

1. FIRST STEPS

The purpose of this section is to demonstrate some characteristic reasoning from nonstandard analysis, and to do this as quickly and simply as possible. We assume familiarity with the basic parts of first-order logic, including the idea of a nonstandard model and how to obtain one. We avoid set theoretic constructions in this section, in order to keep the mathematical context as uncomplicated as possible. We work at the level of undergraduate mathematics in order to present results that are accessible to everyone. We give only some key examples and do not try to develop these ideas systematically in this section.

Let $\mathcal{R} = (\mathbb{R}, \dots)$ be an expansion of the field \mathbb{R} ; we will assume that for every $n \geq 1$, every subset of \mathbb{R}^n and every function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is a primitive of \mathcal{R} . (That is, each such subset or function is the interpretation in \mathcal{R} of some nonlogical symbol in the language of \mathcal{R} .) We let L denote the first-order language for which \mathcal{R} is a structure.

For the rest of this section we fix a proper elementary extension ${}^*\mathcal{R} = ({}^*\mathbb{R}, \dots)$ of \mathcal{R} . (Such extensions exist, by basic results in elementary model theory; for example, we could take ${}^*\mathcal{R}$ to be the ultrapower of \mathcal{R} with respect to a nonprincipal ultrafilter on \mathbb{N} .)

1.1. Terminology. When we use the fact that ${}^*\mathcal{R}$ is an elementary extension of \mathcal{R} , we say that we are using the *Transfer Principle* or that we are arguing *by transfer*.

1.2. Definition (Nonstandard extensions of sets and functions).

(1) Let S be a subset of \mathbb{R}^n for some $n \geq 1$. Take P to be an n -ary predicate symbol in L such that S is the interpretation of P in \mathcal{R} . We define *S to be the interpretation of P in the elementary extension ${}^*\mathcal{R}$.

(2) Similarly, let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and take F to be an n -ary function symbol of L whose interpretation in \mathcal{R} is f . We define *f to be the interpretation of F in ${}^*\mathcal{R}$.

(3) Finally, let $S \subseteq \mathbb{R}^n$ for some $n \geq 1$ and let $g: S \rightarrow \mathbb{R}$ be a function. We define *g by taking $f: \mathbb{R}^n \rightarrow \mathbb{R}$ to be any extension of g and letting ${}^*g: {}^*S \rightarrow {}^*\mathbb{R}$ be the restriction of *f to the set *S .

Note. For the preceding definition to be correct, we must check that the specification of *S does not depend on P , that *f does not depend on F , and that *g does not depend on which extension of g we take for f . So, suppose

P_1, P_2 are two predicate symbols of L such that S is the interpretation of both P_1 and P_2 in \mathcal{R} . Then the L -sentence

$$\forall x_1 \dots \forall x_n (P_1(x_1, \dots, x_n) \leftrightarrow P_2(x_1, \dots, x_n))$$

is true in \mathcal{R} . Therefore, this sentence is true in ${}^*\mathcal{R}$, which ensures that P_1 and P_2 also have the same interpretation in ${}^*\mathcal{R}$. This shows that *S was properly defined in 1.2(1) for any set S . Similar arguments show that *f and *g were also properly defined for any functions f and g as in 1.2(2) and 1.2(3).

The structure $(\mathbb{R}, 0, 1, +, -, \times, <)$ is an ordered field, and the class of all ordered fields is axiomatizable by a finite set of L -sentences. The Transfer Principle ensures that the structure $({}^*\mathbb{R}, 0, 1, {}^*+, {}^*- , {}^*\times, {}^*<)$ is a proper ordered field extension of $(\mathbb{R}, 0, 1, +, -, \times, <)$.

For certain basic operations and relations, including $+, -, \times, <$, we will omit the $*$ even when we are working in ${}^*\mathcal{R}$; this is in keeping with normal mathematical practice. Thus we will write our extension ordered field as $({}^*\mathbb{R}, 0, 1, +, -, \times, <)$ wherever the context makes it clear that we mean to be using ${}^*\mathbb{R}$ rather than \mathbb{R} .

In $({}^*\mathbb{R}, 0, 1, +, -, \times, <)$, as in all ordered rings, we may define the absolute value function by

$$|\alpha| = \max(\alpha, -\alpha)$$

for all $\alpha \in {}^*\mathbb{R}$. We also have a nonstandard extension ${}^*|\cdot|$ of the absolute value function $|\cdot|$, as defined in 1.2(2). In fact these are the same:

$$|\alpha| = {}^*|\alpha|$$

for all $\alpha \in {}^*\mathbb{R}$. Indeed, the following sentence is true in \mathcal{R} :

$$\forall x \forall y (y = |x| \leftrightarrow (x \leq y \wedge -x \leq y \wedge (y = x \vee y = -x))).$$

By the Transfer Principle, we have that for any $\alpha, \beta \in {}^*\mathbb{R}$

$$\beta = {}^*|\alpha| \leftrightarrow (\alpha \leq \beta \wedge -\alpha \leq \beta \wedge (\beta = \alpha \vee \beta = -\alpha)).$$

This shows that ${}^*|\alpha| = |\alpha|$. For that reason, we will usually write $|\cdot|$ rather than ${}^*|\cdot|$ for the nonstandard extension of the absolute value function, just as we are doing for $+, -, \times, <$.

The following result can be used to derive essentially any formal property of our $*$ -notation that one might need. (See the Exercises at the end of

this section.) Its proof is a generalization of the arguments used in the preceding discussion.

1.3. Proposition. *Let $\varphi(x_1, \dots, x_n)$ be any $L(\mathbb{R})$ -formula and suppose $S \subseteq \mathbb{R}^n$ is the set defined by φ in \mathcal{R} ; that is,*

$$S = \{(r_1, \dots, r_n) \in \mathbb{R}^n \mid \mathcal{R} \models \varphi(r_1, \dots, r_n)\}.$$

*Then *S is the set defined by φ in ${}^*\mathcal{R}$.*

Proof. Let P be a predicate symbol of L whose interpretation in \mathcal{R} is the set S , and consider the $L(\mathbb{R})$ -sentence

$$\forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \leftrightarrow \varphi(x_1, \dots, x_n)).$$

This sentence is true in \mathcal{R} , and hence it is true in ${}^*\mathcal{R}$ by transfer; this yields the desired result. \square

Next we introduce some concepts that will be important throughout this course. An element $\alpha \in {}^*\mathbb{R}$ is *finite* if $|\alpha| < r$ for some $r \in \mathbb{R}$, and is *infinitesimal* if $|\alpha| < r$ for all positive $r \in \mathbb{R}$. We write $\alpha \approx \beta$ to mean that $\alpha - \beta$ is infinitesimal; \approx is an equivalence relation on ${}^*\mathbb{R}$. For $r \in \mathbb{R}$ we define the *monad* of r to be the set $\mu(r) = \{\alpha \in {}^*\mathbb{R} \mid \alpha \approx r\}$. The set $\text{fin}({}^*\mathbb{R}) := \{\alpha \in {}^*\mathbb{R} \mid \alpha \text{ is finite}\}$ is a proper convex subring of ${}^*\mathbb{R}$, and $\mu(0)$ is a convex maximal ideal in $\text{fin}({}^*\mathbb{R})$; it follows that $\text{fin}({}^*\mathbb{R})/\mu(0)$ is an ordered field. The order completeness of \mathbb{R} implies that $\text{fin}({}^*\mathbb{R})/\mu(0)$ is isomorphic as an ordered field to \mathbb{R} . Indeed, a homomorphism $\text{st}: \text{fin}({}^*\mathbb{R}) \rightarrow \mathbb{R}$ of ordered rings is defined by $\text{st}(\alpha) = \inf\{r \in \mathbb{R} \mid \alpha < r\}$; that is, $\text{st}(\alpha)$ is the unique real defining the same Dedekind cut in \mathbb{R} as α . Moreover $\ker(\text{st}) = \mu(0)$ and st is the identity on \mathbb{R} . We refer to $\text{st}(\alpha)$ as the *standard part* of the finite element α of ${}^*\mathbb{R}$.

In some texts, (for example in Goldblatt) finite elements of ${}^*\mathbb{R}$ are called *limited*, and the term *unlimited* is used in place of infinite (= not finite).

The set \mathbb{N} of natural numbers is a proper initial segment of ${}^*\mathbb{N}$ and $\text{fin}({}^*\mathbb{R}) \cap {}^*\mathbb{N} = \mathbb{N}$. Therefore ${}^*\mathbb{N} \setminus \mathbb{N}$ is the set of infinite elements of ${}^*\mathbb{N}$ in ${}^*\mathbb{R}$, and it is nonempty. Also, ${}^*\mathbb{N}$ is closed under the addition and multiplication inherited from the field ${}^*\mathbb{R}$. A version of the Archimedean property also holds: for any $\alpha \in {}^*\mathbb{R}$ there exists $H \in {}^*\mathbb{N}$ satisfying $\alpha < H$. For any $H \in \mathbb{N}$, we use the notation $\{0, \dots, H\}$ for the set of all $k \in {}^*\mathbb{N}$ that satisfy

$0 \leq k \leq H$. (As usual, one must be careful about the meaning of ... in mathematical expressions!)

1.4. Exercise. Prove the statements in the previous paragraphs. (For help with this see Goldblatt, pages 49-55, and Davis, pages 43-56.)

A *sequence in \mathbb{R}* is a function $s: \mathbb{N} \rightarrow \mathbb{R}$. Its nonstandard extension ${}^*s: {}^*\mathbb{N} \rightarrow {}^*\mathbb{R}$ is defined in 1.2(3).

Our first result gives a nonstandard criterion for convergence of sequences.

1.5. Proposition. *Let s be a sequence in \mathbb{R} and $r \in \mathbb{R}$. Then $\lim_{n \rightarrow \infty} s_n = r$ if and only if ${}^*s_\alpha \approx r$ for all $\alpha \in {}^*\mathbb{N} \setminus \mathbb{N}$.*

Proof. (\Rightarrow) Take any $\epsilon > 0$ in \mathbb{R} . We need to show $|{}^*s_\alpha - r| < \epsilon$ whenever $\alpha \in {}^*\mathbb{N} \setminus \mathbb{N}$. Since s converges to r , there is $k \in \mathbb{N}$ such that

$$\forall n \in \mathbb{N} (n > k \rightarrow |s_n - r| < \epsilon).$$

This can be expressed by a sentence in the language of \mathcal{R} with constants for the elements of \mathbb{R} . Hence the corresponding property holds of k in ${}^*\mathcal{R}$, by the Transfer Principle. That is, we have

$$\forall \alpha \in {}^*\mathbb{N} (\alpha > k \rightarrow |{}^*s_\alpha - r| < \epsilon).$$

In particular, $|{}^*s_\alpha - r| < \epsilon$ holds whenever $\alpha \in {}^*\mathbb{N} \setminus \mathbb{N}$.

(\Leftarrow) Take any $\epsilon > 0$ in \mathbb{R} and fix $\alpha \in {}^*\mathbb{N} \setminus \mathbb{N}$. We have that

$$\forall \beta \in {}^*\mathbb{N} (\beta > \alpha \rightarrow |{}^*s_\beta - r| < \epsilon).$$

Hence

$$\exists \alpha \in {}^*\mathbb{N} \forall \beta \in {}^*\mathbb{N} (\beta > \alpha \rightarrow |{}^*s_\beta - r| < \epsilon)$$

corresponds to a sentence true in ${}^*\mathcal{R}$. Therefore, by the Transfer Principle,

$$\exists k \in \mathbb{N} \forall n \in \mathbb{N} (n > k \rightarrow |s_n - r| < \epsilon)$$

is true in \mathcal{R} , as desired. \square

1.6. Proposition. *Let $I = (a, b) \subseteq \mathbb{R}$ and let $f: I \rightarrow \mathbb{R}$ be a function. Let $r \in \mathbb{R}$ and $c \in I$. Then $\lim_{x \rightarrow c} f(x) = r$ if and only if*

$$\forall \alpha \in {}^*I ((\alpha \approx c \wedge \alpha \neq c) \rightarrow {}^*f(\alpha) \approx r).$$

Proof. Exercise. The argument is similar to the proof of 1.5. \square

1.7. Corollary. *Let $I = (a, b) \subseteq \mathbb{R}$ and let $f: I \rightarrow \mathbb{R}$ be a function. Then f is continuous on I if and only if*

$$\forall c \in I \forall \alpha \in {}^*I (\alpha \approx c \rightarrow {}^*f(\alpha) \approx {}^*f(c)).$$

Proof. Immediate from the previous result. \square

1.8. Proposition. *Let $f: (a, b) \rightarrow \mathbb{R}$ be a continuous function and suppose $a < c < d < b$. Then, on the closed interval $[c, d]$, the function f achieves maximum and minimum values and has the intermediate value property.*

Proof. We first show that f achieves a maximum value on $[c, d]$. Let $H \in {}^*\mathbb{N} \setminus \mathbb{N}$. Set

$$c_k = c + \frac{d - c}{H}k$$

for $k \in {}^*\mathbb{N}$ in the interval $0 \leq k \leq H$. There exists k_0 such that if $k \in {}^*\mathbb{N}$ satisfies $0 \leq k \leq H$, then ${}^*f(c_k) \leq {}^*f(c_{k_0})$. This follows from the corresponding property of \mathcal{R} using the Transfer Principle. Let $e = \text{st}(c_{k_0}) \in [c, d]$. We will show that $f(e)$ is the maximum value of f on $[c, d]$. Given $x \in [c, d]$, we need to show $f(x) \leq f(e)$. The Transfer Principle ensures that there exists $k \in {}^*\mathbb{N}$ such that $0 \leq k \leq H$ and $c_k \leq x < c_{k+1}$. Now, $c_{k+1} - c_k = (d - c)/H \approx 0$, so $x \approx c_k$. By the continuity of f , we have using the previous result that $f(x) \approx {}^*f(c_k) \leq {}^*f(c_{k_0}) \approx {}^*f(e)$. Thus $f(x) \leq f(e)$, as needed.

The argument for proving that f achieves its minimum value on $[c, d]$ is similar.

To indicate how to prove the intermediate value property, assume we have the special (but typical) situation $f(c) < v < f(d)$ for some $v \in \mathbb{R}$. We want to find $u \in \mathbb{R}$ such that $c \leq u \leq d$ and $f(u) = v$. We construct the same elements c_0, \dots, c_H of ${}^*\mathbb{R}$ as above. We have ${}^*f(c_H) = f(d) > v$ and therefore, by the Transfer Principle, there exists a least $k \in \{0, \dots, H\}$ such that ${}^*f(c_k) > v$. For this value of k we have ${}^*f(c_k) > v \geq {}^*f(c_{k-1})$. Setting $u = \text{st}(c_k)$, we have $f(u) \geq v \geq f(u)$, since $c_k \approx c_{k-1} \approx u$ and f is continuous at u . \square

Next we indicate some ways in which nonstandard analysis interacts with elementary topology.

1.9. Proposition. *Let $A \subseteq \mathbb{R}$ and $c \in \mathbb{R}$. Then*

- (1) c is in the interior of $A \Leftrightarrow \mu(c) \subseteq {}^*A$;
 (2) c is in the closure of $A \Leftrightarrow \mu(c) \cap {}^*A \neq \emptyset$;
 (3) c is a limit point of $A \Leftrightarrow (\mu(c) \setminus \{c\}) \cap {}^*A \neq \emptyset$.

Proof. We only prove (1) and leave the others as exercises.

(\Rightarrow) If c is in the interior of A , then there exists an $\epsilon > 0$ in \mathbb{R} such that the interval $(c - \epsilon, c + \epsilon)$ is contained in A . Passing to ${}^*\mathbb{R}$, we get that $\mu(c) \subseteq {}^*(c - \epsilon, c + \epsilon) \subseteq {}^*A$.

(\Leftarrow) Assume $\mu(c) \subseteq {}^*A$ and take any $\delta \in {}^*\mathbb{R}$ with $\delta > 0$ and $\delta \approx 0$. Then ${}^*(c - \delta, c + \delta) \subseteq \mu(c) \subseteq {}^*A$. Thus we have

$$\exists \delta \in {}^*\mathbb{R} (\delta > 0 \wedge {}^*(c - \delta, c + \delta) \subseteq {}^*A).$$

Passing to \mathcal{R} by transfer we get

$$\exists \delta \in \mathbb{R} (\delta > 0 \wedge (c - \delta, c + \delta) \subseteq A),$$

which implies that c is in the interior of A . \square

1.10. Proposition. *Let $A \subseteq \mathbb{R}$ be a nonempty set. Then A is compact if and only if for every $\alpha \in {}^*A$ there exists $c \in A$ such that $\alpha \approx c$.*

Proof. (\Rightarrow) Assume A is compact. Arguing by contradiction, suppose there is $\alpha \in {}^*A$ such that for each $c \in A$ there exists $\delta_c > 0$ in \mathbb{R} such that $|c - \alpha| > \delta_c$. For each $c \in A$, set $I_c = (c - \delta_c, c + \delta_c)$. It is clear that $\{I_c \mid c \in A\}$ covers A . Since A is compact, there exist $c_1, \dots, c_k \in A$ with $A \subseteq \cup_{j=1}^k I_{c_j}$. Therefore $\alpha \in {}^*A \subseteq \cup_{j=1}^k {}^*I_{c_j}$, which is a contradiction.

(\Leftarrow) We prove the contrapositive. So, assume that A is not compact. Then there exists a collection of rational intervals $\{(p_n, q_n) \mid n \in \mathbb{N}\}$ that covers A , but such that no finite subcollection of it covers A . For each $k \in \mathbb{N}$, let c_k be in $A \setminus \cup_{n=1}^k (p_n, q_n)$. In \mathcal{R} we have

$$\forall k \in \mathbb{N} \forall n \in \mathbb{N} \left(n \leq k \rightarrow (c_k \in A \wedge c_k \notin (p_n, q_n)) \right).$$

Transfer this statement to ${}^*\mathcal{R}$ and take $K \in {}^*\mathbb{N} \setminus \mathbb{N}$ and set $\alpha = {}^*c_K$. Then $\alpha \in {}^*A$ and $\alpha \notin {}^*(p_n, q_n)$ holds for all $n \in \mathbb{N}$. Note that for any $c \in \mathbb{R}$, if $c \in (p_n, q_n)$, then $\mu(c)$ is contained in ${}^*(p_n, q_n)$. Since $\{(p_n, q_n) \mid n \in \mathbb{N}\}$ covers A , it follows that there does not exist any $c \in A$ such that $\alpha \approx c$. \square

1.11. Definition. A *norm on \mathbb{R}^n* is a function $N: \mathbb{R}^n \rightarrow \mathbb{R}$ such that for all $x, y \in \mathbb{R}^n$ and $r \in \mathbb{R}$

- $N(x) \geq 0$;
- $N(x) = 0 \Leftrightarrow x = 0$;
- $N(rx) = |r|N(x)$;
- $N(x + y) \leq N(x) + N(y)$.

1.12. Proposition. *Let $N: \mathbb{R}^n \rightarrow \mathbb{R}$ be a norm. If $x = (x_1, \dots, x_n) \in {}^*\mathbb{R}^n$, then ${}^*N(x) \approx 0$ if and only if $x_i \approx 0$ for all $i = 1, \dots, n$.*

Proof. (\Leftarrow) Let $x = (x_1, \dots, x_n) \in {}^*\mathbb{R}^n$ be such that $x_i \approx 0$ for all $i = 1, \dots, n$. For $j = 1, \dots, n$, let $e_j = (0, \dots, 0, 1, 0, \dots, 0)$. (This is the element of \mathbb{R}^n that has the entry 1 in position j and entry 0 in all other positions.) Using vector space operations in ${}^*\mathbb{R}^n$ we have $x = \sum_{j=1}^n x_j e_j$, and hence

$$0 \leq {}^*N(x) = {}^*N\left(\sum_{j=1}^n x_j e_j\right) \leq \sum_{j=1}^n |x_j| N(e_j) \approx 0$$

since $|x_j| \approx 0$ and $N(e_j)$ is finite for each j . Thus ${}^*N(x) \approx 0$.

(\Rightarrow) Suppose there is $x \in {}^*\mathbb{R}^n$ with ${}^*N(x) \approx 0$ but such that some x_j is not infinitesimal. Assume for simplicity that $|x_1| \geq |x_j|$ for all $j = 1, \dots, n$, so we have that $x_1 \notin \mu(0)$ and thus $\frac{1}{|x_1|}$ is finite. We have

$$0 \approx \frac{1}{|x_1|} {}^*N(x) = {}^*N\left(\frac{1}{x_1} x\right).$$

Moreover, $\frac{1}{x_1} x = (1, x'_2, \dots, x'_n)$ where $|x'_j| = \frac{|x_j|}{|x_1|} \leq 1$ for $2 \leq j \leq n$. Let $r_j = \text{st}(x'_j)$, for $j = 1, \dots, n$. We then have

$${}^*N\left(\frac{1}{x_1} x - (r_1, \dots, r_n)\right) = {}^*N((0, x'_2 - r_2, \dots, x'_n - r_n)) \approx 0.$$

By the triangle inequality,

$$0 \approx {}^*N\left(\frac{1}{x_1} x\right) \approx N((1, r_2, \dots, r_n)),$$

but this is a contradiction, since N is a norm and the standard vector $(1, r_2, \dots, r_n)$ is not 0. \square

1.13. Corollary. *If N_1, N_2 are two norms on \mathbb{R}^n , they define the same topology on \mathbb{R}^n . Indeed, there exist constants $c, d > 0$ so that for any $a \in \mathbb{R}^n$ one has $cN_1(a) \leq N_2(a) \leq dN_1(a)$.*

Proof. Consider $S = \{a \in \mathbb{R}^n \mid N_1(a) \leq 1\}$. To prove the existence of d it suffices to show that N_2 is bounded above on S . For this it suffices to show that *N_2 is bounded above on *S by some element of ${}^*\mathbb{R}$, by the Transfer Principle. We can do this by showing that ${}^*N_2(x)$ is finite for each $x \in {}^*S$, since from this it follows that *N_2 is bounded above on *S by any positive infinite number. Arguing by contradiction, suppose there exists x with ${}^*N_1(x) \leq 1$ but with ${}^*N_2(x) = \alpha$ infinite. Then

$${}^*N_1\left(\frac{1}{\alpha}x\right) \leq \frac{1}{\alpha} \approx 0,$$

but

$${}^*N_2\left(\frac{1}{\alpha}x\right) = \frac{1}{\alpha}\alpha = 1,$$

which is a contradiction. \square

1.14. Exercises. (1) Let $S \subseteq \mathbb{R}^n$, $g: S \rightarrow \mathbb{R}$, and let $\Gamma_g \subseteq \mathbb{R}^{n+1}$ be the graph of g , that is

$$\Gamma_g = \{(x_1, \dots, x_n, y) \in \mathbb{R}^{n+1} \mid (x_1, \dots, x_n) \in S \wedge y = g(x_1, \dots, x_n)\}.$$

Show that ${}^*(\Gamma_g)$ is the graph of *g and that ${}^*(\text{range}(g))$ is the range of *g . Show that *S contains S and that *g agrees with g on S .

(2) Show that when it is applied to subsets of \mathbb{R}^n , the map taking S to *S is an embedding of Boolean algebras of sets, and that it is the identity map on finite subsets of \mathbb{R}^n .

(3) Show that there is no $L({}^*\mathbb{R})$ -formula that defines $\mu(0)$ in the structure ${}^*\mathcal{R}$. Likewise for the sets $\text{fin}({}^*\mathbb{R})$ and ${}^*\mathbb{N} \setminus \mathbb{N}$.

(4) Let s be a sequence in \mathbb{R} . Show that s is a Cauchy sequence if and only if ${}^*s_\alpha \approx {}^*s_\beta$ holds for all infinite $\alpha, \beta \in {}^*\mathbb{N}$.

(5) Let $I = (a, b) \subseteq \mathbb{R}$ and let $f: I \rightarrow \mathbb{R}$ be any function. Show that f is uniformly continuous on I if and only if

$$\forall \alpha, \beta \in {}^*I \ (\alpha \approx \beta \rightarrow {}^*f(\alpha) \approx {}^*f(\beta)).$$

Textbooks referenced on the course information sheet also contain many helpful exercises that are at the level of nonstandard analysis that is discussed in this section. For example, see Hurd-Loeb pages 24, 28, 31, 38–39, 44, and 49–50, as well as the material discussed in pages 20–49 of that book.