Nonstandard Analysis in Dynamical systems

O. Knill
Department of Mathematics,
California Institute of Technology, 253-37,
Pasadena, CA, 91125 USA.

1995

Abstract

We apply the language of nonstandard analysis (IST) to topological dynamics. As an illustration, we translate the proof of Furstenbergs multiple recurrence theorem into the framework of nonstandard analysis. This proof had been presented in a Specker seminar around 1990, this document was TeXed up in 1995 while at Caltech.

1 Internal set theory

There exist two major versions of nonstandard analysis.

The first and older one by Abraham Robinson was developed in 1960 at the Institute for Advanced Study in Princeton [8]. The basic idea is to embed every mathematical structure into an extension while inheriting as much as possible from the properties of the old structures. Robinson realized in a rigorous way Leibniz's dream of a mathematical theory of infinitesimals. The D'Alembert-Cauchy' concept of limit and Weierstrass cleansing lead the "epsilontic" used in real analysis courses. Robinson's theory revived "ghosts of departed quantities summoned by infinitesimals" by using partially ordered non-Archimedean ring extensions of the reals to have elements which behave like infinitesimals. It was motivated also work of T. Skolem who had constructed extensions of the natural numbers.

A newer, different and more friendly approach from Ed Nelson [5] is called *Internal Set Theory* or shortly IST. Internal set theory takes the common mathematical structures and distinguishes inside this structures so called "standard elements". A theory called ISTE which combines both approaches of Robinson and Nelson [4] has also been considered. Almost everything is this introduction paragraph can be found in [5] or the exposition [7]. We should also mention [6] in which the theory of stochastic processes is developed using finite sets only. We will here use Nelson's internal set theory.

The axioms of IST consist of the usual Zermelo-Fraenkel axiom system ZFC (the letter C in that acronym stands for the axiom of choice which is usually included) together with three additional axioms which tell, how to use the new unitary predicate *standard*.

A formula in IST is called *internal* if it does not involve the new predicate, otherwise, it is called *external*.

The three axioms are the *transfer principle (T)*, the *principle of idealisation (I)* and the *principle of standarisation (S)*. We use the short notation $\forall^{st} x$ for saying: "for all standard x" and $\forall^{fin} x$ instead of "for all finite x".

$$
\begin{array}{lcl}
(I) & : & (\forall^{st,fin} z \exists x \forall y \in z \ \phi(x,y)) \Leftrightarrow (\exists x \forall^{st} y \ \phi(x,y)), \qquad \phi \ internal \\
(S) & : & \forall^{st} x \exists^{st} b \forall^{st} x \ (x \in b \Leftrightarrow x \in a \ and \ \phi(x)), \qquad \phi \ arbitrary \\
(T) & : & (\forall^{st} x \ \phi(x,u)) \Leftrightarrow (\forall x \ \phi(x,u)), \qquad \phi \ internal, \quad u \ standard
\end{array}
$$

From the idealisation principle (I) one can deduce: (see [6] Theorem 1.2 or [7] 1.3.2):

Proposition 1.1 (Application of (I)) *For any set X, there is a finite set E, which contains all standard elements of X.*

Remarks.
1) A set X is a standard finite set if and only if every element of E is standard.
2) The finite E containing all standard elements of X is nonstandard if and only if X is infinite.
3) There is no smallest infinite set.
4) Every infinite set contains nonstandard elements.

The standardization principle (S) can be used to define sets. One denotes the unique set b in (I) with ${}^S\{x \in a \mid \phi(x)\}$ and read it as 'the standard subset of the set a whose standard elements are those which satisfy ϕ'.

It is important that the criterion for set membership applies only for standard elements. From the standardisation principle, one can deduce

Proposition 1.2 (Application of (S)) *Given standard sets X, Y and a formula ϕ with two free variables. Assume, there exists for all standard $x \in X$ a standard $y \in Y$ such that $\phi(x, y)$. Then there is a standard function $\psi : X \to Y$ such that for all standard x we have $\phi(x, \psi(x))$.*

The transfer principle (T) finally is used to translate between conventional mathematics and nonstandard analysis.

Proposition 1.3 (Application of (T)) *All theorems of conventional mathematics also hold when relativized to standard sets. Every set in conventional mathematics is a standard set.*

Before applying transfer to an assertion, one must verify that it is internal and that all the parameters in it have standard values.

A real number x is called *infinitesimal* if $|x| \leq \epsilon$ for all standard $\epsilon > 0$. It is called *limited* if there exists a standard r such that $|x| \leq r$ otherwise it is called *unlimited*.

Let (X, d) be a standard metric space. Two points $x, y \in X$ are *infinitely close*, denoted by $x \simeq y$, if $d(x, y)$ is infinitesimal.

Given a sequence x_n in a standard metric space X. We say, a_n *S-converges* to $x \in X$ if $x_n \simeq x$ for all unlimited $n \in \mathbf{N}$. Given two standard metric spaces X and Y, A function $f : X \to Y$ is called *S-continuous*, if $\forall^{st} x \in X \; x \simeq y \to f(x) \simeq f(y)$. Here are some facts:

- A standard function $f : X \to Y$ is continuous, if and only if it is S-continuous.
- A standard metric space X is compact if and only if every $x \in X$ is infinitely close to a standard element $y \in X$.
- If X is compact, there exists for each $x \in X$ a unique standard x° which is infinitely close to x.
- A standard sequence x_n of a standard topological space X converges to an element $x \in X$ if and only if it is S-converging to $x \in X$.
- A standard sequence x_n in a standard topological space X has an accumulation point x if there exists an unlimited $n \in \mathbf{N}$ with $x_n \simeq x$.

2 Birkhoff's multiple recurrence theorem

A topological dynamical system (X, G) is a group G acting by homeomorphisms on the compact metric space (X, d). If $G = \mathbf{Z} = \langle T \rangle$ is generated by one map T, the dynamical system is called *cyclic* and we write (X, T). If $G = \mathbf{Z}^l = \langle T_1, \ldots T_l \rangle$ we have a **finitely generated system**. See [2] for terminology in topological dynamics.

Here are some new notions

- We use the notation $x \simeq_K y$ for $K \cdot d(x, y) \simeq 0$.
- We say $T : X \to X$ is called **K-continuous** if for all $x, y \in X, x \simeq_K y \Rightarrow Tx \simeq Ty$.
- (X, G) is called *minimal* if and only if for all $x, y \in X$, there exists $g \in G$ such that $gx \simeq y$.

Lemma 2.1 *If (X, G) is minimal, there exists a finite subset G_0 of G such that for all $x, y \in X$, $\exists g \in G_0$ $T_g x \simeq y$.*

Proof. The compactness of X is equivalent with: for all $\epsilon > 0$ there exists a finite open cover $X = \bigcup_{i=1}^{n} U_i$ with $\text{diam}(U_i) \le \epsilon/2$. The minimality implies that for all (i, j) there exists $g_{ij} \in G$ such that $g_{ij} U_i \cap U_j \neq \emptyset$. The set $G(\epsilon) = \{g_{ij} | i, j = 1, \ldots n\}$ is finite and $\forall x, y \in X$ $\exists g \in G(\epsilon)$ $d(T_g x, y) \le \epsilon$. Define $G_0 = G(\epsilon_0)$, with infinitesimal ϵ_0. □

Of course, this finite subset G_0 is not standard if X is not a finite standard set.

Lemma 2.2 *Given a dynamical system (X, G). For all $g \in G$, there exists $K > 0$ such that T_g is K-continuous.*

Proof. T_g is continuous. Let $h \mapsto K(h)$ be the continuity module of T_g (it exists because of the compactness of X). This means $d(Tx, Ty) \le K(h) \cdot d(x, y)$ if $d(x, y) \le h$. Take h infinitesimal and $K = K(h)$. □

Lemma 2.3 *Given a compact set X and $K > 0$. There exists a finite set $E \subset X$ such that for all $x \in X$ there exists $e \in E$ with $x \simeq_K e$.*

Proof. The compactness of X is equivalent to the claim with K standard, because $x \simeq y \Leftrightarrow x \simeq_K y$. Use transfer. □

For a cyclic system (X,T), a point $x \in X$ is called *recurrent* if there exists $n \geq 1$ with $T^n x \simeq x$. If $(X, \mathbf{Z}^l)$ is a finitely generated dynamical system, a point $x \in X$ is called *simultaneously recurrent*, if there exists $n \geq 1$ such that $T_i^n x \simeq x$ for all $i = 1, \dots l$.

Theorem 2.4 (Birkhoff's recurrence theorem) *For every cyclic system* (X,T) *there exists a recurrent* $x \in X$*.*

Proof. Take $y \in X$ arbitrarily and a finite set $E \subset X$ which contains all the standard elements of X. For all k there exists $e_k \in E$ with $T^k y \simeq e_k \in E$. Because E is finite, there exists $k < m$ such that $T^k y \simeq T^m y$. Define $x := T^k y$ and $n := m - k$. Then $T^n x = T^{m-k} T^k y = T^m y \simeq T^k y = x$. □

Remark. A point $x \in X$ is periodic if and only if there exists a standard $n \in \mathbf{N}$ such that $T^n x \simeq x$. The only difference between recurrent and periodic orbits is that in the periodic case, we have a standard natural number n such that $T^n x \simeq x$ and in the recurrents case, we have a nonstandard natural number n such that $T^n x \simeq x$. We can think therefore of recurrent points as some sort of "nonstandard periodic points".

If (X,T) is cyclic, a subset A of X is called **homogeneous** if there exists (A,G) minimal, such that for all $g \in G$ $T_g T = T T_g$ and $T_g A = A$.

Lemma 2.5 (Main lemma) *Assume* (X,T) *is cyclic and* A *is homogeneous for* T*. Assume there exists* $x_0, y_0 \in A$ *and* $n \geq 1$ *such that* $T^n y_0 \simeq x_0$*. Then there exists* $z \in A$ *and* $n \geq 1$ *such that* $T^n z \simeq z$*.*

Proof. Consider the three statements
(1) $\exists x_0, y_0 \in A \; \exists n \geq 1 \; T^n y_0 \simeq x_0$
(2) $\forall x \in A \; \exists y \in A \; \exists n \geq 1 \; T^n y \simeq x$.
(3) $\exists z \in A \; \exists n \geq 1 \; T^n z \simeq z$
For every $K > 0$ the statements (1),(2),(3) are equivalent to the statements where $\simeq$ is replaced by $\simeq_K$, since all statements are equivalent to internal statements. We have to show (1) $\Rightarrow$ (3).

(1) $\Rightarrow$ (2):
Take a finite set $G_0 \subset G$ such that $\forall x, y \in X$ there exits $g \in G_0$ with $T_g x \simeq y$. Choose $K_0 > 0$ such that $\forall g \in G_0$ the map T_g is K_0-continuous. Take a finite set $E \subset X$ such that $\forall x \in X$ there exists $e \in E$ with $x \simeq_K e$.
Let $T^n y_0 \simeq_{K_0} x_0$. Given $x \in X$. There exists $g \in G_0$ such that $T_g x_0 \simeq x$. Because T_g is K_0-continuous, it follows that $T_g T^n y_0 \simeq T_g x_0$. The homogeneity allows to pick $y := T_g y_0$ with $y \in A$. Then $T^n y = T^n T_g y_0 = T_g T^n y_0 \simeq T_g x_0 \simeq x$.

(2) $\Rightarrow$ (3): Take $x_0 \in A$ and $\epsilon \simeq 0$ and form inductively the sequence $x_i \in A$ satisfying $\exists n_i \geq 1 d(T^{n_i} x_i, x_{i-1}) \leq \epsilon/2^i$
Since X is compact, there exists $k < m$ such that $x_k \simeq x_m$. Define $n = n_{k+1} + \cdots n_{m-1} + n_m$. Then $d(T^n x_k, x_m) \leq \sum_{i=m+1}^{k} d(T^{n_i} x_i, x_{i-1}) \leq \sum_{i=1}^{m-k} 2^{-i}\epsilon \leq 2\epsilon$ so that $T^n x_k \simeq x_m \sim x_k$. □

Theorem 2.6 (Multiple recurrence theorem) *For any finitely generated system* $(X, \mathbf{Z}^l)$*, there exists a multiple recurrent* $x \in X$.

Proof. We can assume without loss of generality that (X, G) is minimal because we can restrict us else to a closed minimal G-invariant set. The proof goes with induction to l: For $l = 1$ it is Birkhoff's recurrence theorem. For the induction step $l - 1 \mapsto l$ consider the system $(X^l, T = T_1 \times T_2 \cdots \times T_l)$ and the diagonal $A = \{(x, x, \cdots x) \mid x \in X\}$ which is homogeneous with respect to G. An element $T_i \in G$ acts on the product X^l by $(T_i \times, \dots, \times T_i)$. From the main-lemma we see that it is enough to find $x_0, y_0 \in A$ and $n \geq 1$, so that $T^n y_0 \simeq x_0$. Define for $i = 1, \dots l - 1$ the transformation $R_i = T_i T_l^{-1}$. The induction assumption gives $\exists z \in X \ \exists n \geq 1 \ \ T_i^n z \simeq z, i = 1, \dots, l - 1$. Define $x_0 = (z, z, \dots z)$ and $y_0 = (T_l^{-n} z, \dots, T_l^{-n} z) \in A$. Then $T^n y_0 = (T_1^n T_l^n z, T_2^n T_l^n z, \dots, z) = (R_1^n z, \dots R_{l-1}^n z, z) \simeq (z, z, \dots z) = x_0$. □

Corollary 2.7 *Given* (X, T) *cyclic,* $x \in X$ *and* $l \in \mathbf{N}$*. There exists* $m \in \mathbf{Z}, l \geq 1$ *such that* $T^m x, \ \ T^{m+n} x, \dots T^{m+ln} x$ *have pairwise infinitesimal distance.*

Proof. The set $Y := \overline{\{T^k x\}}$ is compact and invariant under $T_1 = T, T_2 = T^2, \dots T_l = T^l$. The multiple recurrence theorem gives $\exists y \in Y \ \exists n \geq 1 \ T_i^n y \simeq y, \ i = 1, \dots, l$. There exists $K > 0$ such that T_i^n are K-continuous for $i = 1, \dots l$. The orbit of x is dense. Therefore $\exists z = T^m x \simeq_K y$. We have so $T_i^n z = T^{m+in} x \simeq T^{in} y \simeq y \simeq z$. □

3 More notions in topological dynamics

(X, G) always denotes a topological dynamical system.

- **Topological transitivity**. There exists $x \in X$ such that the orbit Gx is dense in X.

  Nonstandard formulation. There exists $x \in X$ such that for all $y \in X$ there exists $g \in G$ with $T_g x \simeq y$.

- **Expansivity**. There exists $\epsilon > 0$ such that for any pair $x \neq y$, there exists $g \in G$ with $d(T_g x, T_g y) \geq \epsilon$.

  Nonstandard formulation. For any x, and all $x \neq y$, $(x \simeq y)$, there exists $g \in G$ with $T_g x \ \not{s}imeq T_g y$.

- **Sensitive dependence on initial conditions**. There exists $\epsilon > 0$ such that for all x and any neighborhood U of x, there exists $y \in U$ and $g \in G$ with $d(T_g x, T_g y) \geq \epsilon$.

  Nonstandard formulation. For all $x \in X$, there exists $y \simeq x$, and $g \in G$ where $T_g x \simeq T_g y$ fails.

  Remark. A periodic system is expansive but has no sensitive dependence on initial conditions. If X has no isolated points, then expansivity is stronger than sensitive dependence on initial conditions.

- **Periodic orbits are dense**. For all x and $\epsilon > 0$ there exists $y \in X$ and $g \in G$ such that $T_g y = y$ and $d(x, y) < \epsilon$.

  Nonstandard formulation. For all $x \in X$ there exists $g \in G$ and $y \simeq x$ such that $T_g y = y$.

- **Chaotic in the sense of Deveney**. (X, G) is *chaotic in the sense of Deveney*, if it is topological transitive, if periodic orbits are dense and if the system has sensitive dependence on initial conditions.

  Remark. If (X, G) is cyclic, topological transitivity and the density of periodic orbits implies sensitive dependence on initial conditions.

4 The theorem of van der Waerden

Given an arbitrary partition of the entire numbers $\mathbf{Z} = A_1 \cup A_2 \cdots A_q$. Does there exists an A_i, in which there exists arbitrary long arithmetic progressions ?
In 1927, B.L van der Waerden solved this so called Baudet conjecture which had occupied for a long time a substantial number of mathematicians of that time. His proof was combinatorial and not easy. A considerably simpler proof was found by the Russian mathematician M.A. Lukomskaja. Her still elementary proof was published in the beautiful little book of Chinchin [1].

Furstenberg [3] gave a very simple proof of van der Waerden's theorem by reformulating the question as a problem in dynamical systems. In this manner, the result of van der Waerden becomes a direct corollary of a **multiple Birkhoff recurrence theorem**, which tells us that there exists a simultaneous accumulation point for a finite number of continuous commuting transformations on a compact metric space. Using the language of nonstandard analysis, the proof of Birkhoff's multiple recurrence theorem gets simpler. The basic idea is to embed every mathematical structure into an extension while inheriting as much as possible from the properties of the old structures.

Corollary 4.1 (Theorem of van der Waerden) *Given a partition* $\mathbf{Z} = A_1 \cup A_2 \cup \ldots \cup A_q$*. One of the sets* A_i *contains arbitrarily long arithmetic sequences.*

Proof. We show that for all $l \geq 1$ there exists A_j, which contains an arithmetic sequence $\{m, m+n, m+2n, \ldots m+ln\}$. One of the sets A_j occurs then infinitely often. Take the shift dynamical system (X, T) where T is the shift on the compact metric space $X = \{1, \ldots, q\}^{\mathbf{Z}}$. The points of X are in one to one correspondence to the partitions of $\mathbf{Z}$ in q sets: $n \in A_i \Leftrightarrow x_n = i$. Given $x \in X$ and $l \geq 1$. From the above Corollary follows that there exists $m \in \mathbf{Z}$ and $n \geq 1$, such that $d(T^m x, T^{m+n} x, \ldots T^{m+ln} x$ have pairwise the distance < 1. It follows that $x_m = x_{m+n} = x_{m+2n} = \ldots xm + ln$. If $x_m = j$, the sequence $\{m, m+n, m+2n, \ldots m+ln\}$ is in A_j. □

Equipped with the multiple recurrence theorem, it is possible to prove also a multi-dimensional version of the theorem of van der Waerden, which was found by Gruenwald. The proof presented here is also from Furstenberg:

Theorem 4.2 (Gruenwald) *Given a finite partition* $A_1, A_2, \cdots A_q$ *of* $G = \mathbf{Z}^n$*. There exists* A_j *such that for each finite set* $F \subset G$ *there exists* $a \in G$ *and* $b \in \mathbf{Z}$ *such that* $a + bF \subset B_j$*.*

Proof. Write $F = \{f_1, \ldots, f_l\}$. Take as before the set $\Lambda = \{1, \cdots q\}$ and form the compact metric space $\Omega = \lambda^G$ with metric $d(v, w) = inf\{(k+1)^{-1} | \; v_i = w_i$ for $|i_s| < k, s = 1, \cdots n\}$. There is again a bijective correspondence between Ω and partitions of G with q elements. Define the transformations $T_i : w_n \mapsto w_{n+f_i}$. Given a point $w \in \Omega$. We denote with X the smallest closed set in Ω which contains w and is invariant under the transformations T_i. The multiple Birkhoff recurrence theorem gives a point $x \in X$ and $b \in \mathbf{Z}$ such that $d(T_i^b x, x) < 1$ for $i = 1, \cdots l$. This means, that $x_0 = x_{b \cdot e_i}$ and we have therefore an $a \in G$ such that $w_a = w_{a+b_i}$. □

Appendix

Added, 1995 while TeXing this up: I learned about Nonstandard analysis in Logic courses by Ernst Specker and Peter Läuchli at ETH. By coincidence Specker happened also to be my linear algebra teacher for a two semester course while Läuchli was my calc teacher for a two semester course.

Läuchli later offered a course entirely devoted to nonstandard analysis. He started the course with "Hocus Pocus". I was in awe. Many things in calculus would just become simpler. I read Nelsons article and adored the book of Alain Robert. I had shown the above proof of Weierstrass theorem to Jürgen Moser, my undergraduate advisor and he discouraged me to pursue reproving things which are known with new methods and instead focus on problems which are unsolved with traditional methods. I think it was good advise as non-standard analysis is a rather niche topic. The reason for the reservation of many mathematicians in approaches like nonstandard analysis is that there are not many people who can or are willing to read the language. Talking a language which nobody understands can be especially bad, when starting a career as a mathematician.

Nevertheless, the language of nonstandard analysis helped me while teaching, especially to students who do not want to become mathematicians. It also got kept me think a bit like combinatorics person. Nonstandard analysis is how Euler thought. Instead of the $\epsilon - \delta$ definition of continuity, one says "If x is infinitesimally close to y then $f(x)$ is infinitesimally close to $f(y)$". This is a formulation which is intuitive without a formal axiomatic definition, what "infinitesimally close" means and Euler was very successful with that.

References

[1] A.J. Chinchin. *Drei Perlen der Zahlentheorie.* Harry Deutsch, 1984 (reprint).

[2] M. Denker, C. Grillenberger, and K. Sigmund. *Ergodic Theory on Compact Spaces.* Lecture Notes in Mathematics 527. Springer, 1976.

[3] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton University Press, Princeton, N.J., 1981. M. B. Porter Lectures.

[4] R. Lutz and M. Goze. *Nonstandard Analysis*, volume 881 of *Lecture Notes in mathematics.* Springer, 1981.

[5] E. Nelson. Internal set theory: A new approach to nonstandard analysis. *Bull. Amer. Math. Soc*, 83:1165–1198, 1977.

[6] E. Nelson. *Radically elementary probability theory.* Princeton university text, 1987.

[7] A. Robert. *Analyse non standard.* Presses polytechniques romandes, 1985.

[8] A. Robinson. *Non-standard Analysis.* Princeton University Press, 1960, reprint 1994.