in READ, R. (ed.), *'Graph Theory and Computing'*, Academic Press, 1973.

19. ROSE, D.J. and TARJAN, R.E.
"Algorithmic Aspects of Vertex Elimination on Directed Graphs", *SIAM Journal on Applied Mathematics*, Vol. 34 (1978) 176-197.

20. ROSE, D.J. and WILLOUGHBY, R.A. (Eds)
'Sparse Matrices and their Applications', Plenum Press, 1972.

21. TARJAN, R.E.
"Graph Theory and Gaussian Elimination", in [2] (1976) 3-22.

22. VARGA, R.S.
'Matrix Iterative Analysis', Prentice-Hall, 1962.

23. WELSH, J., ELDER, J. and BUSTARD, D.
'Sequential Program Structures', Prentice-Hall, 1984.

24. WILLOUGHBY, R. (Ed.)
'Sparse Matrix Proceedings', IBM Research, 1968.

Management Information Systems Department,
University College, Dublin

AN INTRODUCTION TO NONSTANDARD ANALYSIS

Christopher Thompson

1. Introduction

This note describes the axiomatic approach to nonstandard analysis developed by Nelson and illustrates it by proving the fundamental theorem of algebra and a form of the spectral theorem in finite dimensions. It is based in part on a talk given at the DIAS in April, 1985.

2. The Axioms

We shall be working in a mathematical universe that contains all the usual familiar objects (e.g. numbers 0, 1, $\sqrt{2}$, π etc., sets $\mathbb{N}$, $\mathbb{R}$, $\mathbb{C}$ etc., function spaces ℓ^2, $C[0,1]$ etc.) and in addition contains new and unfamiliar objects such as infinitely large natural numbers and infinitely small positive real numbers. To ensure the presence of the familiar objects we adopt the usual axioms of set theory, for example the Zermelo-Fraenkel axioms together with the axiom of choice.
To make visible the unfamiliar objects we adopt a new undefined unary predicate *standard* and axioms (I), (S) and (T) to govern its use. The resulting theory is called *internal set theory* (IST) and is due to Nelson [7]. Just as the binary predicate "ε" of ZFC has the informal interpretation "is a member of" (although strictly speaking it is undefined and therefore meaningless), so also has the unary predicate "standard" of IST an informal meaning: 'x standard' has the interpretation 'x is a familiar object of classical mathematics'. It is a consequence of the axioms that 0, 1, $\sqrt{2}$, π, $\mathbb{N}$, $\mathbb{R}$ etc. are indeed standard, as we shall see.

A formula of IST may or may not contain the predicate "standard". If it does then it is called an *external* formula, otherwise it is called *internal*. Speaking informally, internal formulas are those that make sense to a classical mathematician. We shall use a system of abbreviations illustrated by

$$\forall^{st}x \quad \text{for} \quad \forall x\ (x \text{ standard}) \Rightarrow$$

$$\exists^{st}x \quad \text{for} \quad \exists x\ (x \text{ standard}) \wedge$$

$$\forall^{fin}x \quad \text{for} \quad \forall x\ (x \text{ finite}) \Rightarrow .$$

Here, "x finite" means that there is a bijection of x with $\{m \in \mathbb{N} : m < n\}$ for some natural number n.

The axioms of IST are the axioms of ZFC and three new axioms called *transfer* (T), *idealisation* (I) and *standardization* (S).

TRANSFER. *Let* A(x) *be an internal formula with free variable* x *and no other free variable:* A(x) *may contain constants but they must be standard. Then*

$$\text{(T)} \qquad \forall^{st}x\ A(x) \Rightarrow \forall x\ A(x).$$

The transfer axiom implies that all "classical" objects are standard. To see this, observe that (T) is equivalent to

$$\exists x\ A(x) \Rightarrow \exists^{st}x\ A(x). \qquad (1)$$

Consequently if there is a unique x such that A(x) then that x must be standard. In particular, taking A(x) to be "x is a complete ordered field", we deduce that $\mathbb{R}$ is standard. In a similar way any uniquely specified classical object such as 0, 1, $\sqrt{2}$, $\mathbb{N}$, $\mathbb{C}$ is standard.

IDEALISATION. *Let* B(x,y) *be an internal formula with free variables* x,y *and possibly other free variables. Then*

$$\text{(I)} \quad \forall^{st\,fin}z\ \exists x\ \forall y \in z\ \ B(x,y) \Leftrightarrow \exists x\ \forall^{st}y\ \ B(x,y).$$

The idealisation axiom implies the existence of infinitely large natural numbers and non-zero infinitesimal real numbers. To see this we need some definitions. If $x \in \mathbb{R}$ or $\mathbb{C}$ then

x is *infinitesimal* ($x \simeq 0$) if $\forall^{st}n \in \mathbb{N}\ |x| < 1/n$,

x is *unlimited* ($|x| \simeq +\infty$) if $\forall^{st}n \in \mathbb{N}\ |x| > n$,

x is *limited* ($|x| \ll +\infty$) if $\exists^{st}n \in \mathbb{N}\ \ |x| \le n$.

We call x and y infinitely close (and write $x \simeq y$) if $|x-y|$ is infinitesimal. If we let B(x,y) be $x \in \mathbb{N} \wedge y \in \mathbb{N} \wedge y < x$ then from (I) we deduce that

$$\exists x \in \mathbb{N}\ \forall^{st}y \in \mathbb{N}\ \ y < x.$$

This says that $\mathbb{N}$ has an unlimited element. Let us fix on one such element and call it ω. Then $\omega \in \mathbb{R}$ and therefore $1/\omega \in \mathbb{R}$. It is easy to see that $1/\omega$ is positive and infinitesimal.

STANDARDISATION. *Let* C(z) *be a formula, internal or external, with free variable* z *and possibly other free variables. Then*

$$\text{(S)} \qquad \forall^{st}x\ \exists^{st}y\ \forall^{st}z \quad (z \in y \Leftrightarrow z \in x \wedge C(z)).$$

In words: given any standard set x and any property C, there is a standard set y whose standard elements are exactly those standard elements of x that satisfy C. The need for this axiom arises because in IST it is illegal to use an external formula C to form a new set y from a given set x by letting $y = \{z \in x : C(z)\}$; for unless C is a formula of ZFC (i.e. unless C is internal) there is no axiom to permit the formation of y. For example let x be $\mathbb{R}$ and let C(z) be "$z \simeq 0$". Then we cannot form the set $\{z \in \mathbb{R} : z \simeq 0\}$ within IST; the set y in (S) in this case is $\{0\}$.

Fortunately there are two consequences of (S) that are much easier to grasp and are sufficient for our purposes. We call the first of these the *standard part* property (SP).

(SP) *Every limited real number is infinitely close to a standard real number.*

For example, $\sqrt{2} + 1/\omega$ is limited and is infinitely close to $\sqrt{2}$ which is standard. If x is limited and $x \simeq y$ where y is standard then y is unique and is called the *standard part* of x, written $y = {}^o x$. A similar property to (SP) is easily seen to hold for complex numbers by considering real and imaginary parts.

The second consequence of (S) is called *External Induction* (EI).

(EI) *Let* $A(x)$ *be any formula, internal or external, with* x *as free variable and perhaps other free variables. Suppose that* $A(0)$ *and for all standard natural numbers* n, *if* $A(n)$ *then* $A(n+1)$. *Then for all standard natural numbers* n *we have* $A(n)$.

We can use this to prove that if z and w are limited complex numbers and $z \simeq w$ then $z^n \simeq w^n$ for all standard n. The induction step is accomplished by noting that if $z_1 \simeq w_1$ and $z_2 \simeq w_2$ (all limited) then $z_1 z_2 \simeq w_1 w_2$. This is because $z_1 = w_1 + \varepsilon_1$ and $z_2 = w_2 + \varepsilon_2$ where ε_1 and ε_2 are infinitesimal and therefore

$$z_1 z_2 - w_1 w_2 = \varepsilon_1 w_2 + \varepsilon_2 w_1 + \varepsilon_1 \varepsilon_2$$

which is clearly infinitesimal. We can extend the result to standard polynomials by a second application of (EI). Let P be a standard polynomial. Then its degree is a standard natural number and its coefficients are standard complex numbers. If z and w are limited and $z \simeq w$ then $P(x) \simeq P(w)$.

3. The Fundamental Theorem of Algebra

The fundamental theorem of algebra states that every non-constant polynomial with complex coefficients has at least one complex root. The following proof is based on a classical one ([3], pp. 53-55), but the availability of infinitesimals greatly simplifies the technical details.

Let $P(z) = a_0 + a_1 z + \ldots + a_n z^n$ be any complex polynomial with $n \geq 1$ and $a_n \neq 0$. We must prove that P has a root. By transfer we may suppose that P is standard. Then n and all the a_j are standard.

(i) We first prove that $|P(z)|$ attains its minimum at some standard point β in $\mathbb{C}$. Observe that if $|z| \simeq +\infty$ then $|P(z)| \simeq +\infty$. The reason is that if $|z| \simeq +\infty$ then

$$\frac{P(z)}{z^n} = a_n + \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \ldots + \frac{a_0}{z^n} \simeq a_n,$$

and a_n is standard and non-zero. Now let $\omega \in \mathbb{N}$ satisfy $\omega \simeq +\infty$ and let F be defined by

$$F = \{(m + in)/\omega : m,n \in \mathbb{Z}, \ |m|, |n| \leq \omega^2\}.$$

Because F is a finite set, the minimum of $|P(z)|$ as z runs through F is attained at some point α of F. By the remark above $|\alpha| << +\infty$ and so by (SP), $\alpha \simeq \beta$ for some standard β in $\mathbb{C}$. We will prove that $|P(z)| \geq |P(\beta)|$ for all z. By transfer, it suffices to prove this for all standard z. If z is standard then $z \simeq$ some point ζ of F, and therefore

$$|P(z)| \simeq |P(\zeta)| \geq |P(\alpha)| \simeq |P(\beta)|.$$

Since the extreme numbers are standard, it follows that $|P(z)| \geq |P(\beta)|$.

(ii) It is an elementary fact that every equation of the form $z^n = c$ has a solution in $\mathbb{C}$. It suffices to write $c = re^{i\theta}$ and let $z = r^{1/n}e^{i\theta/n}$.

(iii) We now prove that $P(\beta) = 0$. This is done by examining $P(\beta+h)$ when $h \in \mathbb{C}$ and $|h|$ is small. Applying the binomial theorem to $(\beta+h)^j$ we obtain

$$P(\beta+h) = P(\beta) + b_1h + b_2h + \dots + b_nh^n \quad (b_n=a_n\neq 0)$$

where the b_j are standard complex numbers. Let b_m be the first non-zero coefficient. Then

$$P(\beta+h) = P(\beta) + b_mh^m(1+Q(h)), \tag{3}$$

where $Q(h)$ is a standard polynomial with no constant term; $Q(0) = 0$. Now let h be a solution of the equation

$$z^m = -\frac{1}{\omega^m}\frac{P(\beta)}{b_m}.$$

Then $h \simeq 0$ and so $Q(h) \simeq Q(0) = 0$ and hence $|Q(h)| < 1$. From (3) we have

$$P(\beta+h) = \left(1 - \frac{1}{\omega^m}\right)P(\beta) - \frac{1}{\omega^m}P(\beta)Q(h)$$

and if $P(\beta) \neq 0$ then we have the contradiction

$$|P(\beta+h)| \leq \left(1 - \frac{1}{\omega^m}\right)|P(\beta)| + \frac{1}{\omega^m}|P(\beta)||Q(h)| < |P(\beta)|.$$

Thus $P(\beta) = 0$ and the proof is complete.

<u>**Remarks**</u>. This proof is elementary in that it avoids the notions of compactness, continuity and complex integration. It is also constructive, at least in spirit, because the argument by contradiction is held back until the end. In practice we have no unlimited natural number ω to help us, but for any

specific polynomial, a suitably large standard value of ω will make the mesh F fine enough to catch the zeros of F approximately.

4. The Spectral Theorem

The following result is a version of the Jordan normal form theorem. The proof is due to Lutz and Goze [6] and it is so natural that it deserves to be widely known.

THEOREM. *Let $f : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator whose distinct eigenvalues are $\lambda_1, \lambda_2, \dots, \lambda_p$, where λ_j has multiplicity m_j, $j=1,\dots,p$. Then there is a basis $u_1, u_2, \dots, u_n$ for $\mathbb{C}^n$ such that*

$$M(f,(u_j)) = \begin{bmatrix} A_1 & & 0 \\ & A_2 & \\ & & \ddots & \\ 0 & & & A_p \end{bmatrix}, \quad A_j = \begin{bmatrix} \lambda_j & * & * & * \\ & \lambda_j & & * \\ & 0 & \ddots & \\ & & & \lambda_j \end{bmatrix}$$

where A_j is an $m_j \times m_j$ matrix.

COROLLARY. *There is a direct sum decomposition $\mathbb{C}^n = F_1 \oplus F_2 \oplus \dots \oplus F_p$ such that $\dim(F_j) = m_j$, $f(F_j) \subseteq F_j$ and $f|_{F_j} = \lambda_j I + N_j$ where N_j is nilpotent.*

PROOF. If f has no repeated eigenvalues so that $m_j = 1$ for all j, then the result is classical and elementary. The difficulty comes from the possibility that some eigenvalues may be repeated. We can get rid of such degeneracies by an infinitesimal perturbation and this is the key idea of the proof, which is in three parts.

(i) Choose a "Good" Infinitesimal Perturbation g of f

Choose a linear map $g : \mathbb{C}^n \to \mathbb{C}^n$ such that $g(x) \approx f(x)$ for all limited x, the eigenvalues of $\mu_1, \mu_2, \dots, \mu_n$ g are distinct and each $\mu_j \approx$ some λ_i. This can be done by making infinitesimal changes in a matrix representing f. Since the eigenvalues of g are distinct, there is a basis $v_1, \dots, v_n$ of $\mathbb{C}^n$ consisting of eigenvectors of g. We normalise so that $|v_j| = 1$ (where $|x|$ denotes the Euclidean norm of x).

(ii) Get a "Good" Matrix Representation of g

Group all μ_j which are infinitely close, and group the corresponding v_j. After relabelling we get

$$\mu_1 \approx \mu_2 \approx \dots \approx \mu_{m_1} \approx \lambda_1 \; ; \; \mu_{m_1+1} \approx \dots \approx \mu_{m_1+m_2} \approx \lambda_2 ; \text{ etc.}$$

Define subspaces G_j by

$$G_1 = \operatorname{span}\{v_1, v_2, \dots, v_{m_1}\}, \quad G_2 = \operatorname{span}\{v_{m_1+1}, \dots, v_{m_1+m_2}\},$$

etc.

Then clearly $\dim G_j = m_j$, $g(G_j) \subseteq G_j$ and $\mathbb{C}^n = G_1 \oplus G_2 \oplus \dots \oplus G_p$. In each G_j, use the Gram-Schmidt process to obtain an orthonormal basis (w_i) from (v_i). We get

$$G_1 = \operatorname{span}\{w_1, w_2, \dots, w_{m_1}\}, \quad G_2 = \operatorname{span}\{w_{m_1+1}, \dots, w_{m_1+m_2}\},$$

etc., and $\langle w_i, w_j \rangle = \delta_{ij}$ within each subspace. By the nature of the Gram-Schmidt process

$$M(g,(w_i)) = \begin{bmatrix} B_1 & & \bigcirc \\ & B_2 & \\ \bigcirc & & \ddots \\ & & & B_p \end{bmatrix}, \quad B_1 = \begin{bmatrix} \mu_1 & * & \dots & * \\ & \mu_2 & & \vdots \\ \bigcirc & & \ddots & \\ & & & \mu_{m_1} \end{bmatrix} \text{ etc.}$$

(iii) Deduce a "Good" Matrix Representation for f

For each j, w_j is limited because $|w_j| = 1$ and therefore $w_j \approx u_j$ for some standard $u_j \in \mathbb{C}^n$. Define subspaces $F_1 F_2, \dots, F_p$ by

$$F_1 = \operatorname{span}\{u_1, u_2, \dots, u_{m_1}\}, \quad F_2 = \operatorname{span}\{u_{m_1+1}, \dots, u_{m_1+m_2}\},$$

etc., so that in a certain sense $F_j \approx G_j$. We pass now from a nonstandard situation:

$$\mathbb{C}^n = G_1 \oplus G_2 \oplus \dots \oplus G_p, \quad g(G_j) \subseteq G_j, \ \dim G_j = m_j, \text{ and}$$

$$M(g,(w_i)) = \operatorname{diag}[B_1, B_2, \dots, B_p],$$

to a standard one. It is easy to believe and not difficult to prove that

$$\mathbb{C}^n = F_1 \oplus F_2 \oplus \dots \oplus F_p, \quad f(F_j) \subseteq F_j, \quad \dim F_j = m_j, \text{ and}$$

$$M(f,(u_i)) = \operatorname{diag}[A_1, A_2, \dots, A_p]$$

where A_j is standard and $A_j \approx B_j$ in the sense that corresponding entries are infinitely close, so that

$$A_j = \begin{bmatrix} \lambda_j & * & * & \dots & * \\ & \lambda_j & & & \vdots \\ & & \ddots & & \\ \bigcirc & & & \ddots & * \\ & & & & \lambda_j \end{bmatrix}$$

5. Concluding Remarks

Nonstandard analysis is also of use in providing new results and in providing the framework for mathematical modelling of physical processes where different orders of magnitude are involved. Two recent examples are moiré patterns [4] and the theory of singular perturbations of ordinary differential equations [6], [1], [9].

It should be mentioned that the original, constructive approach to nonstandard analysis due to Robinson [8] is also flourishing today especially in the fields of functional analysis, measure theory, stochastic differential equations and optimal control [2], [5].

References

1. CARTIER, P.
"Perturbations Singulières des Équations Différentielles et Analyse Non-Standard", Séminaire Bourbaki No. 580, *Astérisque* 92-93 (1982) 21-44.

2. CUTLAND, N.J.
"Nonstandard Measure Theory and its Applications", *Bull. London Math Soc.*, 15 (1983) 529-589.

3. DIENES, P.
'The Taylor Series', Dover, New York, 1957.

4. HARTHONG, J.
"Le Moiré", *C.R. Acad. Sci. Paris*, Sér. A-B 290 (1980), No. 19, A877-A879.

5. HURD, A. (Editor).
'Nonstandard Analysis - Recent Developments', Lecture Notes in Mathematics 983 (Springer, Berlin, 1981).

6. LUTZ, R. and GOZE, M.
'Nonstandard Analysis. A Practical Guide with Applications', Lecture Notes in Mathematics 881 (Springer, Berlin, 1981).

7. NELSON, E.
"Internal Set Theory: A New Approach to Nonstandard Analysis", *Bull. Amer. Math. Soc.*, 83 (1977) 1165-1198.

8. ROBINSON, A.
'Non-Standard Analysis', 2nd Ed, American Elsevier, N.Y. 1974.

9. ZVONKIN, A. and SHUBIN, M.A.
"Non-Standard Analysis and Singular Perturbations of Ordinary Differential Equations", *Russian Math. Surveys*, 39:2 (1984) 69-131.

Faculty of Mathematical Studies,
The University,
Southampton.