ANALYSE NON STANDARD ET MATHÉMATIQUES ORDINAIRES

Claude LOBRY

Introduction

On se souvient peut-être que dans les années 1980-2000 il y eut, en France, une discussion assez vive sur les mérites de l'Analyse Non Standard (ANS). Pour les uns il s'agissait d'une révolution scientifique qui devait s'imposer dans tous les domaines des mathématiques et dès l'enseignement élémentaire et pour les autres, l'immense majorité, il s'agissait d'une sous-culture à laquelle il fallait résolument tourner le dos. La revue Repères-IREM a publié plusieurs articles de partisans de la méthode [1] ([8, 14, 15, 27]). Qu'en est-il un quart de siècle plus tard ?

L'ANS, c'est-à-dire la pratique des infinitésimaux, est maintenant reconnue comme légitime. Il existe même de très grands mathématiciens comme T. Tao [2] ou M. Gromov [3] qui en préconisent l'usage pour certaines activités de recherche particulièrement pointues. Je ne mentionne pas ces grands mathématiciens pour ouvrir une polémique comptabilisant les autorités qui se déclarent "pour" ou "contre" l'ANS, mais pour que ceux qui pourraient penser que la querelle des années 1980 sur l'usage de l'ANS est une vieillerie intéressant au mieux

Date : 28 février 2021.

1. Mais, me semble-t-il, pas d'opinions de détracteurs.
2. Terence Tao, médaille Fields 2002, s'exprime en abondance sur son site ; voici un exemple où il parle de l'ANS : https://terrytao.wordpress.com/tag/nonstandard-analysis/
3. Prix Abel 2009. Dans l'introduction de l'article *In a Search for a Structure, Part 1 : On Entropy :* http://math.mit.edu/~dspivak/teaching/sp13/gromov--EntropyViaCT.pdf il explique comment certaines idées de physique qui semblaient du "charabia" sont formalisables à travers l'ANS ou la théorie des catégories. Je cite :
Mais de nos jours, ce «charabia» se traduit (presque) automatiquement dans le langage de l'analyse non standard (Abraham Robinson, 1966) et encore plus facilement dans celui de la théorie des catégories (Eilenberg-MacLane-Steenrod-Cartan-Grothendieck, 1945-1957).

les historiens des sciences, sachent qu'il existe en ce moment des mathématiciens qui l'utilisent et en préconisent l'usage. Cependant l'ANS reste une pratique minoritaire et, il ne semble pas qu'elle soit enseignée actuellement de façon significative à quelque niveau que ce soit. Je me propose, dans cet article, de mettre en évidence quelques bloquages psychologiques à l'origine des réticences à l'usage naturel des quantités infinitésimales.

Pour cela je commencerai par développer rapidement l'idée, largement partagée, que l'activité mathématique se construit d'abord avec des images mentales personnelles plus ou moins transmissibles avant de se traduire par des démonstrations en bonne et due forme ; sans une image mentale la soutenant une démonstration est incompréhensible. Ensuite je présenterai les deux principales versions de l'ANS : celle des extensions, associée au nom de Abraham Robinson (1918-1974), puis l'axiomatique, associée à celui de Edouard Nelson (1932-2014), et nous verrons en quoi elles posent problème en matière d'image mentale. En espérant avoir ainsi levé les obstacles conceptuels, je proposerai (section 4) quelques exercices qui permettent de se familiariser avec I.S.T. J'aborderai ensuite ce que j'appelle le "grand malentendu ". Le grand malentendu est de croire que l'ANS est une technique particulière pour (re)démontrer des résultats mathématiques "conventionnels" (ou traditionnels), c'est-à-dire n'utilisant pas directement les notions floues de "grand", "petit"ou "du même ordre de grandeur" si présentes dans le discours des physiciens et que la mathématique traditionnelle ne sait traduire que de façon très indirecte. Je donnerai quelques exemples avant de conclure.

Une façon de lutter contre ce malentendu est de proposer une ANS qui ne permette pas cette traduction en langage classique, une ANS où les énoncés avec des infinitésimaux n'ont pas d'équivalent traditionnel. C'est ce qu'ont proposé à peu près simultanément Nelson [28] Lutz [24] et Callot [6]. Je ne développe pas ici ce point vue, ce qui m'obligerait de sortir du cadre traditionnel, que j'ai scrupuleusement respecté, de ce qu'on appelle la "rigueur mathématique". J'espère le faire dans un autre article où je me permettrai de sortir de ce cadre pour parler des liens entre l'A.N.S. et l'intuitionnisme tels que Harthong et Reeb [18] les voyaient.

Le présent article est en partie issu de [23] où j'essaye d'expliquer les raisons du mauvais accueil de l'ANS dans la communauté mathématique. Cette question d'image mentale que je développe ici en est un des aspects. Il y en a d'autres.

Remerciements. Une première version de cet article a bénéficié d'une relecture minutieuse d'Augustin Fruschard qui y a corrigé un nombre de coquilles qui, par son ampleur, donne une excellente idée de ce qu'est un entier infiniment grand au sens de l'ANS. Une lecture non moins minutieuse des relectrices-teurs de la revue a contribué grandement à en améliorer la forme et le fond. Je les remercie tous.

1. — La place de l'image mentale en mathématiques.

On a beaucoup discuté et l'on discutera encore longtemps sur l'existence et la nature de l'intuition mathématique. Pour les uns, elle est une forme d'accès à un monde des objets mathématiques qui existe de toute éternité, pour d'autres elle n'existe pas, parce que le monde des idées mathématiques n'existe tout simplement pas. Comme je ne souhaite pas faire dépendre ma démonstration de l'issue d'une question qui n'est pas près d'être tranchée, je n'utiliserai pas cette expression au profit de celle d'image mentale qui, je pense, devrait être acceptée sans difficulté par tous les mathématiciens. Par mathématicien j'entends, dans cette discussion, quelqu'un qui a le niveau d'une licence de mathématiques, pas nécessairement plus, et qui les pratique dans le cadre de sa profession. Ce terme inclut donc à la fois des chercheurs, des ingénieurs et des enseignants.

Dans la pratique des mathématiques il faut distinguer deux activités. D'une part, celle où l'individu se crée une intime conviction sur la vérité de telle ou telle proposition, d'autre part, celle où il partage ses convictions avec d'autres mathématiciens. Ainsi l'activité mathématique se décompose-t-elle en une succession d'efforts solitaires, où l'on se creuse la tête pour comprendre, alternant avec des périodes d'échange où l'on s'efforce de transmettre et recevoir[4]. La part collective du travail mathématique est très spécifique et bien comprise : on échange à l'aide d'un langage particulier qui permet des aller-retours rapides entre des formulations un peu vagues, mais parlantes, en d'autres termes des formulations qui ont du "sens", et les formulations plus précises que permettent le formalisme, formulations qui aident à lever les ambiguïtés [21]. Tout mathématicien reconnaitra volontiers que, lorsqu'il travaille à la recherche de la solution d'un problème, son esprit ne fonctionne pas de façon formelle dans un cadre axiomatique précis, mais de façon tout à fait informelle et peu rigoureuse par association d'images mentales diverses. Pour un même objet, l'image mentale qui se présente dans la tête du mathématicien diffère sensiblement d'un individu à un autre. Toutefois, lorsqu'il s'agit d'objets aussi fondamentaux que **N** ou **R** on peut penser que les images mentales des uns et des autres sont très proches. Dans les écrits mathématiques, puisque précisément on souhaite transmettre des connaissances non ambiguës reconnaissables par un individu convenablement éduqué, on ne s'attarde que peu, sauf à titre de motivations, sur les images mentales. Ainsi les images mentales sont-elles pratiquement absentes du discours mathématique écrit, un peu moins du discours parlé, mais bien présentes dans la tête des mathématiciens.

Le grand géomètre W. Thurston (1946-2012) nous en parle avec beaucoup de talent dans *On proof and progress in mathematics*[5]([34]).

> La pensée humaine et la compréhension ne fonctionnent pas en suivant une seule voie, comme un ordinateur avec une unique unité centrale. Nos cerveaux et nos esprits semblent organisés en dispositifs fonctionnels relativement séparés, aux aptitudes puissantes et variées. Ces dispositifs travaillent ensemble de façon assez lâche, "parlant" à chacun des autres à de hauts niveaux plutôt qu'à de bas niveaux d'organisation.

Il passe en revue ces dispositifs fonctionnels :

(1) *le langage humain* : [...] Notre aptitude linguistique est un outil important pour penser, et non simplement pour communiquer...»

(2) *La vision, la perception spatiale, la kinesthésique* : Les gens possèdent des dispositifs très efficaces pour acquérir des informations de manière visuelle [...]

(3) *Logique et déduction* : [...] Apparemment, et en général, les mathématiciens ne s'appuient pas sur les règles formelles de la déduction pendant qu'ils pensent [...]

(4) *Intuition, associations, métaphores* [...] Personnellement, je consacre des efforts importants à "écouter" mes intuitions [...] Cela implique simultanément une sorte de

4. Voir sur ce sujet l'essai de G. Wallet et S. Neuwirth [37] https://hal.archives-ouvertes.fr/hal-01824928v2/document

5. La traduction en français que j'utilise ici a été publiée dans le n 21 de Repères-IREM

quiétude et de concentration de mon esprit. Les mots, la logique, les figures précises qui crépitent autour peuvent inhiber les intuitions et les associations.

(5) *Stimulus-réponse* […]

(6) *Processus et temps* […]

qui s'agitent dans la tête du mathématicien pour se demander « comment la compréhension mathématique se communique-t-elle ?» Il conclut :

> … C'est difficile et compliqué. Par conséquent, pour analyser la compréhension humaine des mathématiques, il est important de préciser **qui** comprend **quoi**, et **quand** [6].

C'est à ce dernier exercice que je vais tenter de me livrer a propos de l'ANS.

2. — Les nombres réels, hyper-réels et le continu

"Un enclos est un espace de terrain entouré d'une clôture qui sert à contenir des animaux domestiques" (wikipedia). L'image du berger de la haute antiquité qui contemple son troupeau de moutons enfermé dans un enclos contient une des plus remarquables conquêtes des mathématiques : avoir relié les deux phénomènes si immédiatement présents et si différents que sont la sensation de l'individu et celle de la clôture, le nombre entier et la courbe, le discret et le continu. La chose ne s'est pas faite instantanément, il a fallu quelques millénaires d'efforts dont on peut suivre la trace dans l'enseignement.

— On commence par apprendre une récitation qui ne se termine pas Un, deux, trois, …, **dix**, onze, douze, …, **vingt**, vingt et un, vingt et deux, …, …, **cent**, …, …, …
L'addition, la multiplication, les tables…

— Puis on s'attaque aux "grandeurs" avec les inégalités, les points ordonnés sur une droite, les fractions qui nous disent ce qui revient à chacun de deux tartes à partager entre trois enfants…

— On apprend à se servir des "nombres avec virgule", de la représentation décimale, de l'art de la division …

N

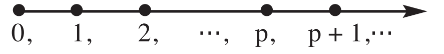


Figure 1. Une image mentale de **N**.

Mais, à côté de ces activités qui se rangent dans le champ des mathématiques, il ne faut pas oublier le dessin si essentiel dans l'enseignement pré-élémentaire. C'est avec lui que très tôt on apprend à distinguer l'intérieur de l'extérieur séparés par une frontière imperméable : le continu infranchissable de la ligne tracée sur le papier. Ce n'est que bien plus tard que l'on apprend à mesurer la part d'illusion qu'il y a dans cette vision. Associée à cette question du continu, mais dans un champ plus mathématique, on trouve la question de la *grandeur* qui n'est pas un nombre entier mais qui y ressemble (ça se compare, ça s'ajoute). C'est ainsi que, en classe terminale, l'élève a une maîtrise suffisante des nombres et du calcul pour suivre le programme de physique où il est question de mesure, d'incertitude, de vitesse etc.

Arrivé à ce stade, le futur mathématicien possède déjà une image mentale très forte des nombres réels comme des points ordonnés sur une droite ; il a conscience qu'il y en a beaucoup et il sent plus ou moins clairement que ceci a à voir avec le fait que, quand on trace sur le papier un enclos, pardon une courbe fermée sans point double, alors, une courbe qui part de l'intérieur pour aller à l'extérieur, la coupe forcément en un

6. Emphase de l'auteur

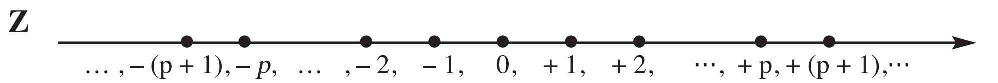


Figure 2. Une image mentale de **Z**.

point. Mais il ne sait pas encore que cette vague intuition qu'il a des nombres se théorise et porte le nom de complétude.

C'est à partir de la licence [7] que l'on va reprendre tout ça dès le début pour lui donner des fondements plus solides et l'*image mentale* qui a déjà une certaine consistance va être renforcée par l'enseignement proposé. Je prie le lecteur de m'excuser d'en rappeler les détails, qui lui sont certainement familiers, mais j'aurai besoin qu'il les ait bien à l'esprit quand j'en viendrai à l'ANS ; toutefois une lecture attentive n'est pas nécessaire et il peut, sans dommage, se contenter de survoler le paragraphe suivant, juste histoire de se rafraîchir la mémoire.

2.1. *Une construction traditionnelle de* **R**.

On commence par donner des rudiments de théorie naïve des ensembles et l'on explique que la règle du jeu va être de tout construire à partir de rien (ou presque, de l'ensemble vide !). Je passe sur cette phase qui est probablement vécue de façon diverse selon la quantité plus ou moins grande de formalisme mis en place à ce moment, pour passer tout de suite au stade où l'ensemble **N** est construit.

> L'*image mentale* de **N** est celle d'un ensemble infini totalement ordonné, où tout élément a un successeur, toute partie non vide a un plus petit élément, dont on se sert pour indexer des suites (u_n), ou des propriété $P(n)$ que l'on démontre par récurrence, c'est-à-dire que : de $\big(P(0)$ et de $P(n) \Rightarrow P(n+1)\big)$ on infère $\forall n\ P(n)$.

7. Je ne sais pas exactement à quel moment, mais cela n'a pas d'importance pour ma démonstration.

Si l'on excepte l'enrichissement du raisonnement par le principe de récurrence qui commence à avoir son petit parfum de formalisme, cette *image mentale* est en tout point conforme à celle qui était issue de l'enseignement primaire (voir la figure 1).

Ensuite on construit les entiers relatifs. On ne dira pas qu'il s'agit de prendre une autre copie de **N** précédée du signe –, ce que l'on imagine naturellement, mais qui est une opération non permise en théorie des ensembles. On construit donc en respectant les règles de constructions ensemblistes :

(1) On considère l'ensemble $E = \mathbf{N} \times \mathbf{N}$ des couples d'entiers.

(2) On munit E de la relation d'équivalence $(a, b) \sim (c, d) \Longleftrightarrow a + d = b + c$ et on "passe au quotient" $\mathbf{Z} = E / \sim$.

(3) On "importe" l'addition, la multiplication, l'ordre sur **Z** et on "injecte canoniquement" **N** dans **Z**.

On ne retrouve pas exactement l'image mentale du nombre négatif − 3 dans l'ensemble des paires a, b telles que $b = a + 3$ et l'on préfère une image mentale du genre de celle de la figure 2 où les entiers relatifs s'ordonnent le long d'une droite, ce qui est d'ailleurs conforme au fait que **Z** est totalement ordonné. On passe ensuite au corps des rationnels.

(1) On considère l'ensemble $F = \mathbf{Z} \times \mathbf{Z}^*$ des couples d'entiers relatifs.

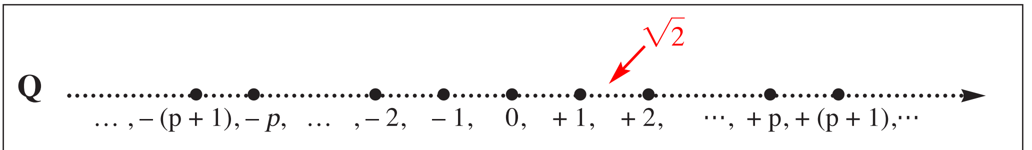


Figure 3. Une image mentale de **Q**.

(2) On munit F de la relation d'équivalence $(p, q) \sim (r, s) \Longleftrightarrow ps = qr$ et on "passe au quotient" $\mathbf{Q} = F / \sim$.

(3) On "importe" l'addition, la multiplication, l'ordre sur **Q** et on "injecte canoniquement" **Z** dans **Q**.

On fait remarquer, au passage, la similitude des deux constructions. Que trouve-t-on maintenant dans l'image mentale de **Q** ? Certainement l'idée de fraction $\frac{p}{q}$ mais aussi celle de nombre écrit dans une base décimale (un nombre à virgule), et la densité : entre deux rationnels il y en a toujours un troisième. Mais on devine qu'on n'a pas tous les nombres puisque le développement décimal d'une fraction est fini ou périodique ; qui sont donc ces nombres dont le développement est arbitraire ? Et que veut dire un développement décimal "infiniment long" ? Et puis il y a le scandale de $\sqrt{2}$ qui n'est pas un nombre rationnel. On insiste beaucoup, à juste titre, sur cette affaire : il faut absolument un nombre pour désigner la longueur de la diagonale du carré de côté unité. Il n'est quand même pas acceptable que l'hypoténuse du triangle rectangle $(3 \perp 4\ , 5)$ ait une longueur et pas celle du triangle rectangle $(1 \perp 1\ , ?)$. Au rebelle pragmatique qui trouve bien suffisant d'avoir des nombres rationnels dont le carré est aussi proche que l'on veut de 2, on demande alors ce qu'il pense de la possibilité de s'échapper du disque de rayon 1 sans rencontrer le cercle en passant par le point (qui n'existe pas encore) de coordonnées $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ et il convient alors que, là, c'est bien un problème. On ne doit pas pouvoir s'échapper du disque unité sans rencontrer le cercle de rayon 1 ; sauf à être capable de sauter par-des-

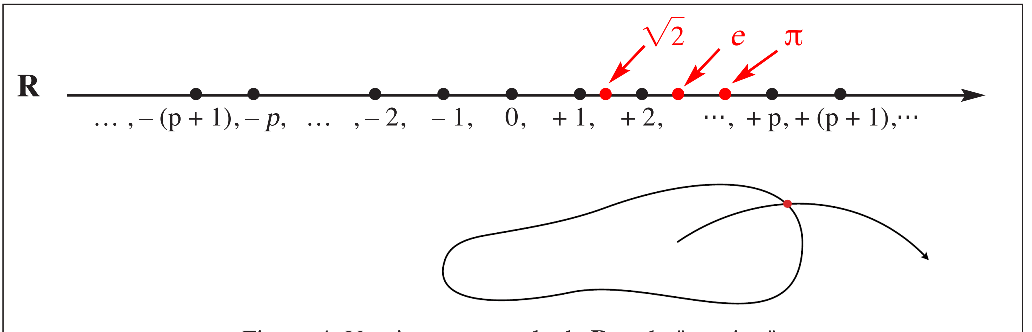


Figure 4. Une image mentale de **R** et du "continu".

sus la clôture, le mouton idéalisé, qui est un point sans dimension, ne doit pas pouvoir sortir. L'image qui va s'imposer pour **Q** est donc celle d'une sorte de pointillé "partout dense" (entre deux rationnels il y en a un troisième), qui contient **N**, mais finalement "plein de trous" ; c'est la figure 3 où le $\sqrt{2}$ désigne une absence.

Maintenant que nous avons bien compris que **Q** ne saurait remplir le rôle d'une clôture sans trou ni mesurer toutes les grandeurs, nous sommes prêts à accepter un pas de plus dans l'abstraction : la construction de **R** par les suites de Cauchy (on pourrait sans peine faire le même discours à propos de la construction par les coupures, mais celle par les suites de Cauchy est mieux adaptée à l'introduction de l'ANS).

(1) On considère l'ensemble

$$G = \left\{(u_n)\ u_n \in \mathbf{Q} : (u_n) \text{ est de Cauchy}\right\}$$

(2) On munit G de la relation d'équivalence $(u_n) \sim (v_n) \Longleftrightarrow (u_n - v_n) \to 0$ et l'on "passe au quotient" $\mathbf{R} = G/\sim$.

(3) On "importe" l'addition, la multiplication, l'ordre sur **R** et l'on "injecte canoniquement" **Q** dans **R**.

On démontre, cette fois, que **R** est complet, ce qui entraîne que le graphe d'une fonction continue qui a des valeurs négatives puis positives s'annule forcément et on énonce le théorème de Jordan qu'on démontre parfois (sous une forme plus ou moins forte). Cette fois c'est bien gagné : *il n'y a plus de trous.* Nous avons au final :

$$\mathbf{N} \hookrightarrow \mathbf{Z} \hookrightarrow \mathbf{Q} \hookrightarrow \mathbf{R}$$

et une *image mentale* de **R** qui est celle d'une droite (on dit la "droite réelle") où tous les nombres, les entiers, les fractions, les irrationnels et les transcendants ont leur place (figure 4). Sur cette figure $\sqrt{2}$ ne désigne plus une absence mais un point. Le lien est fait entre le *continu* de la grandeur « où l'on suppose entre les éléments du continu une sorte de lien intime qui en fait un tout, où le point ne préexiste pas à la ligne, mais la ligne au point »[8], et le *discret* de l'énumération.

Je ne prétends pas que l'image mentale de **R** que je viens de proposer était exactement celle de tous les mathématiciens de ma génération. C'est *mon image mentale* et le propre de l'image mentale est d'être personnelle. Mais comme l'image mentale se construit tout doucement au cours de l'éducation mathématique et que cette dernière est à peu près la même pour tous, il me semble que ce que je viens de décrire doit être assez représentatif.

2.2. *L'ensemble* **R** *et les ordres de grandeurs.*

Oublions quelques temps la question de la "mesure des grandeurs" pour nous pencher sur celle des "ordres de grandeurs". Si un élève d'un cours élémentaire nous dit que les racines de $x^2 + 0{,}01x - 1 = 0$ sont -2 et 4 nous savons immédiatement qu'il s'est trompé parce que nous savons *a priori* que les racines sont deux nombres proches respectivement de $+1$ et -1. Nous le savons parce que le coefficient de x, 0,01, est *petit* par rapport aux autres coefficients, que nous pouvons *négliger* ce terme et que les racines sont donc *proches* de celles de $x^2 - 1 = 0$ (des valeurs approchées des racines sont : 0.995012499921876 et $-$ 1.00501249992188).

Dans un autre ordre d'idées, pour connaître la pente de la tangente à $y = x^2$ au point x_0 nous évaluons la pente de la sécante qui passe par les deux points "infiniment proches" (x_0, x_0^2) et $(x_0 + \mathrm{d}x, (x_0 + dx)^2)$ soit :

$$\frac{x_0^2 + 2x_0dx + dx^2 - x_0^2}{x_0 + dx - x_0} = \frac{dx(2x_0 + dx)}{dx} = 2x_0 + dx$$

8. H. Poincaré, *Le continu mathématique*, Revue de métaphysique et de morale, 1(1), 26-34 (1893)

à condition que dx ne soit pas nul, car sinon ce calcul n'aurait pas de sens, mais si dx est infiniment petit on le néglige et la pente de la tangente est $2x_0$. Cette question des ordres de grandeurs est mise en avant par Leibniz dans une célèbre lettre à Varignon du 2 février 1702 :

> [...] c'est ainsi qu'une parcelle de la matière magnétique qui passe à travers du verre n'est pas comparable avec un grain de sable, ni ce grain avec le globe de la terre, ni ce globe avec le firmament.

pour justifier l'usage qu'il fait des infinitésimaux. Il insiste dans cette même lettre sur le fait que l'infini qu'il introduit en mathématiques n'a rien de métaphysique [9] :

> Je ne me souviens pas assez des expressions dont je m'y puis estre servi [à propos du calcul infinitésimal], mais mon dessein a esté de marquer, qu'on n'a point besoin de faire dépendre l'analyse Mathematique des controverses metaphysiques [...]

Considérons donc des nombres réels dx qui soient infiniment petits. Le problème avec notre infiniment petit positif non nul dx est que, d'une part, c'est un nombre strictement positif, donc il existe un entier n tel que $ndx > 1$ (**R** est archimédien) et, d'autre part, pour qu'il soit utile, le calcul avec des infiniment petits doit être tel que la somme de deux infiniment petits soit un infiniment petit. Comme 1 n'est pas infiniment petit, il y a un problème avec l'entier n_0 qui est le premier tel que n_0dx ne soit pas infiniment petit. Alors $(n_0 - 1)dx$ est infiniment petit et $(n_0 - 1)dx + dx = n_0dx$ est infiniment petit. C'est un problème sérieux ! On ne peut pas être à la fois une chose et son contraire. C'est le fameux et vieux [10] problème du "tas de sable" : "être un tas de sable" est une propriété "floue". Il n'y a pas de nombre de grains de sable bien défini au delà duquel, un ensemble de grains de sable mérite le nom de "tas".

En mathématiques traditionnelles le "flou" n'a pas sa place. Il n'y a pas de droites à peu près parallèles, de cercles à peu près concentriques, de pentagones à à peu près cinq côtés. Une propriété, si c'est une propriété mathématique bien formulée — nous reviendrons sur cette notion de "propriété mathématique bien formulée" — ne peut pas être vraie et fausse à la fois, ni approximativement vraie. Il semble qu'il y ait un obstacle insurmontable à accepter des nombres qui seraient *infiniment petits* non nuls (et donc dont les inverses seraient *infiniment grands*).

Pourtant, comparer puis négliger ce qui est petit est la pratique courante du physicien. Pendant des siècles, des calculs utilisant le concept d'ordre de grandeur ont donné d'excellents résultats bien qu'ils ne soient pas logiquement bien fondés. En fait ils l'étaient, mais on ne le savait pas. Ou plutôt, il est plus correct de le dire ainsi : *On a trouvé, au milieu du siècle* [11] *dernier des fondements logiques impeccables (aux yeux des exigences actuelles de la rigueur) au concept flou d'ordre de grandeur* [12], c'est ce qu'on appelle l'ANS.

9. Il existe d'innombrables travaux historiques et philosophiques sur le concept d'infini chez Leibniz. Une tendance actuelle (le lecteur intéressé pourra trouver dans [33] et sa bibliographie une piste de réflexion) serait de distinguer chez lui deux sortes d'infini : d'une part l'infini pragmatique (une fiction utile) pour faciliter les calculs, de l'autre un infini metaphysique (qui contiendrait "tout"). Il va de soi que l'évolution du concept d'infini, l'histoire qui en est faite et l'évolution de cette histoire, ont leur importance pour comprendre la réception de l'ANS mais je n'aborde pas ce point ici.

10. On attribue à Eubulide (∼ 400) de nombreux paradoxes dont celui du "tas".

11. L'acte de naissance de l'ANS est le livre de Robinson [32] de 1966 mais bien entendu les idées étaient en germe dans de nombreux travaux précédents. Voir Reeb *La mathématique non standard vieille de soixante ans ?* [30].

12. Sur la possibilité de représenter mathématiquement le flou on pourra lire l'ouvrage [36] de van den Berg et Dinis, *Neutrices and External Numbers : A Flexible Number System* .

Voyons une première possibilité pour fonder l'analyse non-standard.

2.3. *Les hyper-réels.*

Une présentation courante des infinitésimaux est celle de la construction d'une extension notée $^*\mathbf{R}$ du corps des réels :

$$\mathbf{R} \hookrightarrow {}^*\mathbf{R}$$

Cette construction n'est pas difficile du tout. En voici une idée rapide.

Un ultrafiltre sur **N** est une famille de parties $\mathcal{F}$ telle que :

(1) $\emptyset \notin \mathcal{F}$

(2) $A \in \mathcal{F} \;\; B \in \mathcal{F} \Longrightarrow A \cap B \in \mathcal{F}$

(3) Si A est une partie de **N**, soit A, soit son complémentaire est dans $\mathcal{F}$.

La famille des complémentaires des parties finies de **N** satisfait 1) et 2) mais pas 3). La famille des parties de **N** qui contiennent un élément a est un ultrafiltre dit dégénéré. Il existe (ça se montre à partir de l'axiome du choix) des ultrafiltres non dégénérés mais on ne sait pas les décrire. On démontre assez facilement qu'un ultrafiltre contient les complémentaires des parties finies.

Nous fixons dans ce qui suit un ultrafiltre $\mathcal{F}$ non dégénéré de **N**. On fait la construction suivante :

(1) On considère l'ensemble $H = \mathbf{R}^{\mathbf{N}}$ des suites (u_n) de nombres réels.

(2) On définit l'équivalence $\sim_{\mathcal{F}}$:

$$(u_n) \sim_{\mathcal{F}} (v_n) \Longleftrightarrow \{n : u_n = v_n\} \in \mathcal{F}$$

et on note $^*\mathbf{R}$ l'espace quotient $H / \sim_{\mathcal{F}}$.

(3) On identifie **R** à une partie de $^*\mathbf{R}$ et l'on "importe" dans $^*\mathbf{R}$ les opérations $(+ \, . \leq)$.

C'est exactement le même genre de construction que celui de **R** à partir de **Q**. On vérifie facilement que les propriétés de filtre font que $\sim_{\mathcal{F}}$ est bien une relation d'équivalence, ce qui donne le point 2). On injecte **R** dans $^*\mathbf{R}$ en associant au nombre réel u la classe de la suite constante $(u_n = u)$ et l'on définit les opérations sur $^*\mathbf{R}$ de la façon habituelle :

$$[(u_n)] + [(v_n)] = [(u_n + v_n)]$$

où $[(u_n)]$ désigne la classe d'équivalence de la suite (u_n).

Pour la relation d'ordre, une petite remarque. Supposons que pour relation d'équivalence nous ayons choisi $(u_n) \sim (v_n)$ si et seulement si $u_n = v_n$ à "partir d'un certain rang". L'ordre qui étend l'ordre de **R** sur les classes d'équivalence est alors $[(u_n)] \leq [(v_n)]$ si et seulement si $u_n \leq v_n$ à partir d'un certain rang ; on voit que l'ordre ainsi défini n'est pas total car, par exemple, la suite $(+1, -1, +1, -1, \ldots, \pm 1, \ldots)$ n'est ni plus grande ni moins grande que la suite $(0, 0, 0, \ldots)$. C'est la propriété 3) de l'ultrafiltre, qui veut que, pour toute partie, soit elle, soit son complémentaire, lui appartient, qui fait que l'ordre est total sur $^*\mathbf{R}$. Donc, selon que l'ensemble des entiers pairs est ou non dans l'ultrafiltre cet hyper-réel sera positif ou négatif, mais nous ne le saurons jamais.

C'est une conséquence du caractère tout à fait non constructif de la démonstration de l'existence d'un ultrafiltre qui fait que certains trouvent cette théorie peu satisfaisante. Mais cela n'explique pas le rejet de la plupart des mathématiciens qui ne se piquent généralement pas de constructivisme [13].

13. Faute de compétence je ne me suis pas penché sur cette question essentielle de l'ANS et du constructivisme qui à elle seule mériterait tout un article ; pour me faire pardonner je propose dans les commentaires bibliographiques quelques pistes sur ce sujet de plus en plus actuel.

Nous avons donc construit un corps ***R** qui contient **R** (ou plus exactement une image isomorphe de **R** via l'application $u \longmapsto [(u_n)]$) ; nous appellerons ces éléments de ***R** qui sont dans l'image de **R** des réels *standard.* Nous avons la situation suivante :

$$\forall u > 0 \in \mathbf{R} \ \ [(u_n = \frac{1}{n})] < [(u_n = u)]$$

En effet, l'ensemble des n tels que $\frac{1}{n} < u$ est le complémentaire d'une partie finie, donc un élément de $\mathcal{F}$. Nous avons donc un hyper-réel, l'élément $\varepsilon = [(u_n = \frac{1}{n})]$, qui n'est pas nul, mais qui est plus petit que tout hyper-réel standard. Nous dirons qu'il est *infiniment petit.* À partir de là nous pouvons développer toute une analyse basée sur l'existence d'infiniment petits dans l'ensemble ***R** des hyper-réels.

Malheureusement le travail n'est pas tout à fait terminé. Nous avons appris, par exemple, que pour tout nombre réel $x \in \mathbf{R}$:

$$cos^2 x + sin^2 x = 1$$

Cette égalité est-elle encore vraie pour un hyper réel $x \in {}^*\mathbf{R}$? La réponse est oui. Tout ce qui était vrai pour les nombres traditionnels peut être transféré aux nouveaux nombres et plus généralement tout ce qui était vrai dans l'ancien monde "standard " peut être *transféré* dans le nouveau monde "non-standard ". Cela n'est pas évident mais peut se démontrer.

2.4. *Quelle image mentale pour* ***R** *?*

La construction que nous venons de faire de ***R** n'est pas difficile ; elle ressemble comme deux gouttes d'eau à celle de **R** à partir de **Q** par les suites de Cauchy et elle nous fournit un cadre rigoureux pour manipuler les ordres de grandeur. Si la preuve de la possibilité du transfert est plus délicate, il est en revanche facile de l'admettre.

Alors comment expliquer que l'utilisation de ***R** soit encore confidentielle de nos jours dans la pratique mathématique ordinaire ?

Je propose l'explication suivante :

R *remplit complètement le continu naïf,*

qui est une sorte de retournement du slogan [14] de Reeb :

Les entiers naïfs ne remplissent pas **N**.

Nous avons une image naïve du continu géométrique telle que, comme nous l'avons vu, l'image mentale que nous nous faisons de **R** est celle d'un ensemble de points qui remplit totalement la droite géométrique. Toute la construction de **R** est faite pour que nous ayons cette image mentale, ce qui fait que nous ne pouvons plus y mettre les hyper-réels de ***R**. *Il n'y a plus de place* pour des nouveaux nombres sur la droite géométrique naïve une fois que nous y avons installé **R**.

Or, si nous voulons utiliser les nombres de ***R** de façon familière, nous devons nous en faire une image mentale. Puisqu'il n'y a plus de place sur la droite géométrique il faut en sortir. Une image possible qui vient à l'esprit est celle qui a bien marché pour les nombres complexes, celle d'un couple (x, ε) le premier élément est un réel ordinaire, le second un infiniment petit. Donc ***R** pourrait se voir comme la bande :

$$B = \mathbf{R} \times \,]-1\ ,+1[$$

la "fibre" $\{x\} \times \,]-1\ ,+1[$ au dessus de x représentant les nombres infiniment proches de x, mais cette vision rentre en contradiction avec notre vision de la *dimension*. Dans notre

14. Sur lequel je renvoie à [23] pour plus de détails.

image mentale classique une bande telle que B est un *continu à deux dimension* ; ainsi les réels standard de l'intervalle [0 , 1] seraient un continu à une dimension et les hyper-réels du même intervalle un continu à deux dimensions. Ça ne marche pas très bien. Je ne continue pas dans cette direction parce que, personnellement, je n'ai jamais trouvé dans la littérature d'image mentale que je trouve satisfaisante de $^*\mathbf{R}$, ni été capable de m'en construire une.

Cela ne discrédite cependant pas l'usage de $^*\mathbf{R}$ pour fonder l'ANS. On peut envisager deux solutions.

Une première est, à force d'exercices, de s'en forger petit à petit une image mentale ; tout objet mathématique s'apprivoise à force d'exercices et, de même que l'espace de Hilbert des fonctions de carré intégrable est parfaitement familier pour un mathématicien du niveau de la licence, $^*\mathbf{R}$ pourrait le devenir en se donnant le temps nécessaire.

Une autre solution, plus ambitieuse serait de réfléchir à ne pas former cette image mentale de $\mathbf{R}$ remplissant tout le continu naïf. Je dis plus ambitieuse et probablement déraisonnablement ambitieuse, car une grande partie de cette image mentale est déjà construite avant la fin de l'enseignement secondaire. Il faudrait donc modifier en profondeur l'enseignement élémentaire, voire pré-élémentaire des mathématiques. La réforme des "mathématiques modernes" nous a instruits sur les dangers de la hâte dans ce genre d'opération.

Je considère donc que, pour longtemps encore, nous devrons faire avec une représentation à la Cauchy-Cantor-Dedekin-Weirstrass du continu naïf. Heureusement, comme nous allons le voir, cela n'empêche en aucune façon la pratique de l'ANS.

3. — Un autre regard sur les mathématiques traditionnelles.

L'idée est de ne rien changer à toutes nos habitudes concernant les nombres, qu'ils soient entiers ou réels, et d'en garder l'image mentale que nous en avons. Il ne s'agit plus d'étendre les réels en des entités nouvelles mais d'enrichir le langage qui permet de parler de $\mathbf{R}$. Plus généralement nous gardons donc tous les objets mathématiques traditionnels tels que nous les connaissons mais, en revanche, nous en parlons dans une langue plus riche.

Mais que veut dire "enrichir" ? Pour le comprendre il faut avoir une idée un peu plus précise de ce que veut dire utiliser un langage formel, c'est-à-dire faire un tout petit peu de logique.

3.1. *Les mathématiques et la formalisation.*

Quand nous disons n est un "entier pair" nous disons que n possède la qualité P : "être un entier pair", que nous traduisons dans notre tête par :

$$P(n) \Longleftrightarrow \exists\ p \in \mathbf{N} : \quad n = 2p$$

Un autre exemple, la qualité $\prod(n)$: "être premier" :

$$\prod(n) \Longleftrightarrow \forall p\left((\exists q : pq = n) \Longrightarrow (p = 1 \lor p = n)\right)$$

on voit que, dans les deux cas, "avoir une qualité" pour l'entier n, c'est rendre vraie une formule dont n est une variable libre (non liée à un quantificateur) et où les objets mathématiques qui interviennent dans la formule sont clairement identifiés (appartenir à $\mathbf{N}$) ainsi que les opérations qui portent sur eux (multiplication, addition, ...). Les mêmes règles sont valables pour des objets plus compliqués que les nombres entiers. Par exemple pour une fonction f de $\mathbf{R}$ dans lui-même, la qualité B : "être bornée" :

$$B(f) \Longleftrightarrow \exists m \in \mathbf{R} : (\forall x \in \mathbf{R}\ |f(x)| \leq m)$$

Ici il est sous entendu que f est une "fonction" de $\mathbf{R}$ dans lui-même ce qui se définit à nouveau avec des symboles. Une "fonction" de $\mathbf{R}$ dans lui-même, est une partie F de $\mathbf{R} \times \mathbf{R}$ telle que :

$$\forall\, x \in \mathbf{R} \;\; \exists!\, y : (x, y) \in F$$

et ainsi de suite. Toute qualité d'un objet mathématique peut être au bout du compte traduite en une suite des symboles en nombre réduit ($+ \times \exists \forall \ldots$) dont la manipulation obéit à une syntaxe précise. C'est ce qu'on appelle un langage formel, notion que la pratique des langages de programmation nous a rendue familière. Inutile d'insister mais précisons un peu dans le cas qui nous intéresse de la théorie des ensembles.

3.2. *La théorie formelle* ZFC *et son langage* $\mathcal{L}$.

En théorie formelle des ensembles, on procède comme suit.

(1) On définit un langage. Pour cela on spécifie un alphabet,

$$a, b, c, \;\ldots\;, \neg \vee \wedge = \forall \exists$$

et des règles d'assemblage qui permettent de former des mots.

(2) *On définit des règles de déduction* qui permettent de passer (on dit inférer) d'une formule à une autre. Par exemple de $\neg(a \vee b)$ on a le droit d'inférer $(\neg a) \wedge (\neg b)$ (en langue naturelle : (non (a ou b)) entraîne ((non a) et (non b))

Ces deux premiers points qui spécifient le système logique dans lequel on entend travailler (en général le système dit du calcul des prédicats du premier ordre) sont communs à la plupart des théories formalisées. Pour la théorie des ensembles, on se donne en plus :

(3) un symbole spécifique : le prédicat à deux places $\in$, ($x \in y$ qui se lit "x est un élément de y") ce qui donne le langage $\mathcal{L}$.

(4) On donne une liste de formules réputées vraies : les axiomes.

Un théorème est démontré lorsqu'il se déduit des axiomes par application successive d'une des règles de déduction. Il existe d'innombrables ouvrages qui traitent de cette théorie (dite

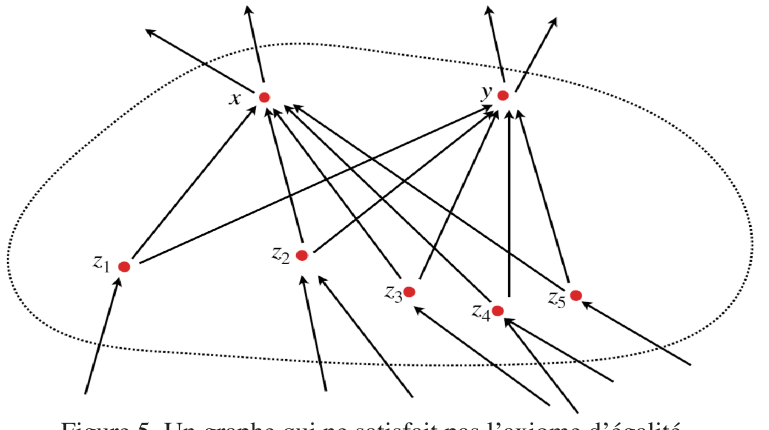


Figure 5. Un graphe qui ne satisfait pas l'axiome d'égalité.

de Zermelo Fraenkel avec Choix (ZFC)) et le lecteur qui voudrait avoir tous les détails est prié de se reporter à l'un d'eux (pour ma part j'ai appris (puis oublié) dans [20]) ici nous nous contenterons de quelques remarques.

On peut voir la théorie des ensembles comme un graphe dont les noeuds sont les ensembles et les arêtes sont des flèches de x vers y lorsque $x \in y$. Tout n'est pas permis ; par exemple on ne trouve nulle part dans le graphe la structure de la figure 5. parce qu'elle contredit l'*axiome d'égalité* :

$$\forall x \forall y \big(\forall z (z \in x \Longleftrightarrow z \in y) \Longrightarrow x = y\big)$$

qui se lit "deux ensembles qui ont les mêmes éléments sont égaux" ; sur la figure les noeuds (ensembles) différents x et y ont les mêmes éléments $z_1, \ldots, z_5$.

Un autre axiome :

$$\exists x (\forall y : y \notin x)$$

Cet ensemble qui ne contient pas d'éléments est l'ensemble vide Ø qui est le premier ensemble de toute la construction.

Voyons maintenant un axiome de ZFC qui est essentiel pour comprendre IST. C'est l'axiome :

$$\forall x \exists y \big(\forall z (z \in y \Longleftrightarrow (z \in x \wedge \Phi(x))\big)$$

qui dit que les ensembles z de x qui satisfont la formule (propriété) Φ constituent un ensemble. C'est l'axiome dit de *compréhension* qui donne le statut d'ensemble à la collection des éléments (d'un ensemble donné) qui satisfont une formule du langage $\mathcal{L}$.

Par exemple la propriété pourrait être

$$\Phi(x) \longleftrightarrow x \text{ est racine de } P(x) = x^7 - 2x^5 + 3$$

et nous écrivons, sans réfléchir :

$$E = \{x \in \mathbf{R} : P(x) = 0\}$$

que nous lisons *ensemble des x de* **R** *tels que* $P(x) = 0$, pour désigner l'ensemble des racines de P.

C'est la fameuse phrase : *on note F l'ensemble des x de E tels que* et patati et patata :

$$F = \{x \in E : \text{et patati et patata}\}$$

Et il va sans dire que le discours *et patati et patata* doit être écrit, si l'on travaille dans le cadre informel ordinaire, dans un langage mathématique précis, ou, quand on travaille dans le cadre formel, en une formule Φ autorisée par la langue $\mathcal{L}$.

Il est clair qu'on ne fait plus des mathématiques si l'on dit, par exemple : « nous considérons l'ensemble des nombres réels qui sont joyeux le matin et tristes après midi ». Si l'on travaille dans le cadre de la théorie des ensembles, tout ce qui est dit de façon informelle doit pouvoir, en principe, être dit dans le langage $\mathcal{L}$.

Supposons maintenant que l'on décide d'ajouter à $\mathcal{L}$ une nouvelle qualité, qui ne peut pas se réduire à une formule du langage. Ajoutons quelques axiomes et nous aurons une nouvelle théorie qui englobe la précédente, on dit que c'est une *extension* de $\mathcal{L}$. C'est ce que nous allons faire.

3.3. *Une extension de* ZFC *: le système* IST.

On peut donc dire que la théorie formelle des ensembles ZFC est un *graphe*, dont les noeuds sont en perpétuelle construction, et une *langue* qui permet d'énoncer des propriétés sur les noeuds, par exemple "tel noeud (pardon, ensemble) est fini" ; mais rien n'interdit de garder les noeuds tels qu'ils sont et d'enrichir le langage qui en parle.

Ainsi, dans son article de 1977 ([28]) Nelson [15] propose d'ajouter à la langue $\mathcal{L}$ de ZFC un prédicat unaire noté $st(x)$ qui se lit *x est standard*. Dans un instant nous serons amenés à dire :

$$\exists\ \omega \in \mathbf{N}\left(\forall n \in \mathbf{N}\ (st(n) \Longrightarrow n < \omega)\right)$$

qui se lit : "il existe un entier plus grand que tous les entiers standard".

Il faut bien comprendre que la formule "qui est plus grand que tous les entiers standard" ne peut pas être comprise dans ZFC car le mot *standard* n'appartient pas au langage $\mathcal{L}$ de ZFC. La langue $\mathcal{L}'$ de IST contient strictement la langue de ZFC. Mais IST ne contient pas d'objet mathématique nouveau, ce sont toujours les mêmes ensembles, c'est-à-dire les mêmes noeuds du graphe, et tout se passe comme si maintenant nous avions la possibilité de les colorier : en bleu les noeuds standard, en rouge les autres. A partir de maintenant nous dirons qu'une formule écrite dans le langage usuel $\mathcal{L}$ est une formule *interne* et une formule qui fait usage de *st* est une formule *externe*.

Standard est une qualité que possèdent ou non les ensembles alors qu'*interne* est une qualité qui porte sur les propositions. Il reste maintenant à donner les axiomes qui régissent l'emploi de *st*.

On utilise les abréviations suivantes :

$$\forall^{st}\, x \quad \longleftrightarrow \quad \forall\ x\ (x \text{ standard}) \Longrightarrow$$

$$\forall^{fin}\, x \quad \longleftrightarrow \quad \forall\ x\ (x \text{ fini})\ ) \Longrightarrow$$

$$\forall^{st\,fin}\, x \quad \longleftrightarrow \quad \forall^{st}\ x\ (x \text{ fini}) \Longrightarrow$$

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

15. Il faut signaler ici que, indépendamment et presque simultanément, Karel Hrbaceck publiait une présentation de l'ANS proche de celle de Nelson [19] ; je ne sais vraiment pas expliquer pourquoi la version de Nelson a eu plus de succès.

Insistons ici sur *fini* : Le mot *fini* est pris dans son sens mathématique usuel : un ensemble *x* est *fini* ssi toute injection de *x* sur lui-même est surjective. On a les trois axiomes suivants :

— ***Idéalisation*** : Soit B(x; y) une formule *interne* :

$$\forall^{st\,fin} z\ \exists x\ \forall y \in z\ B(x, y) \Longleftrightarrow \exists x\ \forall^{st} y\ B(x, y)$$

— ***Standardisation*** : Soit $C(z)$ une formule, interne ou non :

$$\forall^{st} x\ \exists^{st} y\ \forall^{st} z\ (z \in y \Longleftrightarrow z \in x \wedge C(z))$$

— ***Transfert*** : Soit $A(x, t_1, \ldots, t_k)$ une formule *interne* sans autre variable libre que $(x, t_1, \ldots, t_k)$. Alors :

$$\forall^{st} t_1 \ldots \forall^{st} t_k \left(\forall^{st} x\ A(x, t_1, \ldots, t_k) \Longrightarrow \forall x\ A(x, t_1, \ldots, t_k)\right)$$

et par contraposition :

$$\forall^{st} t_1 \ldots \forall^{st} t_k \left(\exists\ x\ A(x, t_1, \ldots, t_k) \Longrightarrow \exists^{st}\ x\ A(x, t_1, \ldots, t_k)\right)$$

d'où le nom donné à la théorie : IST, qui vaut aussi pour Internal Set Theory.

Reconnaissons que, pour qui n'y est pas habitué, de telles écritures ne sont pas du tout parlantes mais elles ne sont pas pires que celles qui interviennent dans l'écriture des axiomes de ZFC. Toutefois, le mathématicien ordinaire n'a pas présent à l'esprit les axiomes de la théorie des ensembles (à supposer qu'il les ait jamais connus) mais des conséquences, érigées à leur tour en axiomes, ou en tournures stylistiques que tout le monde comprend. Il en est de même pour tous ceux qui ont appris IST. À force d'exercices avec le langage ils ont acquis une "façon de parler" qu'ils savent être rigoureuse et ne recourent à la formalisation stricte que lorsqu'ils ont un doute. Ainsi celui qui pratique IST finit-il par oublier les axiomes et les remplace par des tournures qui font sens pour lui et que les praticiens de IST savent être correctes.

Mais il est vrai que ce n'est pas immédiat et que des exercices sont nécessaires. Nous ne ferons pas ces exercices [16] et je me contenterai ici d'expliquer à quoi servent ces axiomes.

Mais avant d'en arriver là, une question essentielle se pose. Qu'est-ce qui nous prouve que cette théorie n'est pas contradictoire ? Rien puisqu'elle contient ZFC dont nous savons qu'on ne peut pas en démontrer la non contradiction (Gödel). Cela veut-il dire qu'on peut étendre ZFC par n'importe quoi ? Certainement pas. On demande à une extension honnête de ZFC d'être conservative. Ceci veut dire que, si P est une proposition interne démontrée dans IST, alors il existe une démonstration de P dans ZFC, ce qui montre la consistance relative de IST à ZFC. En effet, s'il y avait une contradiction dans IST tous les énoncés seraient alors des théorèmes, en particulier $1 = 2$ qui, ayant une démonstration dans IST en aurait une dans ZFC qui de ce fait serait contradictoire. Dans [28] il est démontré que IST est conservative.

Si maintenant on se demande d'où viennent ces axiomes, on peut retourner à la théorie de Robinson où nous avions pu définir, sans sortir du langage de ZFC, des ensembles dits "internes" (par exemple des réels internes comme l'image dans $^*\mathbf{R}$ de $\mathbf{R}$ par l'injection canonique). Ce sont les propriétés de ces ensembles internes qui donnent naissance au système de Nelson.

> Les axiomes de la théorie des ensembles internes sont simplement les propriétés de base des ensembles internes dans l'approche traditionnelle de l'analyse non standard (...) [28].

16. L'article séminal ([28]) de Nelson qui est parfaitement accessible, même à celui qui n'a jamais fait sérieusement de logique, en propose. Il existe d'autre part de nombreuses et excellentes présentations didactiques de IST, par exemple [9, 10, 26, 31].

3.4. *Premières conséquences des axiomes* IST.

3.4.1. *Le transfert.*

Si nous appliquons l'axiome de transfert à la formule interne :

$$\exists\, x\, (\forall\, y : y \notin x)$$

il vient que cet x est standard. Or cet x est l'ensemble vide. Nous avons démontré que l'ensemble vide est standard. L'ensemble vide est la première brique de la construction ensembliste. A partir de lui on montrera que tous les objets mathématiques de la mathématique traditionnelle tels que $\mathbf{N}, \mathbf{Z}, \mathbf{R}, \ldots$ sont des objets standard (attention ça ne veut pas dire que les éléments de ces ensembles sont tous standard !).

Prenons maintenant la suite des ordinaux :

$$\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots , \{\ldots\{\emptyset,\{\emptyset\}\} \ldots \}, \ldots$$

que nous appelons

$$0, 1, 2, \ldots , n, \ldots$$

Toujours par l'usage de l'axiome de transfert nous obtenons que si n est standard alors $n + 1$ est standard.

3.4.2. *L'idéalisation.*

Il découle de l'idéalisation le :

> *Théorème.* Soit X un ensemble. Tous les éléments de X sont standard si et seulement si X est standard et fini.

Démonstration. (Traduction littérale de [28], p. 1167.) Soit $B(x, y)$ la formule $x \in X \wedge x \neq y$. Alors le membre de droite de l'axiome **I** est équivalent à $\exists x \in X \neg(x \text{ standard})$. En prenant la négation on a :

$$\begin{aligned}&\forall x \in X\,(x \text{ standard})\\ &\quad\Longleftrightarrow \exists^{st\,fin} z\, \forall x\, \exists y \in z(\neg(x \in X) \vee x = y)\\ &\quad\Longleftrightarrow \exists^{st\,fin} z\,(X \subset z)\end{aligned}$$

Si X est un ensemble standard et fini, tout élément de X est standard car nous pouvons prendre $z = X$. Réciproquement, si tout élément de X est standard alors $x \subset z$, ou z est un ensemble standard et fini. Donc $x \in P(z)$ où $P(z)$, l'ensemble des parties de l'ensemble des parties de z, est lui aussi fini. Ceci, en vertu de ce que nous avons déjà montré, prouve que X est standard et fini comme sous ensemble d'un ensemble fini. Q.E.D.

La compréhension de cette démonstration demande que l'on se concentre un peu et elle s'oublie vite. Peu importe, le théorème lui est clair. Nous en déduisons immédiatement qu'il existe dans **N** (puisque **N** est infini), un élément ω qui n'est pas standard et donc, puisque 0 est standard, 1 est standard, $st(n) \Longrightarrow st(n+1)$ nous avons :

$$0 < 1 < 2 \ldots < n < \ \ldots\ < (\omega - 1) < \omega < (\omega + 1), \ldots$$

Nelson utilise la métaphore de l'image en couleur : "être standard" c'est comme "être bleu", "être non standard" c'est comme "être rouge". On peut donc proposer l'image mentale :

$$0, 1, 2, \ldots, n, \ldots \ldots (\omega - 1), \omega, (\omega + 1), \ldots$$

en se rappelant que la frontière entre les bleus et les rouges n'est pas "nette" : il n'existe pas de dernier entier standard ni de premier entier non standard. Les entiers tels que ω qui sont plus grands que tous les entiers standard sont évidemment appelés des entiers infiniment grands. Il faut noter ici que les standard n'ont pas de plus grand élément, puisque le successeur d'un standard est standard, mais qu'on ne peut pas leur appliquer le principe de récurrence (puisque la qualité "être standard" (ou bleu) n'appartient pas à la mathématique traditionnelle).

3.4.3. *La standardisation.*

Lorsqu'une formule C est externe, elle n'appartient pas au langage de ZFC et donc nous ne pouvons pas appliquer l'axiome de *compréhension* et nous ne pouvons pas affirmer, pour un x donné, l'existence d'un ensemble des $z \in x$ tels que $C(z)$. Nous ne pouvons plus affirmer que :

$$\forall x\, \exists y\, \forall z\,\big(z \in y \Longleftrightarrow z \in x \wedge C(z)\big)$$

mais nous pouvons le faire à condition de ne pas sortir des standard :

$$\forall^{st} x\, \exists^{st} y\, \forall^{st} z\,\big(z \in y \Longleftrightarrow z \in x \wedge C(z)\big)$$

De cet axiome nous pouvons déduire immédiatement que si E est un sous-ensemble d'un ensemble standard (par exemple **R**) il existe un unique ensemble standard ${}^{S}E$ dont les éléments standard sont exactement les éléments standard de E :

$$\forall^{st} x \ \ (x \in E \Longleftrightarrow x \in {}^{S}E)$$

Le standardisé d'un ensemble peut être vide, comme par exemple, dans **N** le standardisé de $[n, n+1, \ldots\ ]$ lorsque n est infiniment grand.

Un peu plus subtil que le standardisé, nous avons l'*ombre* d'un ensemble qui se définit ainsi dans **R** (mais qui pourrait l'être dans n'importe quel espace métrique).

Si E est un sous-ensemble de **R** on dit que "x est dans le *halo* de E s'il existe y dans E tel que $x - y$ est infiniment petit. On notera $x \approx y$ lorsque $x - y$ est infiniment petit". L'*ombre* de E, notée ${}^{\circ}E$ est le standardisé du halo ; c'est l'ensemble standard tel que :

$$\forall^{st} x\,\big(x \in {}^{\circ}E \Longleftrightarrow \exists y \in E \wedge x \approx y\big).$$

Grace à l'axiome de standardisation on

montre que tout nombre réel limité x est infiniment proche d'un unique réel standard $^{\circ}x$. L'unicité est évidente ; l'existence tient au fait que **R** est complet (tout sous-ensemble majoré non vide possède une borne supérieure). En effet, si x est limité le standardisé ^{S}E de l'ensemble :

$$E = \left\{ t \in \mathbf{R} : t < x \right\}$$

est un ensemble majoré non vide, il possède donc une borne supérieure. La borne supérieure de ^{S}E est la borne supérieure d'un ensemble standard, donc elle est standard ; on vérifie aisément qu'elle est infiniment proche de x.

3.5. *Une image mentale pour terminer.*

Sur la partie gauche de la figure 6, on a dans l'encadré de droite l'univers $\mathcal{U}$ des mathématiques traditionnelles et dans celui de gauche l'univers $^*\mathcal{U}$ construit à partir de $\mathcal{U}$ en prenant des suites d'éléments modulo un ultrafiltre. Les flèches de couleur bleues *ne représentent pas des appartenances* entre éléments des deux univers *mais des morphismes* entre objets structurés comme les injections canoniques de **N** dans ***N** et **R** dans ***R**[17]. A droite, l'univers de Nelson, est l'univers des mathématiques traditionnelles, mais coloré par l'usage du prédicat $st(x)$. Il n'y a pas de nouveaux nombres réels, infiniment petits ou grands. Ils étaient déjà présents dans la droite réelle mais nous ne le savions pas. Nelson conclut [28] par :

> Un fois que l'on sera remis du choc d'avoir été informé que les infinitésimaux et autres éléments idéalisés étaient déjà présents dans les ensembles avec lesquels nous sommes familiers, et que l'on aura appris à éviter la formation illégale d'ensembles, on trouvera notre approche très facile à utiliser. Tous les résultats familiers sont disponibles sans chan-

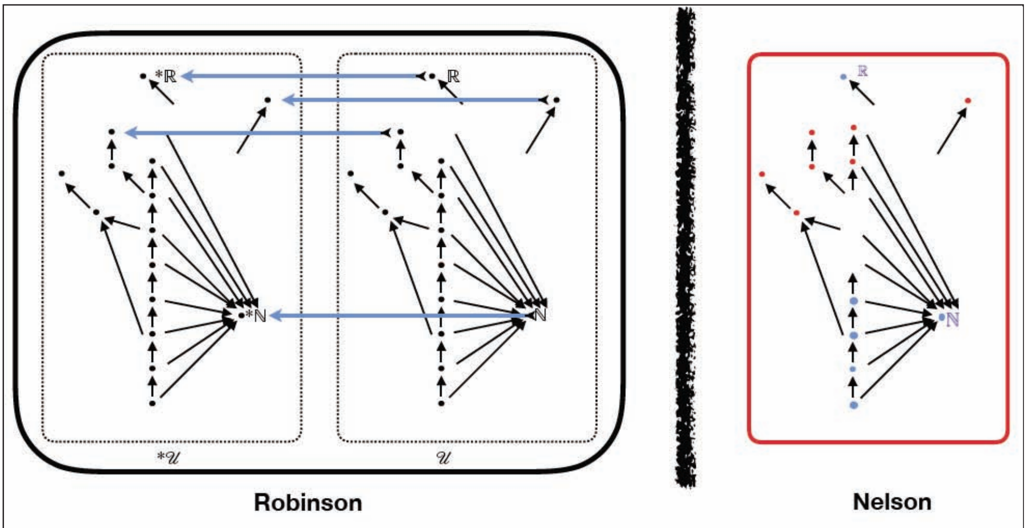


Figure 6. Les mondes de Robinson et de Nelson

17. On remarquera ici une ressemblance avec la théorie des catégories. Il y a effectivement des liens mais c'est une autre histoire.

gement ; ils ne doivent pas être étendus à un nouveau système. (...)

C'est ce que nous allons immédiatement vérifier.

4. — Un peu de pratique ANS

Nous nous plaçons dans toute cette section dans le cadre de IST. Quand nous utilisons les mots "suite" ou "fonction" c'est dans le sens précis que ces mots ont dans ZFC, ce qui veut dire que le discours qui les définit doit être un discours *interne*. Il n'y a pas de sens dans IST à parler de la fonction qui à x infiniment petit associe 0 et 1 sinon.

4.1. *Vocabulaire.*

Il existe donc dans $\mathbf{N}$ un entier ω tels que

$$\forall n\ (st(n) \Longrightarrow n < \omega)$$

Un tel entier mérite de s'appeler *infiniment grand*. Maintenant il faut s'habituer à ce que :

L'ensemble : $[0, 1, \ldots, n]$
est toujours un *ensemble fini* (au sens classique, qui n'est pas en bijection avec une de ses parties propres) même lorsque l'entier n est *infiniment grand* au sens de IST qui vient d'être défini.

C'est évidemment un problème terminologique sérieux. A. Deledic avait bien cerné cette difficulté dans [8] où il évitait de dire "infiniment grand" ; il avait même proposé, *idéalement grand* qui est mieux adapté et reste compatible avec *i.g* qui peut se lire indifféremment *infiniment-grand* ou *idéalement-grand*. Malheureusement les mauvaises habitudes étaient prises et il n'a pas eu de succès.

Comme les réels sont construits de proche en proche à partir des entiers on imagine comment on pourrait définir des réels standard mais ce ne sera pas utile ici. Avec la qualité $st(n)$ on a tout ce qu'il faut pour définir le vocabulaire de l'analyse infinitésimale. En effet :

— Un réel x qui est plus grand que tout entier standard sera dit *infiniment grand* ;

$$x \text{ i.g.} \Longleftrightarrow \forall^{st} n : n < |x|$$

— un nombre qui n'est pas infiniment grand est *limité* ;

— L'inverse d'un *infiniment grand* sera dit *infiniment petit* ;

$$x \text{ i.p.} \Longleftrightarrow \forall^{st} n : |x| < \frac{1}{n}$$

— $x \approx y$ (x est *infiniment proche* de y) équivaut à $|x - y|$ est *infiniment petit*.

4.2. *Les séries.*

Si n est un entier la somme : $\sum_{i=0}^{\omega} u_i$ est un symbole formel défini de façon récursive :

$$\sum_{i=0}^{0} u_i = u_0 \qquad \sum_{i=0}^{n+1} u_i = \Big(\sum_{i=0}^{n} u_i\Big) + u_{n+1}\,.$$

Tout ce qui est dit est parfaitement interne, donc le symbole $\Sigma_{i=0}^{n} u_i$ pour tout n de $\mathbf{N}$, fût-il non standard ou non. Le symbole $\Sigma_{i=0}^{n} u_i$ désigne un nombre réel parfaitement défini au sens de la mathématique traditionnelle ; maintenant il peut être standard ou pas, limité ou pas. Par exemple :

$$\Big(\sum_{i=1}^{n} \frac{1}{i}\Big)$$

est un réel infiniment grand dès que n est infiniment grand, alors que :

$$\Big(\sum_{i=1}^{n} \frac{1}{2^i}\Big)$$

est limité quel que soit n standard ou non.

Mais attention, le fait que $\Sigma_{i=0}^{\omega} u_i$ avec ω infiniment grand soit toujours défini ne veut pas dire qu'il n'y a plus de théorie des séries. Il y a toujours des séries convergentes. Ce sont celles pour lesquelles les valeurs des sommes partielles $S_n = \Sigma_{i=0}^{n} u_i$ tendent vers une limite et maintenant, on dira que S_n tend vers une limite l ssi :

$$n \text{ infiniment grand} \Longrightarrow S_n \approx l.$$

4.3. *La S-continuité.*

Soit $x \longmapsto f(x)$ une fonction de $\mathbf{R}$ dans lui-même.

Définition. On dit que f est *S-continue* en un point $x \in \mathbf{R}$ si $\forall x \in \mathbf{R}$

$$\big(h \approx 0 \Longrightarrow (f(x+h) - f(\mathrm{x})) \approx 0\big)$$

Théorème. Soit f une fonction standard. Elle est continue en tout point si et seulement si elle est S-continue en tout point standard.

Démonstration. Montrons que S-continu entraîne continu. Soit x standard. Par hypothèse :

$$h \approx 0 \Longrightarrow (f(x+h) - f(\mathrm{x})) \approx 0$$

Par définition de $\approx$

$$\forall^{st} m \left(h \approx 0 \Longrightarrow |f(x+h) - f(\mathrm{x})| < \frac{1}{m}\right)$$

ce qui entraîne (en prenant un entier n tel que $1/n \approx 0$) :

$$\forall^{st} m \left(\exists n \forall h \left(h < \frac{1}{n} \Longrightarrow |f(x+h) - f(\mathrm{x})| < \frac{1}{m}\right)\right)$$

Dans la parenthèse qui suit m les deux seules variable libres sont f et x qui sont supposées standard, donc, en utilisant l'axiome de transfert nous avons :

$$\forall m \, \exists n \forall h \left(h < \frac{1}{n} \Longrightarrow |f(x+h) - f(\mathrm{x})| < \frac{1}{m}\right)$$

qui est une définition de la continuité au point standard x et, à nouveau par transfert, nous pouvons lever cette restriction sur x. Q.E.D.

Nous laissons au lecteur qui le souhaite le soin de montrer que la continuité traditionnelle entraîne la S-continuité en tout point standard.

Nous retiendrons de cette démonstration que pour pouvoir passer d'une définition externe (la S-continuité) à son homologue interne (la continuité) il faut faire usage de l'axiome de transfert.

D'une manière générale on peut définir toutes sortes de S-propriétés, qui appliquées à des objets standard, sont équivalentes à une propriété classique. Ainsi :

— l est la S-limite de la suite (u_n) ssi n i.g. $\Longrightarrow u_n \approx l$. Si (u_n) est standard, l est la S-limite de u_n ssi $lim_{n \longrightarrow \infty} u_n = l$

ou encore

— (u_n) est S-Cauchy ssi p, q i.g. $\Longrightarrow u_p \approx u_q$

Sur cette dernière notion on voit bien l'économie en termes de quantificateurs puisque être de Cauchy est, par définition :

$$\forall \varepsilon > 0 \; \exists n : p > n, q > n \Longrightarrow |u_p - u_q| < \varepsilon$$

Cette économie est une des vertus que T. Tao [18] attribue à l'ANS :

> Les arguments en analyse "dure" sont connus pour leur profusion "d'epsilons et deltas". Dans les arguments les plus sophistiqués de ce type, on peut finir par avoir toute une armée d'epsilons ε_1, ε_2, ε_3, ... qu'il faut gérer, [...] Cet art de la gestion des epsilons, une fois maîtrisé, n'est pas terriblement difficile [...] Pour ceux qui pratiquent l'analyse "dure" pour vivre (comme moi), il est natu-

18. Sur son blog du 25 juin 2007 : https://terrytao.wordpress.com/2007/06/25

rel de se demander si l'on peut en quelque sorte "nettoyer" ou "automatiser" toute la gestion des epsilons que l'on est obligé de faire. [...] Une étape importante dans cette direction a été le développement de divers types de notation asymptotique, tels que la notation Hardy utilisant des constantes non spécifiées C, la notation Landau utilisant O() et o(), [...] Mais si nous essayons de formaliser cela en essayant de créer l'ensemble $A := \{x \in \mathbf{R} : x = O(1)\}$ de tous les nombres bornés, et en affirmant que cet ensemble est alors fermé sous addition et multiplication, nous disons un non-sens ; la notation O() ne peut pas être utilisée dans le schéma axiome de la spécification, et donc la définition ci-dessus de A n'a pas de sens.

Il existe cependant un moyen de rendre des concepts tels que "l'ensemble de tous les nombres bornés" précis et pleins de sens en utilisant l'analyse non standard, [...].

4.4. *Dérivée et intégrale.*

Voyons maintenant la notion de dérivée. Soit $f : I \longrightarrow \mathbf{R}$ une fonction standard et x un élément standard de l'intervalle I.

Définition. On dit que le nombre standard $f'(x)$ est la S-dérivée de f au point x si, pour tout $x_1 \approx x$ et tout $h \approx 0$ on a :

$$(1) \qquad \frac{f(x_1 + h) - f(x_1)}{h} \approx f'(x)$$

On a à nouveau une S-notion dont on vérifie facilement l'équivalence avec la notion traditionnelle de dérivée.

On peut alors développer aisément les règles et les résultats de base du calcul infinitésimal (dérivée d'une somme, d'un produit, etc.) jusqu'à la formule de Taylor en quelques pages (moins de trois pages dans [28]).

L'intégrale d'une fonction continue sera :

Définition. Soit f continue ; pour f, a et b standard et pour $h \approx 0$ on pose :

$$(2) \qquad \int_a^b f(\theta)d\theta = {}^\circ\Big(\sum_{j\,:\,a \le jh < b} f(jh)h\Big)$$

L'absence de h dans la notation avec le signe $\int$ laisse supposer que le membre de droite ne dépend pas de h, pourvu qu'il soit infiniment petit ou, ce qui revient au même étant donné la définition de l'ombre d'un réel, que :

$$(3) \quad h_1 \approx 0\ ,\ h_2 \approx 0 \Longrightarrow$$

$$\sum_{j\,:\,a \le jh_1 < b} f(jh_1)h_1 \approx \sum_{j\,:\,a \le jh_2 < b} f(jh_2)h_2$$

ce qui se vérifie immédiatement. A partir de là on démontre évidemment la :

Proposition. Si f est une fonction standard continue sur I ; pour a et x dans I, on pose :

$$F(x) = \int_a^x f(t)dt$$

Alors, pour tout x, F est dérivable et :

$$F'(x) = f(x)$$

4.5. *Le théorème d'existence des solutions d'une équation différentielle.*

Soit $x \longmapsto f(x)$ une fonction continue bornée de $\mathbf{R}$ dans lui-même.

Théorème. Pour tout x_0 dans $\mathbf{R}$ il existe une fonction dérivable $t \geq 0 \longmapsto x(t)$ telle que $x(0) = x_0$ et $x'(t) = f(x(t))$.

Démonstration (adaptée de Nelson [28]). Par transfert on peut toujours se ramener à f et x_0 standard. Puisque f est bornée, il existe C tel que

$\forall x \in \mathbf{R}\ |f(x)| < C$ et puisque f est standard, on peut supposer C standard, donc limité.

Soit dt fixé infiniment petit. Notons $\mathbf{R}^+_{dt}$ le sous- ensemble de $\mathbf{R}$ constitué des points kdt pour $k \in \mathbf{N}$.

$$\mathbf{R}^+_{dt} = \{0, 1dt, 2dt, \dots, ndt, \dots\}$$

Considérons la fonction définie sur $\mathbf{R}^+_{dt}$ par :

$$\xi(0) = x_0$$

$$\xi((k+1)dt) = \xi(kdt) + dtf(\xi(kdt))$$

On reconnaît bien entendu le classique schéma d'Euler. La fonction $ndt \longmapsto \xi(ndt)$ est S-continue. En effet, soient $m < n$ tels que $(n - m)dt$ soit infiniment petit. 0n vérifie immédiatement que :

$$|\xi(ndt) - \xi(mdt)| \leq (n - m)dtC$$

Puisque C est limité, si $(n - m)dt$ est infiniment petit il en est de même de $(n - m)dtC$ et donc $ndt \longmapsto \xi(ndt)$ est bien S-continue. De la même manière, puisque x_0 est standard, donc limité, on voit que $\xi(ndt)$ est limité.

Soit (par standardisation) $t \longmapsto x(t)$ la fonction qui à tout standard t de $\mathbf{R}^+$ associe $°\xi(kdt)$ où k est le plus grand entier tel que $kdt < t$. On a :

$$x(t) \approx \xi(\mathrm{k}dt) = x_0 + \sum_{j=0}^{k} f(\xi(jdt))$$

$$\approx x_0 + \int_0^t f(x(\theta))d\theta$$

et comme x_0 et $x_0 + \int_0^1 f(x(\theta))d\theta$ sont tous deux standard, ils sont égaux.

Ce qui démontre le théorème.Q.E.D.

On voit ici que, la partie difficile dans la démonstration traditionnelle : le théorème d'Ascoli, est totalement absente de cette démonstration, ce qui fait dire à Nelson dans [28] :

> La croyance largement répandue selon laquelle on ne ne peut pas obtenir quelque chose en échange de rien est une superstition.

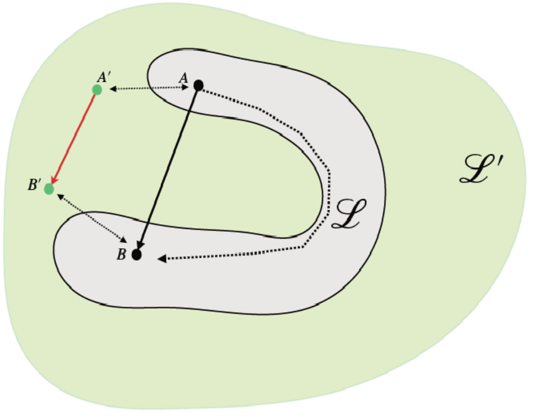


Figure 7. Formulations internes et externes.

5. — Le grand malentendu

Voici deux extraits, toujours de l'article de Nelson [28] que je vais commenter à l'aide de la figure 7.

> § 2. (...) Certains des théorèmes que nous prouvons sont externes. Pouvons-nous les reformuler pour qu'ils deviennent internes ? Les définitions d'objets standards faites au moyen de (S) peuvent impliquer des notions externes. Pouvons-nous trouver des formulations internes équivalentes de telles définitions ? Nous montrerons que la réponse aux deux questions est oui (...)

> § 8. Cette annexe est consacrée à montrer que IST est une extension conservatrice de ZFC ; c'est-à-dire que toute proposition interne qui peut être prouvée dans IST peut être prouvée dans ZFC.

Sur la figure, ce qui peut être dit de façon interne à ZFC est représenté en gris et est inclus dans tout ce qui peut être dit dans IST qui est représenté en vert. La première citation nous dit que les énoncés externes $A' \Longrightarrow B'$ peuvent être traduits en des énoncés internes $A \Longrightarrow B$ équivalents. La seconde citation nous dit que, s'il existe une preuve externe de $A \Longrightarrow B$ (la flèche noire de A vers B) alors il en existe une interne, la flèche pointillée qui ne quitte pas le gris. De cette présentation des choses on retire l'impression que l'ANS n'apporte rien puisque, si $A' \Longrightarrow B'$ est un théorème dans IST il existe un théorème classique qui lui est formellement équivalent. A partir de là la discussion va porter sur le prix à payer pour des démonstrations plus simples. On admettra volontiers que pour des démonstrations très techniques, comme celles auxquelles Tao fait allusion sur son blog [19], le prix du ticket d'entrée dans l'ANS est acceptable, mais pour des mathématiques plus ordinaires on hésitera à faire l'investissement. Malheureusement, cette présentation ne rend pas justice à l'ANS car, même si son pouvoir effectif dans la simplification de preuves complexes est réel, même si elle a permis de résoudre des conjectures (voir figure 8), ou de découvrir des phénomènes nouveaux comme les solutions "canards" de certaines équations différentielles qui peuvent être étudiés par des méthodes classiques ou non [20], le propre de l'ANS est de permettre de *parler directement d'objets non standard*. Voyons sur quelques exemples ce que je veux dire.

5.1. *La S-continuité n'est pas la continuité.*

Sur la figure 9 j'ai tracé les graphes de deux fonctions non standard :

— *Une fonction S-continue qui n'est pas continue :*

La fonction :

$x \geq 0 \longmapsto f_{dt}(x) = kdt$ si $x \in [kdt, (k+1)dt[$

dont le graphe est un escalier discontinu en tous les points d'abscisse kdt ; cette fonction est *S*-continue dès que dt est infiniment petit, donc non standard, mais elle n'est jamais continue en les points kdt.

— *Une fonction continue qui n'est pas S-continue*

La fonction :

$$x \longmapsto th(\omega x) = \frac{e^{\omega x} - e^{\omega x}}{e^{\omega x} + e^{\omega x}}$$

est continue pour toute valeur de ω et pour ω infiniment grand (100 sur la figure), elle n'est pas *S*-continue en 0 où elle "saute brusquement" de la valeur -1 à la valeur $+1$ alors que la fonction $f_{dt}(x)$ ne fait que des "petits sauts" de valeur infinitésimale.

19. https://terrytao.wordpress.com/2007/06/25

20. voir l'article de Scholarpedia Canards http://www.scholarpedia.org/article/Canards

PACIFIC JOURNAL OF MATHEMATICS
Vol. 16, No. 3, 1966

SOLUTION OF AN INVARIANT SUBSPACE PROBLEM
OF K. T. SMITH AND P. R. HALMOS

ALLEN R. BERNSTEIN AND ABRAHAM ROBINSON

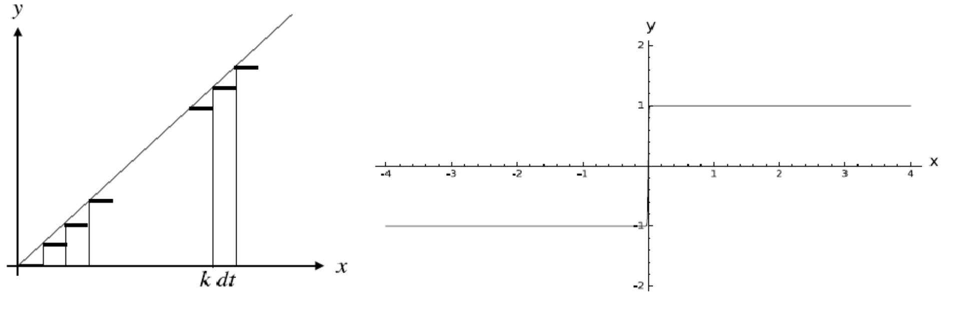
The following theorem is proved.

Let T be a bounded linear operator on an infinite-dimensional Hilbert space H over the complex numbers and let $p(z) \neq 0$ be a polynomial with complex coefficients such that $p(T)$ is completely continuous (compact). Then T leaves invariant at least one closed linear subspace of H other than H or $\{0\}$.

For $p(z) = z^2$ this settles a problem raised by P. R. Halmos and K. T. Smith.

The proof is within the framework of Nonstandard Analysis.

Figure 8. Le théorème de Bernstein-Robinson : Dans cet article du Pacific Journal of Mathematics, Bernstein et Robinson proposent la réponse à une question posée par Smith et Halmos. Halmos devait publier dans le même numéro du journal une preuve conventionnelle imitant la preuve nonstandard. Certains pensent que Halmos, qui était éditeur de la revue, ne s'est pas conduit de façon irréprochable à cette occasion (voir [5]).



$f_{dt}(x)$ $\quad$ $\mathrm{th}(\omega x) \qquad \omega = 100$

Figure 9. *S*-ontinuité et continuité

La définition traditionnelle de la continuité exprime la complétion de **R** à travers le théorème de la valeur intermédiaire : pour passer de – 1 à +1 on passe par toutes les valeurs intermédiaires. La fonction $f_{dt}(x)$ n'a pas cette propriété mais en revanche elle ne fait que des sauts infinitésimaux.

5.2. *Les systèmes différentiels nonstandard.*

La dépendance continue des solutions d'une équation différentielle pas rapport aux conditions initiales est une forme de continuité pour laquelle la *S*-continuité diffère aussi de la continuité. Considérons, par exemple, dans $\mathbf{R}^2$ le système différentiel :

$$\frac{dx}{dt} = f(x) \tag{4}$$

où f est une fonction de classe C^1 et telle que les solutions soient définies pour tout t.

Le théorème de dépendance continue des solutions dit :

Théorème. Soit $x(t, \xi, f)$ la valeur à l'instant t de la solution de (4) telle que $x(0) = \xi$. Notons :

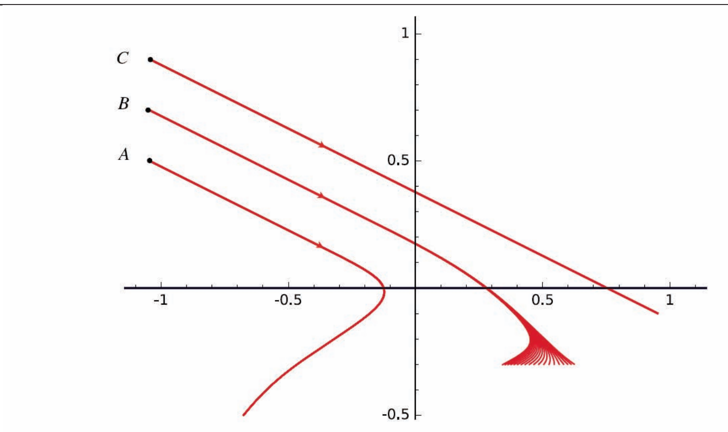
$$\delta_t(\xi, \xi_0, f) = \max_{s \in [0,1]} \| x(s, \xi, f) - x(s, \xi_0, f) \|$$

Alors pour tout t la fonction $\xi \longmapsto \delta_t(\xi, \xi_0, f)$ est continue en ξ_0.

La figure 10 “illustre” bien mal ce théorème à travers une simulation des trajectoires du système :

$$\Sigma_{10} \begin{cases} \dfrac{dx_1}{dt} = 2th\,(10\,(x_1^2 + x_2)) \\ \dfrac{dx_2}{dt} = -1 \end{cases} \tag{5}$$

En effet. Sur la figure on observe, issues res-



$A = \{(-1.000000, 0.5 + k10^{-4}) : k = [-10, -9, \cdots 9, 10]\}$

$B = \{(-1.046804, 0.7 + k10^{-4}) : k = [-10, -9, \cdots 9, 10]\}$

$C = \{(-1.000000, 0.9 + k10^{-4}) : k = [-10, -9, \cdots 9, 10]\}$

Figure 10. Dispersion de trajectoires du système (5)

pectivement des points A et C, deux trajectoires de (5) ; en réalité il s'agit pour chaque point de 20 trajectoires d'une durée de 1 unité de temps, issues de 20 points très proches les uns des autres (dans un disque de rayon 10^{-3}), mais à l'échelle de représentation choisie, elles ne sont pas discernables. C'est bien la "dépendance continue" des solutions par rapport à la condition initiale qui est observée. Mais les 20 trajectoires issues d'un voisinage du point B de même rayon 10^{-3} restent indiscernables pendant une durée d'environ 0,8 unité de temps, puis se dispersent. Il est assez difficile de décrire ce phénomène de façon classique ; pour le faire on considère, non plus le système différentiel (5), mais toute une famille de systèmes :

$$(6) \quad \Sigma_\omega \quad \begin{cases} \dfrac{dx_1}{dt} = 2th\,(\omega\,(x_1^2 + x_2)) \\ \dfrac{dx_2}{dt} = -1 \end{cases}$$

indexée par un paramètre ω que l'on fait tendre vers l'infini. Une façon de traduire le phénomène de dispersion sera alors :

Définition. Dispersion : Soit ξ_0 une condition initiale. On dit que T est un temps de dispersion (pour ξ_0) si :

$$\exists(\omega_n) \longrightarrow +\infty, \exists(\xi_n) \longrightarrow \xi_0, \exists(t_n) \longrightarrow T, \exists\alpha > 0$$

tels que :

$$\delta_T(\xi_n, \xi_0, \Sigma_{\omega_n}) \longrightarrow 0$$

$$||x(t_n, \xi_n, \Sigma_n) - x(t_n, \xi_0, \Sigma_n)|| > \alpha$$

Si l'on accepte de rester dans le cadre de formulations externes, le même phénomène sera idéalisé de la façon suivante :

Définition. Soit le système (4) (avec f éventuellement non standard). On dit que T limité est un temps de dispersion de ξ_0 limité si :

(1) $\xi \approx \xi_0 \Longrightarrow \delta_T(\xi, \xi_0, f)\;\; 0$

(2) $\forall^{st}\lambda > 0\;\; \exists t \in [T, T+\lambda]$
$\exists\xi \approx \xi_0 : x(t, \xi, f) \not\approx x(t, \xi_0, f)$.

Et même mieux, de manière informelle, mais encore plus parlante :

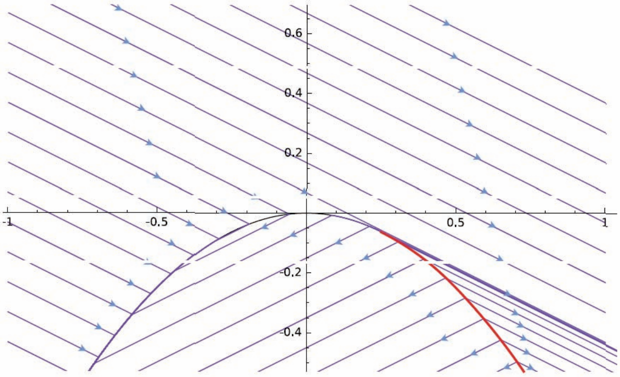


Figure 11. Dispersion de trajectoires du système (6) avec ω infiniment grand ; en rouge, les poins de dispersion.

Définition. le temps de *dispersion* de x_0 est le premier instant où la fonction $\xi \longmapsto x(t, \xi, f)$ cesse d'être S-continue en x_0.

Du théorème de dépendance continue on peut déduire que si f est un champ standard alors il n'existe pas de points de dispersion à distance limitée. Cette définition de la dispersion, qui traduit bien le phénomène observé — jusqu'au temps T les trajectoires de condition initiales infiniment proches restent infiniment proches (condition 1) mais pas au-delà (condition 2) — n'a donc d'intérêt que pour les systèmes différentiels non standard.

Le système simulé Σ_{10} est une illustration de la démonstration (facile) de l'existence d'un point de dispersion (le point où la droite de pente $-\frac{1}{2}$ est tangente à la parabole $x_1^2 + x_2 = 0$) pour Σ_ω lorsque ω est infiniment grand ; l'existence de ce point découle évidemment de ce que la fonction $(x_1^2, x_2) \longmapsto \mathrm{th}(\omega(x_1^2 + x_2))$ n'est pas S-continue le long de la parabole ce qui fait que, au dessus, les trajectoires sont infiniment proches de droites de pente $-\frac{1}{2}$ et, au dessous, infiniment proches de droites de pente $+\frac{1}{2}$; la figure 11 est le portrait de phase “idéalisé” de (6) lorsque ω est infiniment grand. Tout ceci est assez trivial, mais pratiquement indicible dans le langage traditionnel.

Dans la théorie qualitative des systèmes différentiels standard on définit divers éléments caractéristiques : équilibre stable, équilibre instable, cycle limite, attracteur, etc. autour desquels on organise la description des trajectoires. Pour les systèmes non standard il faudra ajouter les points de dispersions. Il faut remarquer que non standard ne veut pas dire pathologique ; le système (5) n'a rien d'extraordinaire.

5.3. *Les marches de pas infiniment petit.*

Au paragraphe précédent nous proposions d'utiliser un langage nouveau pour parler d'un objet classique : les solutions d'une équation différentielle. Maintenant nous allons un peu plus loin en reconsidérant notre vision de ce qu'est une équation différentielle.

5.3.1. *Le continu-discret.* Nous pouvons remplacer la notion traditionnelle de fonction définie sur l'intervalle continu [0 ; 1] par une notion discrète : on se donne un ω infiniment grand et l'on pose $dx = 1/\omega$. L'intervalle [0; 1] est "discrétisé" en les points :

$$[0, 1]_\omega = \{0, dx, 2dx, \dots , ndx, \dots , \omega dx = 1\}$$

et une "fonction" sur $[0; 1]_\omega$ est une suite

$$f(0), f(dx), f(2dx), \dots , f(ndx), \dots , f(\omega dx)$$

Cet objet est familier à qui a fait un tout petit peu de programmation (donc, avec le temps qui passe, à tout le monde). Pour tracer le graphe d'une fonction on fait afficher à l'ordinateur la suite des points $(ndx, f(ndx))$. Avec les infinitésimaux on peut parler utilement de "continuité" pour de telles fonctions alors que la définition classique est sans intérêt (toutes les fonctions sont continues pour la topologie discrète) :

Définition. La fonction f est S-continue sur $[0, 1]_\omega$ si et seulement si :

$$mdx \approx ndx \Longrightarrow f(mdx) \approx f(ndx)$$

On démontre immédiatement le S-théorème de la valeur intermédiaire :

Théorème. Soit f une fonction S-continue sur $[0, 1]_\omega$ telle que $f(0) < 0$ et $f(1) > 0$. Alors il existe n tel que $f(ndx) \approx 0$.

Preuve. En effet considérons l'ensemble E des entiers tels que $f(ndx) < 0$. Il est non vide, donc possède un plus grand élément n^* qui d'après l'hypothèse est strictement plus petit que ω et $f((n^* + 1)dx) \geq 0$. Comme d'après la S-continuité $f(n^*dx) \approx f((n^* + 1)dx)$ ces deux dernières valeurs sont nécessairement infiniment petites. QED

On voit bien ici que le S-théorème de la valeur intermédiaire (qui est une trivialité) ne dit pas tout à fait la même chose que le "vrai" théorème de la valeur intermédiaire pour les fonctions de $\mathbf{R}$ dans $\mathbf{R}$ qui, lui, exprime que $\mathbf{R}$ représente bien le continu géométrique dans lequel, pour aller de A à B, on passe par toutes les valeurs intermédiaires.

L'analyse fonctionnelle traditionnelle est l'étude de l'espace vectoriel de dimension infinie (au sens traditionnel) des fonctions définies sur [0, 1]. L'analogue discret-continu sera l'étude de l'espace vectoriel de dimension finie $\mathbf{R}^{\omega}$. Pour cet espace toutes les normes sont équivalentes, on a toujours $||x||_1 \leq C\,||x||_2$ mais, bien entendu, ce n'est pas la même chose selon que la constante C est, ou non, infiniment grande.

5.3.2. *Marches de* $\mathbf{R}^n$ *de pas infiniment petit.*

On se donne une application f de $\mathbf{R}^n$ dans lui-même qui ne satisfait *a priori* aucune condition de régularité. Tout ce qui est demandé est qu'à chaque x donné la fonction f "renvoie" une valeur bien définie $f(x)$. On se donne un $dt > 0$ infiniment petit :

$$(7) \qquad \Sigma_{dt} \begin{cases} x(0) & = x_0 \\ x(t + dt) & = x(t) + dt\, f(x(t)) \end{cases}$$

La règle qui vient d'être écrite définit clairement une unique fonction $t \longmapsto x(t)$ pour t appartenant à l'ensemble discret :

$$0, dt, 2dt, \ldots , ndt, (n + 1)dt, \ldots$$

C'est la *marche de pas dt* de condition initiale x_0 associée à f. Si f est majorée par une constante C on a immédiatement que

$$|x((n + k)dt) - x(ndt)| \leq Ck\, dt$$

et donc, si dt est infiniment petit et C limité, $n \longmapsto x(ndt)$ est S-continue puisque, si l'accroissement du temps kdt est infiniment petit, l'accroissement de x, $|x((n + k)dt) - x(ndt)|$ l'est aussi.

Nous venons de reproduire le tout début de la démonstration I.S.T. du théorème d'existence des solution de l'équation différentielle :

$$(8) \qquad \text{EDO} \qquad \begin{cases} x(0) & = x_0 \\ \dfrac{dx}{dt} & = f(t) \end{cases}$$

Comme nous l'avions vu au § 4.5, les axiomes S. et T. vont nous permettre d'en déduire le théorème classique.

Mais nous pouvons voir les choses tout autrement.

> Au lieu de considérer (7) comme un auxiliaire pour démontrer l'existence de solution de (8) nous le considérons pour lui-même, nous en faisons un objet à part entière dont nous cherchons les propriétés.

Par exemple, nous venons de constater que si f est majorée par une constante limitée, la marche de pas infiniment petit associée est S-continue. On voit qu'il est absolument inutile de supposer que f est une fonction continue ; la marche de pas infiniment petit associée est toujours S-continue dès que f est majorée par une quantité limitée. Ceci permet de jeter un regard nouveau sur la vielle question des équations différentielles à second membre discontinu, comme il a été proposé dans [22].

5.4. *Vers un point de vue plus radical ?*

Nous venons de voir en quoi il n'est pas forcément souhaitable de rétablir dans la langue (interne) traditionnelle les énoncés obtenus dans la langue I.S.T. mais la tentation restera toujours grande de le faire... sauf si nous nous privons des moyens de le faire ! Dans I.S.T. oublions le S. et le T. et ne gardons que l'idéalisation qui nous permet le langage des infinitésimaux. Dans ce système axiomatique faible il ne sera plus possible d'atteindre tous les résultats traditionnels, mais en serons nous plus pauvres pour autant ? Personne n'a, pour le moment, la réponse définitive à cette question mais il nest pas interdit de se la poser et, éventuellement, de prendre parti.

Reeb et Nelson l'on fait, chacun à sa manière, le premier en nous rappelant, avec son ami J. Harthong que la doctrine intuitionniste de Brouwer avait encore toute son actualité [18], le second en nous proposant une théorie radicalement élémentaire [29] du mouvement Brownien. Mais en parler nous obligerait de parler des "vérités mathématiques" et de l'usage qu'on veut en faire, donc faire de la philosophie des mathématiques, ce qui nous ferait sortir du cadre de cet article.

6. — Pour finir, quelques éléments historico-bibliographiques

Il s'agit d'une d'une bibliographie sommaire très biaisée par mon point de vue de mathématicien français et, de plus, même si l'on se restreint au plan hexagonal elle reste incomplète.

On fait généralement démarrer l'ANS moderne au livre de 1966, *Non-standard Analysis* de A. Robinson [32] qui a été salué dès 1968 par le philosophe A. Badiou dans un article intitulé *La subversion infinitésimale* [2]. G. Reeb découvre ce livre en 1969 [21] ; séduit, il crée un groupe qui travaille sur diverses applications et, avec des collègues philosophes, entame une réflexion sur les fondements des mathématiques.

Il publie dans la collection IRMA de l'Université de Strasbourg un court essai *La mathématique non standard vieille de soixante ans ?* [30] ; avec A. Fuchs il rédige un cours de logique en 1981 non publié mais disponible au centre de documentation de l'IRMA ([13]). L'activité strasbourgeoise autour de Reeb se concrétise rapidement par une douzaine de thèses de mathématiques, un livre de cours en 1981, *Non standard Analysis : A practical guide with applications*, par R. Lutz et M. Goze [26] et un ouvrage collectif édité par H. Barreau et J. Harthong, qui aborde l'ANS sur les plans mathématique, historique et philosophique [3] ; sur ce dernier point l'article de Harthong et Reeb [18] jette un regard *intuitionniste* sur l'ANS. En 1977 E. Nelson publie *Internal Set Theory : a new approach to nonstandard analysis* [28] ; indépendamment et presque simultanément K. Hrbacek [19] publie une version axiomatique très proche ; la méthode axiomatique de Nelson est immédiatement reconnue et adoptée par divers mathématiciens dont l'école de Strasbourg. Je ne sais pas expliquer pourquoi la version de K. Hrbacek a eu moins de succès que celle de Nelson. L'article de revue [12] tente une description presque exhaustive des systèmes qui permettent de faire de l'analyse non standard en respectant les

21. Selon son élève R. Lutz c'est en 1973, à la bibliothèque de l'institut de mathématiques de Strasbourg que Reeb a découvert le livre de Robinson. Voir *La saga du bon maître Reeb*, L'Ouvert n 76, spécial G. Reeb. Disponible en ligne à l'adresse :
http://publimath.univ-irem.fr/biblio/IST94005.htm
Mais, selon Reeb lui même, dans sa préface à l'ouvrage collectif *La Mathématique Non Standard* (*Barreau-Harthong Eds.*), Editions du C.N.R.S. (1989) ; [13], ce serait en 1969.

canons usuels de la rigueur mathématique formelle.

Les années 1980 voient se développer de nombreuses applications : en physique mathématique [1, 17], mathématiques appliquées [11, 16], autour des méthodes asymptotiques [35] et les perturbations singulières d'équations différentielles [4, 7] [22]. Dans d'autres régions du monde et sur d'autres questions, de nombreuses applications de l'ANS se développent. Le premier ouvrage important à les recenser est [38] suivi de [1].

Références

[1] Sergio Albeverio, Jens Erik Fenstad, Raphael Høegh-Krohn, and Tom Lindstrøm. *Nonstandard methods in stochastic analysis and mathematical physics.* Academic Press, 1986.

[2] Alain Badiou. *La subversion infinitesimale. Cahiers pour l'analyse,* 9 : 118–137, 1968.

[3] Hervé Barreau and Jacques Harthong. *La mathématique non standard.* Editions du CNRS, Paris, 1989.

[4] Eric Benoît, Jean-Louis Callot, Francine Diener, and Marc Diener. *Chasse au canard.* Collectanea Mathematica, 32(2) :37–119, 1981.

[5] Błaszczyk, Piotr and Borovik, Alexandre and Kanovei, Vladimir and Katz, Mikhail G and Kudryk, Taras and Kutateladze, Semen S and Sherry, David *A non-standard analysis of a cultural icon : The case of Paul Halmos,* Logica Universalis, 10,4 : 393–405, 2016.

[6] JL Callot. Trois leçons d'analyse infinitésimale. In *Le Labyrinthe du Continu (Colloque de Cerisy),* Jean-Michel Salanskis, Hourya Sinaceur (eds), Springer-Verlag 1992.

[7] Pierre Cartier. *Perturbations singulières des équations différentielles ordinaires et analyse non-standard.* Séminaire Bourbaki, 24 :21–44, 1981.

[8] André Deledicq. *Est-il possible d'enseigner l'analyse aujourd'hui ?* Repères IREM, 24 : 79-102, 1996.

[9] André Deledicq and Marc Diener. *Leçons de calcul infinitésimal.* Armand Colin, 1989.

[10] Francine Diener and Georges Reeb. *Analyse non standard.* Hermann, Edition conjointe CNRS (Paris) et OPU (Alger), 1989.

[11] M Diener and C Lobry. *Analyse non standard et représentation du réel.* 1985.

[12] Peter Fletcher, Karel Hrbacek, Vladimir Kanovei, Mikhail G Katz, Claude Lobry, and Sam Sanders. *Approaches to analysis with infinitesimals following Robinson, Nelson, and others.* Real Analysis Exchange, 42(2) :193–252, 2017.

22. dont la découverte des "canard" qu'on peut retrouver sur Scholarpedia : http://www.scholarpedia.org/article/Canards et sur le site du CNRS : http://images.math.cnrs.fr/Des-canards-dans-mes-neurones.html

[13] Aimé Fuchs and Georges Reeb. *Logique*. IRMA, Université de Strasbourg,, 1981.

[14] Thérèse Gilbert. *L'enseignement de la continuité et de la dérivabilité en analyse nonstandard*. Repères IREM, (13), 1993.

[15] Thérèse Gilbert. *Qu'est-ce que l'analyse non-standard*. Repères IREM, (11), 1993.

[16] Jacques Harthong. *Le moiré*. Advances in Applied Mathematics, 2(1) :24–75, 1981.

[17] Jacques Harthong. *Etudes sur la mécanique quantique*. Astérisque n111, Société mathématique de France, 1984.

[18] Jacques Harthong and Georges Reeb. *Intuitionnisme 84*. In *La mathématique nonstandard*, CNRS, pages 307–329, 1989.

[19] Karel Hrbacek. *Axiomatic foundations for nonstandard analysis*. Fundamenta Mathematicae, 98 :1–19, 1978.

[20] Jean-Louis Krivine. *Théorie axiomatique des ensembles*. 1969.

[21] Claude Lobry. *La méthode des élucidations successives*. Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées, 9 :171–193, 2008.

[22] Claude Lobry and Tewfik Sari. *Equations différentiellesa second membre discontinu*. Contrôle non linéaire et Applications, Travaux en Cours, (64) :255–289, 2004.

[23] Claude Lobry. *L'analyse non standard en France 1975-1995 : une dispute avortée*. Doctoral dissertation, Université Côte d'Azur
https://www.theses.fr/2019AZUR2017

[24] Robert Lutz. *Rêveries infinitésimales*. Gazette des mathématiciens, (34) :79–87, 1987.

[25] Robert Lutz. *La force modélisatrice des théories infinitésimales faibles. Le Labyrinthe du Continu*, Jean-Michel Salanskis, Hourya Sinaceur (eds), Springer-Verlag 1992.

[26] Robert Lutz and Michel Goze. *Nonstandard Analysis. : A Practical Guide with Applications.*, volume 881. Springer, 1981.

[27] Robert Lutz, Abdelnacer Makhlouf, and Etienne Meyer. *L'enseignement de l'analyse en termes d'ordre de grandeur : nécessité et fondements*. REPERES-IREM, (24) :103–116, 1996. wblock Princeton University Press, 1967.

[28] Edward Nelson. *Internal set theory : a new approach to nonstandard analysis*. Bulletin of the American Mathematical Society, 83(6) :1165–1198, 1977.

[29] Edward Nelson. *Radically Elementary Probability Theory.* (AM-117), volume 117. Princeton University Press, 1987.

[30] Georges Reeb. *La mathématique non standard vieille de soixante ans ?* IRMA, Université Louis Pasteur de Strasbourg, P30, 1979.

[31] Alain Robert. *Analyse non standard.* PPUR presses polytechniques, 1985.

[32] Abraham Robinson. *Non-standard analysis.* Princeton University Press (réédition), 2016.

[33] Tzuchien, THO *Equivocation in the foundations of Leibniz's infinitesimal fictions*, Society and Politics, 62 : 70–98, 2012

[34] WP Thurston. *On proof and progress in mathematics.* Bull. Am. Math. Soc., 30 (math. HO/9404236) :161–177, 1994.

[35] Imme Van den Berg. *Nonstandard asymptotic analysis,* volume 1249. Springer, 2006 (1987).

[36] Imme van den Berg and Bruno Dinis. *Neutrices and External Numbers A Flexible Number System.* Springer, 2019.

[37] GuyWallet and Stefan Neuwirth. *Enquête sur les modes d'existence des êtres mathématiques (version augmentée).* hal.archives-ouvertes.fr/hal-01824928v2/document.

[38] Zvonkin, A. K., and Shubin, M. A. *Non-standard analysis and singular perturbations of ordinary differential equations.* Russian Mathematical Surveys, 39(2), 69., 1984

Centre de Recherche en Histoire des Idées, Université de Nice Sophia-Antipolis, 98, Bd Edouard Herriot, BP 3209- 0624 Nice-cedex