions can be constructed according to the following

Proposition 17. *Let* $\mathcal{P} = (a_1, b_1), \ldots, (a_n, b_n)$ *be a finite standard sequence of pairs of real numbers with* $a_1 < \cdots < a_n$. *Then there exists a uniformly continuous standard function* $f = f_{\mathcal{P}} : \mathbb{R} \to \mathbb{R}$ (the polygonial function determined by $\mathcal{P}$) *such that*

(a) $f(x) = b_1$ *for* $x \leqslant a_1$,

(b) $f(x) = b_n$ *for* $x \geqslant a_n$,

(c) $f(x) = \dfrac{x - a_{i+1}}{a_i - a_{i+1}}(b_i - b_{i+1}) + b_{i+1}$ *for* $a_i \leqslant x \leqslant a_{i+1}$.

PROOF. — Define $g(a, b, x) = ba^{-1}(\max(0, x) - \max(0, x - a))$, and let

$$f(x) = b_1 + \sum_{k=1}^{n-1} g(a_{k+1} - a_k, b_{k+1} - b_k, x - a_k).$$

It is easy to check that this function satisfies the conditions. □

Example 3. Let $f_N : [0, 1] \to \mathbb{R}$ be the polygonial standard function given by the points

$$(0, 0), (N^{-1}, N^{-2}), \ldots, (kN^{-1}, k^2N^{-2}), \ldots, (1, 1).$$

Then the internal function f_Ω is S-differentiable on $[0, 1]$ with S-derivative $g(x) = 2x$.

An S-continuous function $f : X \to Y$ is said to be of class S^0, if $f(x)$ is near standard, for every near standard $x \in X$.

Theorem 18. *Let* $[a, b] \subseteq \mathbb{R}$ *be a standard interval. If* $f : [a, b] \to \mathbb{R}$ *is of class* S^0, *then there exists a unique continuous standard function* $g : [a, b] \to \mathbb{R}$ *such that* $f(x) \simeq g(x)$, *for all standard* $x \in [a, b]$.

PROOF. — Unicity is clear. To prove existence, let m be an S-continuity modulus of f. For every standard $n > 0$, let $a = a_0 < \cdots < a_k = b$ be a uniform standard subdivision of $[a, b]$ such that $a_{i+1} - a_i \leqslant m(1/n)$. For $i = 0, \ldots, k$, take b_i to be standard with $b_i \simeq f(a_i)$. Let g_n be the polygonial function determined by the points (a_i, b_i), $i = 0, \ldots, k$. Clearly $|f(x) - g_n(x)| \leqslant n^{-1}$ for all standard $x \in [a, b]$. Let g be the limit of the Cauchy sequence (g_n). □

Example 4. $f(x) = (-1)^{\Omega}x$ defines an S-continuous bounded function $[0,1] \to \mathbb{R}$ which is not near any standard function.

5. Existence of solutions to differential equations

The existence theorem for ordinary differential equations is constructively valid under a Lipschitz condition. Beeson [1] noted that Picard's successive approximation technique works in this case. The numerically (slightly) more interesting method of Euler is representative for many step methods. We shall give two nonstandard variants of an existence proof using this method. The first is close to the usual nonstandard proof (Laugwitz [10], Diener and Reeb [5]). The second is closer to a standard proof by Henrici [8].

Let $(\mathbb{R}^d, \|\cdot\|)$ be the usual euclidean space of dimension d; vectors in this space are denoted by boldface letters $\mathbf{x} = (x^{(1)}, \dots, x^{(d)})$. Let $\mathbf{f} : [a,b] \times \mathbb{R}^d \to \mathbb{R}^d$ be a standard, continuous function, and let $\mathbf{x}_0 \in \mathbb{R}^d$ be a standard point. A solution to the *initial value problem* with these data is a standard, continuously differentiable function $\mathbf{v} : [a,b] \to \mathbb{R}^d$ such that

$$\begin{cases} \mathbf{v}(a) = \mathbf{x}_0 \\ \mathbf{v}'(t) = \mathbf{f}(t, \mathbf{v}(t)) \quad (a \leqslant t \leqslant b). \end{cases} \tag{11}$$

Recall that the polygonial approximations according to Euler's method are constructed as follows. Let $N > 0$ be a standard natural number, and let $h = (b-a)/N$ be the step size. Define lattice points on $[a,b]$ by $t_n = a + nh$ (sometimes written $t_{N,n}$ to emphasise the subdivision factor). We define a polygonial function $\mathbf{u} = \mathbf{u}_N$ such that

$$\begin{aligned} \mathbf{u}(t_0) &= \mathbf{x}_0 \\ \mathbf{u}(t_{n+1}) &= \mathbf{u}(t_n) + h\mathbf{f}(t_n, \mathbf{u}(t_n)). \end{aligned}$$

(Technically this is achieved using Proposition 17. Let $\mathbf{u}_N(t) = (u_N^{(1)}(t), \dots, u_N^{(d)}(t))$ be the vector valued function given by the component polygonial functions $u_N^{(i)}(t)$, in turn determined by the points $(t_0, a_0^{(i)}), \dots, (t_N, a_N^{(i)})$, where $\mathbf{a}_j = (a_j^{(1)}, \dots, a_j^{(d)})$ and $\mathbf{a}_0 = \mathbf{x}_0$, $\mathbf{a}_{k+1} = \mathbf{a}_k + h\mathbf{f}(t_k, \mathbf{a}_k)$.)

A direct consequence of this definition is

$$\begin{aligned}\mathbf{u}(t_{n+k}) - \mathbf{u}(t_n) &= \sum_{i=0}^{k-1}(\mathbf{u}(t_{n+i+1}) - \mathbf{u}(t_{n+i})) \\ &= \sum_{i=0}^{k-1} h\mathbf{f}(t_{n+i}, \mathbf{u}(t_{n+i})). \qquad (12)\end{aligned}$$

Now lifting this statement and letting $N = \eta$ be infinite, $\mathbf{u}_\eta$ should be some kind of nonstandard solution. Indeed, suppose that there is a standard continuous function $\mathbf{v} : [a, b] \to \mathbb{R}^d$ with $\mathbf{v}(t) \simeq \mathbf{u}_\eta(t)$ for all $t \in [a, b]$. Let $n = 0$. Then

$$\mathbf{v}(t_k) - \mathbf{v}(t_0) + \delta = \sum_{i=0}^{k-1} h\mathbf{f}(t_i, \mathbf{v}(t_i)) + \sum_{i=0}^{k-1} h\varepsilon_i,$$

where $\varepsilon_i = \mathbf{f}(t_i, \mathbf{u}_\eta(t_i)) - \mathbf{f}(t_i, \mathbf{v}(t_i))$ and δ are all infinitesimal. We have $|\sum_{i=0}^{k-1} h\varepsilon_i| \leqslant kh \max_i |\varepsilon_i|$ and the lefthand side of this inequality is infinitesimal. Consequently by the hyperfinite sum expression for the integral

$$\mathbf{v}(t_k) - \mathbf{v}(t_0) \simeq \int_{t_0}^{t_k} \mathbf{f}(s, \mathbf{v}(s))ds.$$

Now $\mathbf{v}(t_0) = \mathbf{x}_0$, and to every standard $t \in [a, b]$, there is a k with $t \simeq t_k$ (Lemma 12). Hence

$$\mathbf{v}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{v}(s))ds.$$

This is clearly a solution to the initial value problem.

The problem is thus to find such a standard continuous function. In *classical* nonstandard analysis it can readily be obtained by taking standard parts. Let $\mathbf{v}(t) = {}^{\circ}\mathbf{u}_\eta(t_n)$, where $t_n \leqslant t < t_{n+1}$; continuity follows from (12) under some boundedness condition on $\mathbf{f}$. This yields Peano's existence theorem. However, it is known to be constructively unprovable, without further conditions on $\mathbf{f}$ (Beeson [1, p.15]). A constructive possibility is to prove that the sequence of approximating polygonial functions is a Cauchy sequence. Then its limit is a function $\mathbf{v}$ of the required kind (Proposition 10). This can be achieved by imposing a Lipschitz condition on $\mathbf{f}$, which also gives uniqueness of the solution.

Theorem 19

(a) *Let* $\mathbf{f} : [a, b] \times \mathbb{R}^d \to \mathbb{R}^d$ *be a standard, continuous function satisfying the Lipschitz condition*

$$\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\|,$$

where $L > 0$ *is standard. Let* $\mathbf{x}_0$ *be any standard point in* $\mathbb{R}^d$*. Then there exists a unique, standard solution* $\mathbf{v} : [a, b] \to \mathbb{R}^d$ *to the initial value problem (11).*

(b) *If* $\mathbf{f}$ *is required to satisfy the conditions of* (a) *only on the standard set*

$$K(b_0) = \{(t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^d : a \leqslant t \leqslant b_0, \|\mathbf{x} - \mathbf{x}_0\| \leqslant \beta\},$$

then there is a unique standard solution $\mathbf{v}$ *to the initial value problem (11) with* b *replaced by* $b_1 = \min(b_0, a + \beta/B)$*, where* $B = \sup\{\|\mathbf{f}(\mathbf{w})\| : \mathbf{w} \in K(b_0)\}$*, and where the graph of* $\mathbf{v}$ *lies within* $K(b_0)$*.*

PROOF of uniqueness. — Suppose that $\mathbf{u}$ and $\mathbf{v}$ are two solutions to (11). Their difference can be written as

$$\delta(t) = \mathbf{u}(t) - \mathbf{v}(t) = \int_a^t \mathbf{f}(s, \mathbf{u}(s)) - \mathbf{f}(s, \mathbf{v}(s))ds,$$

for $a \leqslant t \leqslant b$. By the Lipschitz condition we get

$$\|\delta(t)\| \leqslant L \int_a^t \|\mathbf{u}(s) - \mathbf{v}(s)\|ds = L \int_a^t \|\delta(s)\|ds. \tag{13}$$

Thus by Lemma 13, page 83, $\|\delta(t)\| = 0$ for all $t \in [a, b]$. Noting that $b_1 \leqslant b_0$ this argument works also in case (b), with b replaced by b_1.

For the proofs of existence (we will actually give two variants) some lemmata are needed. The estimates in the proof are based on Henrici's [8, pp. 15–26, 108–118] presentation. The approximating polygonial functions $\mathbf{u}_N$ can be bounded by a standard estimate:

Lemma 20. *For all standard $t \in [a,b]$ and all standard $n \in \mathbb{N}$,*

$$\|\mathbf{u}_n(t)\| \;\leqslant\; e^{(a-b)L}\|\mathbf{x}_0\| + \frac{e^{(a-b)L}-1}{L}c,$$

where $c = \sup\{\|\mathbf{f}(t,\mathbf{0})\| : a \leqslant t \leqslant b\}$.

PROOF. — Henrici's [8, pp.18–19] proof is constructive and gives the bound at the lattice points of $\mathbf{u}_n$. □

Let Y be this uniform bound of $\|\mathbf{u}_n(t)\|$ on $[a,b]$ and put $D = [-Y,Y]^d$; it is sufficient to consider $\mathbf{f}$ on the domain $[a,b]\times D$. The continuity modulus ω defined in the next lemma is crucial in various estimates to follow.

Lemma 21. *The standard function $\omega : [0,\infty) \to \mathbb{R}$ defined by*

$$\omega(\varepsilon) = \sup\big\{\|\mathbf{f}(x,\mathbf{y}) - \mathbf{f}(x',\mathbf{y})\| : |x-x'| \leqslant \varepsilon;\ x,x' \in [a,b], \mathbf{y} \in D\big\}$$

is uniformly continuous and subadditive, i.e. $\omega(\varepsilon+\varepsilon') \leqslant \omega(\varepsilon)+\omega(\varepsilon')$, *for* $\varepsilon,\varepsilon' \geqslant 0$.

PROOF. — The argument uses only standard objects. Note that ω is welldefined since the supremum is taken over a continuous image of a compact set.

First we establish uniform continuity. Let m be the continuity modulus of $\mathbf{f}$. Let $\delta > 0$ and $|\varepsilon-\varepsilon'| \leqslant m(\delta)$. We show that $\omega(\varepsilon) \leqslant \omega(\varepsilon')+\delta$ (the proof is similar for $\omega(\varepsilon') \leqslant \omega(\varepsilon)+\delta$). Consider $x,x' \in [a,b]$ with $|x-x'| \leqslant \varepsilon$. Thus

$$|x-x'| \;\leqslant\; \varepsilon' + |\varepsilon-\varepsilon'| \;\leqslant\; \varepsilon' + m(\delta).$$

We can choose $u \in [a,b]$ with $|x-u| \leqslant \varepsilon'$ and $|u-x'| \leqslant m(\delta)$. Hence

$$\begin{aligned}\|\mathbf{f}(x,\mathbf{y}) - \mathbf{f}(x',\mathbf{y})\| &\leqslant \|\mathbf{f}(x,\mathbf{y}) - \mathbf{f}(u,\mathbf{y})\| + \|\mathbf{f}(u,\mathbf{y}) - \mathbf{f}(x',\mathbf{y})\| \\ &\leqslant \|\mathbf{f}(x,\mathbf{y}) - \mathbf{f}(u,\mathbf{y})\| + \delta.\end{aligned}$$

Since y and $|x-x'| \leqslant \varepsilon$ were arbitrary, $\omega(\varepsilon) \leqslant \omega(\varepsilon')+\delta$.

We now turn to subadditivity. It suffices to prove that for each $\alpha > 0$,

$$\omega(\varepsilon+\varepsilon') \;\leqslant\; \omega(\varepsilon)+\omega(\varepsilon')+\alpha. \tag{14}$$

Let $|x - x'| \leqslant \varepsilon + \varepsilon'$. We construct, in turn, $u, u' \in [a, b]$ such that $|x - u| \leqslant \varepsilon$, $|u - u'| \leqslant \varepsilon'$, $|u' - x'| \leqslant m(\alpha)$. Thus

$$\|\mathbf{f}(x, \mathbf{y}) - \mathbf{f}(x', \mathbf{y})\| \leqslant \|\mathbf{f}(x, \mathbf{y}) - \mathbf{f}(u, \mathbf{y})\| + \|\mathbf{f}(u, \mathbf{y}) - \mathbf{f}(u', \mathbf{y})\| + \alpha.$$

Since y and $|x - x'| \leqslant \varepsilon + \varepsilon'$ were arbitrary, this proves (14). □

Lemma 22. $\mathbf{v}_n = \mathbf{u}_{2^n}$ *defines a uniform Cauchy sequence.*

Proof. — The proof is standard and is almost a copy of the proof in [8, pp. 21–22, 113–115]. Let n and $\ell > 0$ be natural numbers. We estimate the norm of

$$d(x) = \mathbf{v}_m(x) - \mathbf{v}_n(x),$$

where $m = n + \ell$. It clearly suffices to consider the lattice points of $\mathbf{v}_m$. Write

$$t_i = t_{2^n,i}, \quad t'_i = t_{2^m,i}, \quad h = 2^{-n}(b - a), \quad h' = 2^{-m}(b - a).$$

Let $i = k2^\ell + j$, $0 \leqslant j < 2^\ell$. It is easy to check that

$$\mathbf{v}_m(t'_{i+1}) - \mathbf{v}_m(t'_i) = h'\mathbf{f}(t'{}_i, \mathbf{v}_m(t'_i)), \tag{15}$$

$$\mathbf{v}_n(t'_{i+1}) - \mathbf{v}_n(t'_i) = h'\mathbf{f}(t_k, \mathbf{v}_n(t_k)). \tag{16}$$

Thus

$$\begin{aligned}
\|d(t'_{i+1}) - d(t'_i)\| &= h'\|\mathbf{f}(t'_i, \mathbf{v}_m(t'_i)) - \mathbf{f}(t_k, \mathbf{v}_n(t_k))| \\
&\leqslant h'\|\mathbf{f}(t'_i, \mathbf{v}_m(t'_i)) - \mathbf{f}(t'_i, \mathbf{v}_n(t'_i))\| \\
&\quad + h'\|\mathbf{f}(t'_i, \mathbf{v}_n(t'_i)) - \mathbf{f}(t_k, \mathbf{v}_n(t'_i))\| \\
&\quad + h'\|\mathbf{f}(t_k, \mathbf{v}_n(t'_i)) - \mathbf{f}(t_k, \mathbf{v}_n(t_k))\|.
\end{aligned}$$

Using the Lipschitz condition and the ω-modulus this is seen to be majorised by

$$h'L\|d(t'_i)\| \;+\; h'\omega(|t'_i - t_k|) \;+\; h'L\|\mathbf{v}_n(t'_i) - \mathbf{v}_n(t_k)\|.$$

Now, since $t_k = t'_{k2^\ell}$, $|t'_i - t_k| \leqslant h$, $jh' \leqslant h$ and (16) hold, this is no greater than

$$h'L\|d(t'_i)\| \;+\; h'\omega(h) \;+\; h'LhY.$$

Thereby we have

$$\|d(t'_{i+1})\| \leqslant A\|d(t'_i)\| + B,$$

where $A = 1 + h'L$, $B = h'\omega(h) + h'hLY$. Solving this recursive difference inequality and noting that $d(t'_0) = 0$ we obtain

$$\|d(t'_i)\| \leqslant A^i\|d(t'_0)\| + \frac{A^i - 1}{A - 1}B \leqslant \frac{A^{2^m} - 1}{A - 1}B.$$

Since $1 + h'L < e^{h'L}$ it follows that

$$\|d(t'_i)\| \leqslant \frac{e^{2^m h'L} - 1}{h'L}B = \frac{e^{(b-a)L} - 1}{L}(\omega(h) + hLY).$$

The righthand side does not depend on i, and goes to zero as n grows. □

First proof of Theorem 19. — Let $\mathbf{v}$ be the (uniform) limit of the sequence $\mathbf{v}_n$. Then $\mathbf{v}$ is standard and continuous. By Proposition 10, page 80 we have $\mathbf{v}(t) \simeq \mathbf{v}_\Omega(t) = \mathbf{u}_{2^\Omega}(t)$ for all $t \in [a, b]$. By the remarks preceding Theorem 19, this is a solution to the initial value problem (11).

Second proof of Theorem 19. — Here we do not rely on the fundamental theorem of calculus.

Lemma 23. *Let $N > 0$ be standard and $t_n = t_{N,n}$. If $0 \leqslant m, n \leqslant N$ are standard, then*

$$\begin{aligned}&\|\mathbf{u}_N(t_m) - \mathbf{u}_N(t_n) - \mathbf{f}(t_n, \mathbf{u}_N(t_n))(t_m - t_n)\| \\ &\qquad \leqslant LY|t_m - t_n|^2 + |t_m - t_n|\omega(|t_m - t_n|).\end{aligned} \tag{17}$$

Proof. — Let $g(m, n)$ denote the lefthand side of (17). Write $\mathbf{u}(x) = \mathbf{u}_N(x)$. We consider the case $m \geqslant n$ ($m \leqslant n$ is similar) and prove the result by induction on $k = m - n$. For $k = 0$, this is clear. By unwinding the recursive definition of $\mathbf{u}(t_{m+1})$ one step we get:

$$g(m + 1, n) \leqslant g(m, n) + h\|f(t_m, \mathbf{u}(t_m)) - f(t_n, \mathbf{u}(t_n))\|.$$

The second term can be estimated by

$$\begin{aligned}&h\|f(t_m, \mathbf{u}(t_m)) - f(t_m, \mathbf{u}(t_n))\| + h\|f(t_m, \mathbf{u}(t_n)) - f(t_n, \mathbf{u}(t_n))\| \\ &\quad \leqslant hL\|\mathbf{u}(t_m) - \mathbf{u}(t_n)\| + h\omega(|t_m - t_n|).\end{aligned}$$

by the Lipschitz condition and the modulus. Moreover $\|\mathbf{u}(t_m)-\mathbf{u}(t_n)\| \leqslant Y|t_m - t_n|$, so from the inductive hypothesis we get

$$\begin{aligned} g(m+1,n) &\leqslant LY(|t_m-t_n|^2 + h|t_m-t_n|) + (|t_m-t_n|+h)\omega(|t_m-t_n|) \\ &\leqslant LY|t_{m+1}-t_n|^2 + |t_{m+1}-t_n|\omega(|t_{m+1}-t_n|). \end{aligned}$$

The last step uses also that ω is increasing. □

First we find an "S-solution" to the system (11). Let $\mathbf{v}_n \equiv \mathbf{u}_{2^n}$ and $\mathbf{g}_n(t) = \mathbf{f}(t, \mathbf{v}_n(t))$.

Claim: $\mathbf{g}_\Omega$ *is an S-derivative of* $\mathbf{v}_\Omega$ *on* $T = \{t_n : 0 \leqslant n \leqslant 2^\Omega\}$, *where* $t_n = t_{2^\Omega,n}$.

Proof. — Let $d(\varepsilon) = \min(\varepsilon/(2LY), M(\varepsilon/2))$ where M is a continuity modulus of ω. By lifting the statement of Lemma 23, and putting $N = \Omega$, we have for $|t_m - t_n| \preccurlyeq d(\varepsilon)$,

$$\begin{aligned} &\|\mathbf{v}_\Omega(t_m) - \mathbf{v}_\Omega(t_n) - \mathbf{f}(t_n, \mathbf{v}_\Omega(t_n))(t_m-t_n)\| \\ &\quad\leqslant LY|t_m-t_n|^2 + |t_m-t_n|\omega(|t_m-t_n|) \\ &\quad\preccurlyeq \frac{\varepsilon}{2}|t_m-t_n| + |t_m-t_n|\frac{\varepsilon}{2} \quad = \quad \varepsilon|t_m-t_n|. \end{aligned} \tag{18}$$

Clearly $\mathbf{v}_\Omega(a) = \mathbf{v}_\Omega(t_0) = \mathbf{x}_0$.

Let $\mathbf{v}$ be the (uniform) limit of the standard sequence $\mathbf{v}_n$. We have for all $t \in [a,b]$, $\mathbf{v}(t) \simeq \mathbf{v}_\Omega(t)$, and hence $\mathbf{f}(t,\mathbf{v}(t)) \simeq \mathbf{f}(t, \mathbf{v}_\Omega(t))$. Lemma 12 gives that for any standard $x, y \in [a,b]$ there exist $t_m \simeq x$ and $t_n \simeq y$. Hence, by the S-solution (18), $|x-y| \leqslant d(\varepsilon)$ implies

$$\|\mathbf{v}(x) - \mathbf{v}(y) - \mathbf{f}(y, \mathbf{v}(y))(x-y)\| \leqslant \varepsilon|x-y|.$$

$\mathbf{v}$ is the sought solution to (11).

As for part (b), in both variants of the proof, we need only to check that the graphs of the approximating polygonial functions on $[a, b_1]$ lie within $K(b_1)$. Let $h = (b_1 - a)/N$, $t_k = a + kh$. By induction one easily proves the first inequality of

$$\|\mathbf{u}_N(t_k) - \mathbf{x}_0\| \leqslant khB \leqslant (b_1 - a)B \leqslant \beta,$$

for $k = 0, \ldots, N$. By convexity the result follows. □

6. Concluding remarks

As may have been observed in the above nonstandard arguments, the standard proofs shine through in many cases! The reason for this is the intuitive interpretation of infinite numbers embodied in the limit axioms. In fact, these axioms yield a translation procedure which converts any nonstandard statement into a standard statement. We sketch the method for statements without defined basic predicates. The first limit axiom can be used to eliminate all quantification over nonstandard objects. Using the second limit axiom the Ω's can finally be eliminated. The method is simple compared to Nelson's [13] procedure for classical internal set theory. This gives further evidence that the difference between nonstandard analysis and standard analysis is greater in the classical case than in the constructive approach presented here.

Acknowledgement. The author is grateful to Thierry Coquand and Viggo Stoltenberg-Hansen for useful comments on an earlier version of this paper.

References

[1] M. Beeson. *Foundations of Constructive Mathematics.* Springer, Berlin – Heidelberg – New York, 1985.

[2] E. Bishop. Compiling Mathematics into Algol. Unpublished manuscript.

[3] E. Bishop. *Foundations of Constructive Analysis.* McGraw-Hill, New York, 1967.

[4] E. Bishop and D. Bridges. *Constructive Analysis.* Springer, Berlin – Heidelberg – New York, 1985.

[5] F. Diener et G. Reeb. *Analyse Non Standard.* Hermann, Paris, 1989.

[6] H. Friedman. Set Theoretic Foundations for Constructive Analysis. *Ann. of Math.*, Vol. 105, pp. 1–28, 1977.

[7] N.D. Goodman and J. Myhill. The formalization of Bishop's constructive mathematics. In F.W. Lawvere, editor, *Toposes, Algebraic*

Geometry and Logic, volume 274 of *Lecture Notes in Mathematics*. Springer, Berlin – Heidelberg – New York, 1972.

[8] P. Henrici. *Discrete Variable Methods in Ordinary Differential Equations*. Wiley, New York, 1962.

[9] H.J. Keisler. *Elementary Calculus*. Prindle, Weber and Schmidt, 1976.

[10] D. Laugwitz. *Zahlen und Kontinuum*. Bibliographisches Inst., Mannheim – Wien – Zürich, 1986.

[11] P. Martin-Löf. Mathematics of Infinity. In P. Martin-Löf and G.E. Mints, editors, *COLOG-88 Computer Logic*, volume 417 of *Lecture Notes in Computer Science*, pp. 146 – 197. Springer, Berlin – Heidelberg – New York, 1989.

[12] E. Nelson. Internal Set Theory: a new approach to nonstandard analysis. *Bull. Amer. Math. Soc.*, Vol. 83, pp. 1165–1198, 1977.

[13] E. Nelson. The syntax of nonstandard analysis. *Ann. Pure Appl. Logic*, Vol. 38, pp. 123–134, 1988.

[14] E. Palmgren. A constructive approach to nonstandard analysis. *Ann. Pure Appl. Logic*, Vol. 73, pp. 297–325, 1995.

[15] E. A. Palyutin. Categorical Horn classes I. *Algebra and Logic*, Vol. 19, pp. 377–400, 1980.

[16] A. Robinson. On generalized limits and linear functionals. *Pac. J. Math.*, Vol. 14, pp. 269–283, 1964.

[17] C. Schmieden und D. Laugwitz. Eine Erweiterung der Infinitesimalrechnung. *Math. Zeitschr.*, Vol. 69, pp. 1–39, 1958.

[18] F. Wattenberg. Nonstandard analysis and constructivism? *Studia Logica*, Vol. 47, pp. 303–309, 1988.

Department of Mathematics
Uppsala University
PO Box 480
S-751 06 Uppsala
Sweden
e-mail: `Erik.Palmgren@math.uu.se`