Recherche Coopérative sur Programme n° 25

Pierre Cartier

Analyse non standard : nouvelles méthodes infinitésimales en analyse. Application à la géométrie et aux probabilités

Les rencontres physiciens-mathématiciens de Strasbourg - RCP25, 1985, tome 35 « Conférences de : P. Cartier, A. Guichardet et G.A. Viano », , exp. n° 1, p. 1-21

<http://www.numdam.org/item?id=RCP25_1985__35__1_0>

ANALYSE NON STANDARD : NOUVELLES MÉTHODES INFINITÉSIMALES EN ANALYSE . APPLICATION À LA GÉOMÉTRIE ET AUX PROBABILITÉS

par Pierre CARTIER (Ecole Polytechnique)

1. Historique des infinitésimaux

De Cavalieri , au 17^e siècle , à Cauchy , au 19^e siècle , les mathématiciens firent un grand usage des infiniment petits et des infiniment grands . On définit une intégrale comme une somme d'un nombre infiniment grand de quantités infiniment petites , une tangente à une courbe est la droite qui joint deux points infiniment voisins de la dite courbe , une fonction est continue si un accroissement infiniment petit de la variable produit un accroissement infiniment petit de la fonction . Les propriétés qui nous sont bien familières , concernant les fonctions ou les figures géométriques , s'établissent en général de manière simple et relativement intuitive . Malheureusement , on ne peut éviter certains paradoxes ; comme le remarque avec vivacité l'évêque Berkeley , la principale difficulté est qu'une quantité infiniment petite (dite aussi infinitésimale) sera considérée selon les cas comme nulle , ou comme non nulle . Prenons par exemple le calcul de la dérivée de la fonction x^2 de la variable x (je devrais dire de la fonction $x \longmapsto x^2$); on a donc $y = x^2$, et si l'on donne un accroissement infiniment petit dx à x , on trouve $y + dy = (x + dx)^2 = x^2 + 2x.dx + (dx)^2$, d'où $dy = (2x + dx).dx$ et finalement

$$\frac{dy}{dx} = 2x + dx = 2x \quad .$$

Le résultat $dy/dx = 2x$ est bien conforme à l'orthodoxie , mais le paradoxe est que l'on doit supposer dx non nul pour faire la division , alors qu'on doit le supposer nul pour écrire l'égalité finale $2x + dx = 2x$. La difficulté était double :

--- quelle est la nature des infinitésimaux ?

--- comment codifier avec précision les règles de calcul et de raisonnement sur ces objets ?

A la première question , Leibniz répond qu'il s'agit de nombres idéaux , que l'on adjoint aux nombres "réels" . On peut donc concevoir qu'après les extensions multiples de la notion de nombre --- qui ont conduit des entiers positifs aux nombres complexes en passant par les nombres fractionnaires , les nombres négatifs , les nombres irrationnels , puis imaginaires --- il faille une étape supplémentaire . Ces nombres seront manipulés se-

lon les règles usuelles , et toutes les identités valables pour les nombres usuels sont encore valables pour ces nombres généralisés . Selon la terminologie algébrique contemporaine , les nombres réels généralisés forment un corps commutatif totalement ordonné . Mais les infinitésimaux ne sont que des intermédiaires de démonstration ; s'ils sont légitimes dans les étapes intermédiaires , ils doivent disparaître du résultat final .

La réponse la plus complète à la seconde question est celle de Lazare Carnot : dans un ouvrage qui fut fameux en son temps "Réflexions sur la métaphysique du calcul infinitésimal" --- et dont la seconde édition parut en 1797 , en pleine crise révolutionnaire , alors que Lazare Carnot occupe des fonctions politiques éminentes --- il introduit la distinction capitale entre les équations imparfaites , où les deux membres diffèrent d'un infiniment petit , et les équations parfaites . Autrement dit , à côté de l'égalité absolue des nombres , notée hier et aujourd'hui $a = b$, il faut considérer une égalité approchée , notée aujourd'hui $a \backsimeq b$ et qui signifie que $a - b$ est infiniment petit . Mais il note encore $a = b$ cette nouvelle relation , et il se peut que ce manque de notation autonome ait empêché la diffusion et l'acceptation de son système .

Quoi qu'il en soit , on ne put jamais , au 18^e et au 19^e siècle , élaborer un système qui fût conérent , complet et commode . Déjà , d'Alembert avait affirmé que le concept fondamental était celui de limite , et que toutes les autres notions du Calcul Infinitésimal peuvent s'y ramener . Cette notion de limite n'est pas non plus aisée à définir , et il fallut attendre Weierstrass pour la clarifier complètement . Aussi , malgré l'effort de Lazare Carnot pour élucider la nature des raisonnements sur les infinitésimaux , malgré la répugnance de Cauchy lui-même à abandonner les infiniment petits , le point de vue d'Abel , Dirichlet , Weierstrass et Kronecker finit par s'imposer . Toutes les notions sont subordonnées à celle de limite , et cette dernière est remplacée --- définie --- par une manipulation assez sophistiquée d'inégalités : les ε et les δ qui sont aujourd'hui bien familiers dans les cours de Mathématiques . Sur le plan de la rigueur , il n'y a rien à reprocher à ces méthodes "nouvelles" (en 1850) ; elles se sont bien incorporées aux idées de Cantor , et elles se trouvent au coeur de la Topologie Générale ou de l'Analyse Fonctionnelle qui forment la version moderne du Calcul différentiel et Intégral dans une Mathématique entièrement remodelée dans le cadre de la Théorie des Ensembles .

2. Mathématique non-archimédienne

A la fin du 19^e siècle , Veronese dans une suite de travaux d'accès difficile tente de dépasser les mathématiques archimédiennes . On les connaît surtout par la ver-

sion simplifiée qu'en donne Hilbert dans ses "Grundlagen der Geometrie" . Dans ce dernier ouvrage , Hilbert donne un système axiomatique aussi complet que possible pour la géométrie euclidienne , indépendamment de l'idée cartésienne des coordonnées , dans la ligne de la tradition d'Euclide . Il examine soigneusement l'indépendance des divers axiomes , et les géométries plus ou moins inhabituelles que l'on obtient en supprimant ou en affaiblissant tel ou tel axiome . Dans l'étude de l'Axiome d'Archimède , il introduit des modèles de la géométrie , où les coordonnées d'un point de l'espace seraient prises dans un système non archimédien . Il s'agit d'élargir l'ensemble des nombres réels $\underset{\sim}{R}$ en un corps totalement ordonné (commutatif) K . Si K est distinct de $\underset{\sim}{R}$, l'axiome d'Archimède est violé et il existe des éléments x de K tels que l'on ait $x > n$ pour tout entier n ; de tels éléments seront appelés "infiniment grands" (positifs) et l'on peut , par changement de signe , introduire des nombres négatifs infiniment grands . Un nombre ε sera infiniment petit si l'on a $\varepsilon = 0$ ou si ε est l'inverse d'un nombre infiniment grand . On peut aussi dire que l'élément ε de K est infiniment petit si l'on a $|\varepsilon| < x$ pour tout nombre réel $x > 0$; il s'agit là d'une version non contradictoire de la fameuse définition d'Abel :

"On appelle infiniment petite toute quantité qui est plus petite que toute quantité assignable" .

Bien entendu , on n'échappe à la contradiction qu'en distinguant soigneusement deux domaines emboîtés $\underset{\sim}{R}$ et K , et en interprétant "quantité assignable" comme les éléments du plus petit domaine $\underset{\sim}{R}$.

Dans ce cadre , on justifie facilement une partie des règles de calcul de Lazare Carnot . Ecrivons $a \simeq b$ pour signifier que $a - b$ est infiniment petit . Si un élément a de K n'est pas infiniment grand (positif ou négatif) , il existe un unique élément a^o de $\underset{\sim}{R}$ tel que $a \simeq a^o$; autrement dit , a est la somme $a^o + \varepsilon$ d'un nombre réel a^o et d'un nombre infiniment petit ε . Si a et b sont deux nombres non infiniment grands , l'égalité approchée $a \simeq b$ équivaut à l'égalité vraie $a^o = b^o$. Le calcul donné plus haut et qui de $y = x^2$ déduit la dérivée $dy/dx = 2x$ s'interprète comme suit : soit x un nombre réel et soit $y = x^2$; choisissons un nombre x' (dans K) infiniment voisin de x , et posons $y' = x'^2$; les accroissements

$$dx = x' - x \qquad , \qquad dy = y' - y$$

sont infiniment petits , et l'on a

$$\frac{dy}{dx} = 2x + dx \simeq 2x \quad , \quad \text{soit encore} \quad \left(\frac{dy}{dx}\right)^o = 2x \quad .$$

On doit donc interpréter la dérivée comme la partie finie du quotient de l'accroisse-

ment de la fonction par l'accroissement de la variable .

Pour éviter tout malentendu , précisons que rien de tout ce qui précède ne se trouve chez Hilbert ; il se contente de donner un exemple explicite d'extension non archimédienne du corps des nombres réels . Ce n'est que vers 1930 que les algébristes et les logiciens --- Artin , Schreier , Nikodym , Tarski --- feront l'étude systématique des corps ordonnés . La conception des dérivées que nous avons décrite plus haut est essentiellement celle qui se trouve dans le manuel d'Analyse de J. Keisler .

3. Théorie des modèles

Les idées voisines de dérivée , de vitesse ou de tangente à une courbe , sont assez faciles à formuler de manière algébrique --- Descartes lui-même donne une telle définition des tangentes à une courbe algébrique , qui est encore acceptée aujourd'hui . La conception des indivisibles de Cavalieri --- où une aire est conçue comme la somme d'un nombre infiniment grand de rectangles infiniment fins --- oblige à considérer et manipuler des entiers infiniment grands . La situation de l'axiome d'Archimède est particulièrement subtile : on ne nie pas que tout nombre réel (positif) puisse être majoré par un nombre entier , mais l'on admet deux sortes de nombres entiers , les finis et les infinis . Leibniz postule que ces nombres obéissent aux mêmes règles de calcul que les entiers finis ordinaires , et Euler étend à ces entiers les règles de calcul sur les sommes , et les polynômes . Il y aura ainsi des polynômes de degré infini , tels que le polynôme

$$(1 + x/\Omega)^{\Omega} = 1 + x + \frac{1}{2}(1 - 1/\Omega)x + \ldots + x^{\Omega}/\Omega! \quad ,$$

introduit par Euler pour définir l'exponentielle . Mais l'utilisation sans frein de telles sommes à un nombre infiniment grand de termes conduit à des paradoxes ; en voici deux d'ailleurs apparentés . Considérons la somme

$$S_{\Omega} = 2^{-1} + 2^{-2} + \ldots\ldots + 2^{-(\Omega-1)} + 2^{-\Omega} = \sum_{n=1}^{\Omega} 2^{-n} \quad ;$$

si l'on admet les règles usuelles de manipulation des sommes , on est conduit à lui attribuer la somme $1 - 2^{-\Omega-1} < 1$. Quelle valeur faut-il attribuer à la série infinie

$$S = 2^{-1} + 2^{-2} + 2^{-3} + \ldots\ldots + 2^{-N} + \ldots\ldots \quad ?$$

On a évidemment $S < S_{\Omega}$, donc $S < 1$, et pourtant , on a $S > S_N = 1 - 2^{-N-1}$ pour tout entier fini N . La valeur $S = 1$, qui est couramment admise, est la seule possible , et pourtant elle contredit l'inégalité $S < S_{\Omega} < 1$. De manière analogue , la somme

$$\Sigma = 1 + 1 + 1 + \ldots\ldots + 1 + \ldots\ldots$$

satisfait à l'inégalité $N < \Sigma < \Omega$ pour tout entier fini N et tout entier infini Ω ;

ceci conduit à la conséquence absurde $\Sigma = \Sigma + 1$. De tels calculs ont laissé une trace dans les conceptions courantes , où l'on tolère un nombre infini unique , noté ∞ ; il obéit à des règles de calcul telles que

$$\infty + 1 = \infty \quad , \quad \infty \times \infty = \infty \quad , \quad 2^{-\infty} = 0 \quad , \quad 1/\infty = 0 \quad ,$$

qui l'empêchent d'être un nombre comme les autres . A l'inverse , pour Leibniz , Euler et leurs émules , il y a plusieurs nombres infinis distincts , mais ils obéissent aux mêmes règles algébriques que les autres . En géométrie aussi , on n'hésitera pas à remplacer un cercle par un polygône régulier à une infinité de côtés .

Ce n'est qu'au 20^{e} siècle que les logiciens sont parvenus à résoudre de tels paradoxes par l'introduction de l'axiomatique , de la logique du premier ordre et de la théorie des modèles . Pour prendre un exemple , considérons le corps $\underset{\sim}{Q}$ des nombres rationnels , le corps $\underset{\sim}{R}$ des nombres réels et le corps intermédiaire $\underset{\sim}{A}$ des nombres algébriques réels . Il s'agit là de trois exemples de corps ordonnés ; autrement dit , dans chacun de ces systèmes , on sait définir la somme $a + b$, la différence $a - b$ et le produit $a \times b$ (ou ab) de deux éléments , on dispose de deux nombres particuliers 0 et 1 , on définit le quotient a/b lorsque b est différent de 0 , et l'on peut comparer deux nombres a et b par la relation notée $a < b$. On formule sous forme d'axiomes les règles de calculs usuelles , y compris celles qui permettent de manipuler les inégalités . Les notions et propriétés du premier ordre sont celles qui s'expriment en fonction des opérations et relations fondamentales ; par exemple , les notions de carré x^2 , de cube x^3 , de polynôme du second degré $ax^2 + bx + c$, de racine d'un tel polynôme ... sont du premier ordre ; une propriété telle que

« si a et b sont deux nombres tels que $a < b$, on a $a^2 < b^2$ »

est du premier ordre .

Dans les trois exemples $\underset{\sim}{Q}$, $\underset{\sim}{A}$ et $\underset{\sim}{R}$, les notions de base sont définies , et les axiomes sont satisfaits ; on dit que ce sont des modèles de la théorie axiomatique des corps ordonnés . Parmi les propriétés du premier ordre , certaines s'avèrent être vraies dans tous les modèles (par exemple , la propriété $a^2 \geq 0$) , certaines sont fausses dans tous les modèles (par exemple , celle qui est donnée ci-dessus) , et certaines sont tantôt vraies , tantôt fausses . Un théorème fondamental de Herbrand affirme que les propriétés du premier ordre qui sont vraies dans tous les modèles ne sont autres que celles que l'on peut démontrer à partir des axiomes par des raisonnements purement logiques , en faisant abstraction des propriétés particulières et de la nature des éléments du modèle (« théorème de complétude ») . La propriété du premier ordre

« il existe un nombre a tel que $a^2 = 2$ »

est vraie dans $\underset{\sim}{A}$ et dans $\underset{\sim}{R}$, elle est fausse dans $\underset{\sim}{Q}$. On peut donc distinguer les modèles $\underset{\sim}{Q}$ et $\underset{\sim}{R}$ par une propriété du premier ordre, c'est-à-dire par des moyens internes à la théorie des corps ordonnés. Une découverte surprenante de Tarski (vers 1930) est que l'on ne peut pas distinguer $\underset{\sim}{A}$ et $\underset{\sim}{R}$ par une propriété du premier ordre. Pourtant, il existe une propriété simple qui distingue $\underset{\sim}{A}$ et $\underset{\sim}{R}$, à savoir

« pour tout nombre x, il existe des entiers n et $a_0, a_1, \ldots, a_n$ avec $a_0 \neq 0$ et tels que l'on ait $a_0 x^n + a_1 x^{n-1} + \ldots + a_n = 0$ »

qui est satisfaite dans $\underset{\sim}{A}$ et violée dans $\underset{\sim}{R}$. Cette propriété n'est pas du premier ordre à cause du point subtil suivant : dans une propriété du premier ordre, on peut mentionner un entier explicite tel que $2 = 1 + 1$, $4 = 2 \times 2$, $81 = (3 \times 3) \times (3 \times 3)$ (ceux que Reeb appelle les entiers "naïfs"), mais non un entier arbitraire. Autrement dit, on ne peut pas distinguer $\underset{\sim}{A}$ et $\underset{\sim}{R}$ par des moyens internes à la théorie des corps ordonnés, mais seulement à l'intérieur d'une théorie plus riche.

Une situation analogue se présente pour les nombres entiers positifs. Comme il est bien connu, ceux-ci sont décrits par l'axiomatique de Peano ; les notions fondamentales sont celles du nombre 0 et de successeur x' d'un nombre x (égal à $x + 1$) et les axiomes sont les suivants :

a) Pour tout nombre x, on a $x' \neq 0$.

b) Pour tout nombre y tel que $y \neq 0$, il existe un nombre unique x tel que $y = x'$.

c) Si une propriété $\underset{\sim}{P}(x)$ est telle que les relations $\underset{\sim}{P}(0)$ et $\underset{\sim}{P}(x) \Longrightarrow \underset{\sim}{P}(x')$ soient vraies, alors $\underset{\sim}{P}(x)$ est vraie pour tout entier x.

Le point qui demande élucidation est le sens de "propriété" dans l'axiome de récurrence c). A l'intérieur de la théorie des ensembles, on l'interprétera comme suit : si l'on note $\underset{\sim}{N}$ l'ensemble des entiers positifs, on postule que si A est un sous-ensemble de $\underset{\sim}{N}$ contenant 0, et si la relation $x \in A$ entraîne $x' \in A$ pour tout entier x, on a $A = \underset{\sim}{N}$.

Dans ces conditions, à l'intérieur de la théorie des ensembles, deux modèles des axiomes de Peano (sous la forme ensembliste) sont isomorphes. Les logiciens ont décrit avec précision une notion de propriété du premier ordre des entiers assez riche pour contenir la quasi-totalité des propriétés intéressantes. La découverte surprenante de Skolem (aussi vers 1930) est l'existence de deux modèles non isomorphes de l'arithmétique de Peano qui ne peuvent être distingués par aucune propriété du premier ordre.

Autrement dit, à côté du modèle usuel $\underset{\sim}{N}$ des entiers, on dispose d'un modèle non standard $\overset{*}{\underset{\sim}{N}}$ et toutes les propriétés du premier ordre qui sont valables pour les

entiers usuels le sont pour les entiers non standard de $\overset{*}{\mathbb{N}}$. On peut donner une construction simple de $\overset{*}{\mathbb{N}}$ au moyen des ultraproduits introduits par Łoś vers 1950 ; j'en ai donné par ailleurs une interprétation probabiliste .

4. L'analyse non standard de Robinson

Robinson , mort prématurément , était l'un des meilleurs logiciens de notre siècle et l'un des experts mondiaux de l'aérodynamique et de l'aviation . Parfaitement conscient de l'efficacité des méthodes infinitésimales en Mécanique --- où elles subsistent aujourd'hui malgré les tabous de l'orthodoxie mathématique --- il a appliqué son savoir-faire de logicien à les réhabiliter .

La notion-clé de Robinson est celle de superstructure construite sur un ensemble --- notion voisine de la notion d'échelle d'ensembles introduite par Bourbaki dans sa description générale des structures . Partant d'un ensemble X , on définit comme suit[1] la suite des ensembles $V_n(X)$:

$$V_0(X) = X \quad , \quad V_{n+1}(X) = V_n(X) \cup P(V_n(X)) \qquad ;$$

puis l'on note $V(X)$ la réunion de ces ensembles $V_0(X), V_1(X), \ldots$ Interprétons un couple ordonné (a,b) d'éléments comme l'ensemble $\{\{a\},\{a,b\}\}$; si A et B sont deux ensembles, on aura alors $A \times B \in P(P(P(A \cup B)))$; il en résulte que la plupart des objets construits en Analyse se représentent par des éléments de $V_2(\underset{\sim}{R})$ ou $V_3(\underset{\sim}{R})$. La superstructure $V(\underset{\sim}{R})$ est donc apte à décrire l'Analyse classique et moderne .

Robinson imagine donc qu'à côté de la superstructure $V(\underset{\sim}{R})$, on introduise une autre superstructure $V(\overset{*}{\underset{\sim}{R}})$ et une application qui à tout élément a de $V(\underset{\sim}{R})$ fasse correspondre un élément $^*a$ de $V(\overset{*}{\underset{\sim}{R}})$. On postule un <u>principe de transfert</u> : une propriété du premier ordre a lieu entre des éléments $a,b,c,\ldots$ de $V(\underset{\sim}{R})$ si et seulement si elle est satisfaite pour les correspondants $^*a, ^*b, ^*c, \ldots$ Il résulte de ce principe que , $\underset{\sim}{R}$ étant un corps ordonné , il en est de même de $\overset{*}{\underset{\sim}{R}}$: si l'on note par exemple S l'ensemble des triplets de nombres réels (a,b,c) tels que $a = b + c$, ou I l'ensemble des couples (a,b) tels que $a < b$, l'ensemble $^*S$ se compose des triplets (a,b,c) d'éléments de $\overset{*}{\underset{\sim}{R}}$ tels que $a = b + c$, et $^*I$ se décrit de manière analogue dans $\overset{*}{\underset{\sim}{R}}$.

Un objet appartenant à $V(\overset{*}{\underset{\sim}{R}})$ est dit standard s'il est de la forme $^*a$, où a est un objet appartenant à $V(\underset{\sim}{R})$. On pourra parler par exemple des entiers standard , des nombres réels standard . Il résulte du principe de transfert que l'application $a \longmapsto {}^*a$ est injective : la propriété du premier ordre $a = b$ est satisfaite si et seulement si l'on a $^*a = {}^*b$. Il n'y a aucun inconvénient à identifier un nombre a

de $\underset{\sim}{R}$ à son correspondant ${}^*a$ dans ${}^*\underset{\sim}{R}$. De la sorte $\underset{\sim}{R}$ est un sous-corps du corps ordonné ${}^*\underset{\sim}{R}$. Comme on l'a mentionné au n°2 , cela suffit pour permettre l'introduction des nombres infiniment petits ou infiniment grands dans ${}^*\underset{\sim}{R}$, de l'égalité approchée $a \underset{\sim}{\sim} b$ entre éléments de ${}^*\underset{\sim}{R}$ et de la partie finie (ou standard) a° d'un nombre non standard a qui n'est pas infiniment grand (on dira "limité") . On dira qu'une application g de ${}^*\underset{\sim}{R}$ dans ${}^*\underset{\sim}{R}$ est standard si elle est de la forme $g = {}^*f$, où f est une application de $\underset{\sim}{R}$ dans $\underset{\sim}{R}$. En pratique , f et g sont désignés par le même symbole ; si l'on a identifié $\underset{\sim}{R}$ à un sous-ensemble de ${}^*\underset{\sim}{R}$, les fonctions sinus , exponentielle ,... s'étendent en des fonctions de même nom dans ${}^*\underset{\sim}{R}$. En pratique , tous les objets mathématiques courants dont la définition est suffisamment explicite --- qu'il s'agisse de nombres comme e , π , $\sqrt{2}$,... ou de fonctions comme $\sin x$, $\exp x$,... ou de l'espace ℓ^2 des suites $x = (x_1, x_2, \ldots)$ telles que $\sum_{n=1}^{\infty} x_n^2$ soit fini --- ont un correspondant dans $V({}^*\underset{\sim}{R})$ que l'on affuble du qualificatif "standard" .

Si un ensemble A appartenant à la superstructure $V({}^*\underset{\sim}{R})$ est standard , donc de la forme ${}^*B$, où B appartient à $V(\underset{\sim}{R})$, il possédera beaucoup d'éléments standard , à savoir les éléments ${}^*b$, où b parcourt B . Il en possède suffisamment , puisque l'on montre que deux ensembles standard qui ont les mêmes éléments standard sont égaux . Mais un ensemble standard infini contient toujours des éléments non standard --- par exemple , il existe des entiers non standard et des nombres réels non standard . On introduit donc une distinction supplémentaire : un objet appartenant à $V({}^*\underset{\sim}{R})$ est dit interne si c'est un élément d'un ensemble standard . En particulier , les objets standard sont internes , mais la réciproque est fausse ; tout nombre hyperréel (on baptise ainsi les éléments de ${}^*\underset{\sim}{R}$) est interne , et il en est de même des éléments hyperentiers (de ${}^*\underset{\sim}{N}$) . Les ensembles de nombres hyperréels se classent de la manière suivante :

--- ensembles de la forme ${}^*A$, avec $A \subset \underset{\sim}{R}$, qualifiés de <u>standard</u> (par exemple , l'ensemble ${}^*\underset{\sim}{N}$ des nombres hyperentiers , l'intervalle $[0,+\infty[$ de ${}^*\underset{\sim}{R}$) ;

--- ensembles appartenant à ${}^*P(\underset{\sim}{R})$, appelés <u>internes</u> (par exemple , un intervalle de la forme $[a,b]$, où a et b sont deux nombres hyperréels , standard ou non) ;

--- ensembles non internes , et donc baptisés <u>externes</u> (par exemple , l'ensemble des nombres hyperréels infiniment petits) .

De manière parallèle , on distinguera parmi les fonctions :

--- les fonctions standard (exponentielle , sinus ,...) ;

--- les fonctions internes (obtenues telles $t \longmapsto \sin \omega t$ par substitution d'une valeur non standard ω d'un paramètre dans la fonction standard de deux variables $\sin kt$);

--- les fonctions externes (par exemple , celle qui vaut 0 ou x selon que le nom-

bre x est standard ou non) .

Les notions topologiques prennent un tour nouveau . Considérons par exemple un espace métrique (X,d) de la superstructure $V(\underset{\sim}{R})$. Alors X se plonge comme l'ensemble des éléments standard dans l'extension ${}^*X$; la distance sur X , qui est une application d de $X \times X$ dans $\underset{\sim}{R}$, se prolonge en une distance ${}^*d$ sur ${}^*X$. La monade (ou halo) d'un point a de ${}^*X$ sera alors l'ensemble des points x de ${}^*X$ tels que la distance ${}^*d(x,a)$ soit infiniment petite . Un point de ${}^*X$ est dit presque-standard (ou accessible) s'il appartient au halo d'un point standard . Alors l'espace X est compact si et seulement si tout point de ${}^*X$ est accessible , une partie U de X est ouverte si et seulement si ${}^*U$ contient le halo de chacun de ses points standard , l'adhérence $\overline{A}$ d'une partie A de X se compose des points dont le halo rencontre ${}^*A$.

Terminons ce numéro par l'étude de la continuité . Soit f une application de $\underset{\sim}{R}$ dans $\underset{\sim}{R}$, que nous prolongeons en une application standard de ${}^*\underset{\sim}{R}$ dans lui-même , encore notée f . La continuité de f peut se décrire par l'une ou l'autre des propriétés suivantes :

a) (Continuité au sens de Weierstrass) : Etant donnés un point x_0 de ${}^*\underset{\sim}{R}$ et un nombre $\varepsilon > 0$, il existe un nombre $\varepsilon' > 0$ tel que la relation $|x - x_0| < \varepsilon'$ entraîne $|f(x) - f(x_0)| < \varepsilon$.

b) (Continuité infinitésimale) : pour tout point x_0 de $\underset{\sim}{R}$, et tout point x de ${}^*\underset{\sim}{R}$ infiniment proche de x_0 (on a $x \simeq x_0$) , alors $f(x)$ est infiniment proche de $f(x_0)$.

Pour une fonction standard f ces propriétés sont équivalentes , mais il n'en est plus de même pour une fonction interne . Si C est l'ensemble des applications continues de $\underset{\sim}{R}$ dans $\underset{\sim}{R}$ (au sens usuel) , la continuité au sens de Weierstrass caractérise les éléments de l'ensemble standard ${}^*C$. La continuité infinitésimale est une notion nouvelle , souvent baptisée S-continuité . On donne facilement des exemples de fonctions internes qui sont continues en l'un de ces sens , mais pas en l'autre .

On devra de manière analogue distinguer différentiabilité et S-différentiabilité, intégrabilité et S-intégrabilité ...

5. Les successeurs de Robinson

L'héritage de Robinson a été exploité par une école américaine active et nombreuse[2] ; elle est composée de logiciens (Keisler , Helson , et Kreisel d'une manière plus marginale) , de spécialistes de théorie des ensembles et d'analyse fonctionnelle (Luxemburg , Stroyan , Bayod) , par des analystes et des probabilistes (Anderson , Loeb , Wat-

temberg) . Une école française autour de Reeb et Lutz a surtout développé les applications à la géométrie et aux équations différentielles ; nous y reviendrons plus loin . En Allemagne , Laugwitz et ses élèves s'intéressent aux fondements de la théorie ; en Allemagne et en Scandinavie , Hoegh-Krohn , Fenstad et Lindstrom et quelques autres s'intéressent aux applications des méthodes infinitésimales à la Physique Mathématique et aux équations aux dérivées partielles .

Les méthodes de Robinson ne sont pas faciles . Les 100 pages de logique qui ouvrent son livre ont sans doute découragé plus d'un lecteur ; la nécessité de mettre des étoiles partout et de distinguer X et ${}^*X$, la démultiplication des notions (continuité et S-continuité ...) apportent beaucoup de complications , alors qu'on pouvait espérer que l'utilisation des infinitésimaux simplifierait l'Analyse et la rendrait intuitive .

Dans un article fondamental paru en 1977 , E.Nelson a tenté d'axiomatiser l'Analyse non standard , sous le nom de "Théorie des ensembles internes" (avec le sigle IST). L'idée de base est que les ensembles internes de Robinson sont l'objet central ; ce seront "les" ensembles et l'on admettra à leur sujet les règles usuelles , telles qu'elles sont codifiées dans la théorie de Zermelo-Fraenkel (au sigle ZFC) . Dans l'univers des ensembles , on distingue une sous-classe formée des ensembles "standard" . Les nouvelles règles de raisonnement sont au nombre de trois :

--- l'axiome de transfert (T) exprime que si l'on veut prouver une propriété interne (c'est-à-dire dans l'énoncé de laquelle la notion d'ensemble standard ou une notion dérivée telle que celle de nombre infini petit n'interviennent pas implicitement ou explicitement) , il suffit de l'établir sous l'hypothèse supplémentaire que tous les objets considérés sont standard ;

--- l'axiome de standardisation (S) affirme que , étant donnés un ensemble (interne) A et une propriété $\Pi(x)$ de ses éléments , il existe un unique ensemble standard X dont les éléments standard sont les points standard x de A qui satisfont à la propriété $\Pi(x)$;

--- l'axiome d'idéalisation (I) est un principe général de compacité dont une conséquence d'allure paradoxale est l'existence d'un ensemble fini auquel appartienne tous les objets standard .

Le transfert est utilisé fréquemment dans la théorie de Robinson , et l'axiome de standardisation est un substitut de l'opération * .

Nelson n'introduit pas explicitement les ensembles externes , et c'est parfois une limitation gênante . En 1978 , le logicien Hrbacek a décrit une théorie des ensem-

bles à trois niveaux , où l'on considère des ensembles standard , internes et externes . Cette théorie contient l'équivalent des axiomes (I) (S) et (T) et des règles de manipulation des ensembles externes . Hrbacek étudie d'ailleurs plusieurs variantes de sa théorie , leur force respective et leur non -contradiction (relativement à la théorie ZFC prise pour base) .

Aucune de ces théories concurrentes n'est vraiment simple , et je ne suis pas sûr qu'aucune codifie complètement toute la pratique de la mathématique non-standard . En 1984 , à l'occasion de recherches sur le Calcul des Probabilités , Nelson a codifié un petit bout ("a tiny bit" dans ses termes propres) de l'Analyse Non Standard , au moyen d'une mini-axiomatique . C'est une version développée de ce nouveau système que je décris au numéro suivant .

6. Une nouvelle vue sur le continu

L'idée de base est de codifier l'emploi des mots "grand" et "petit" tels qu'ils sont utilisés dans une manière informelle de s'exprimer en Mathématiques . Nous partons de l'ensemble $\underset{\sim}{N}$ des entiers positifs 0,1,2,... Rappelons qu'une partie A de $\underset{\sim}{N}$ est finie si et seulement si elle est contenue dans un intervalle de la forme $[0,n]$ (qui se compose des entiers i tels que l'on ait $0 \leq i \leq n$) et qu'elle contient alors un plus petit et un plus grand élément —— c'est une des formes du principe de récurrence , ou la descente infinie de Fermat .

En Statistique se pose souvent de répartir des individus en deux classes , les petits et les grands . On veut donc partager $\underset{\sim}{N}$ en deux ensembles non vides A et B tels que l'on ait $a < b$ quels que soient a dans A et b dans B . D'après le principe de récurrence , la seule manière de procéder est de choisir un entier L et de ranger dans A les entiers a tels que $0 \leq a \leq L$, et dans B les entiers b tels que $b > L$. Outre l'arbitraire du nombre L , cette manière de faire serait inadéquate pour fonder l'Analyse Infinitésimale . On va donc postuler qu'il existe une partition de $\underset{\sim}{N}$ en deux parties A et B , avec la propriété ci-dessus , qui ne soit pas associée à une limite déterminée L . Il n'y a pas de paradoxe pourvu que nous supposions que les ensembles A et B sont d'un type nouveau . Les ensembles considérés jusqu'ici s'appelleront "internes" , et obéiront au principe de récurrence ; les nouveaux ensembles que nous introduisons seront qualifiés d' "externes" [3] . Une analogie va nous aider à comprendre cette distinction . Imaginons un observateur aveugle aux couleurs ; si on lui demande de mettre à part tous les cubes en bois situés dans cette pièce , la question lui est parfaitement

intelligible --- elle décrit un ensemble interne --- alors que le problème de séparer les sphères rouges des autres n'a strictement aucun sens pour lui --- car elle définit un ensemble externe . Nous admettons que les opérations usuelles : prendre la réunion ou l'intersection de deux ensembles , former leur produit cartésien ,... ont un sens et que, appliquées à des ensembles internes (externes) , elles fournissent des ensembles internes (externes) . Par contre , nous excluons la formation de l'ensemble $P(X)$ de "toutes" les parties d'un ensemble X .

Les nombres de la classe inférieure seront appelés "limités" (ou "bornés" , mais il vaut mieux éviter l'emploi du mot "fini") ; ils forment un ensemble (strictement) externe , que l'on note $\mathbb{N}_{\text{lim}}$; les nombres de la classe supérieure sont dits "très grands" On rappelle qu'on a $n < N$ si n est limité et N très grand , mais qu'il n'existe pas de plus grand nombre limité , ni de plus petit nombre très grand . L'axiome crucial est le suivant :

<u>Axiome de saturation</u> : Soient A un ensemble interne , et B un ensemble externe contenu dans $A \times \mathbb{N}_{\text{lim}}$; il existe alors un ensemble interne B' tel que l'on ait

$$B = B' \cap (A \times \mathbb{N}_{\text{lim}}) \quad .$$

Comme conséquence , toute suite externe $a_0, a_1, \ldots, a_n, \ldots$ d'entiers , définie pour les valeurs limitées de l'indice n , se prolonge en une suite interne $a_0, a_1, \ldots, a_N$. Le point important est que , même s'il est très grand , l'entier N a les propriétés usuelles , et la suite $a_0, a_1, \ldots, a_N$ est une <u>suite finie</u> ; on peut donc définir la somme $a_0 + a_1 + \ldots + a_N$ ou le produit $a_0 a_1 \ldots a_N$. Les formules usuelles , telles que

$$0 + 1 + 2 + \ldots + N = N(N+1)/2 \quad , \quad 1.2.\ldots.N = N!$$

restent valables pour un très grand entier N .

Que la frontière entre les nombres limités et les nombres très grands soit floue s'exprime de plusieurs manières . Par exemple , un ensemble interne qui contient des nombres limités aussi grands que l'on veut (c'est-à-dire que tout nombre limité de l'ensemble en question est suivi d'un nombre limité de l'ensemble encore plus grand) contient des nombres très grands . Si une propriété interne est valable pour tous les très grands entiers , il existe un nombre limité L tel que la propriété soit vraie de tous les entiers n tels que $n \geq L$. De telles propositions permettent de transformer facilement un énoncé de type courant en un énoncé portant sur les très grands nombres . Donnons un exemple : l'assertion

« il existe un entier n_0 tel que l'on ait $n^n > 10.n!$ pour tout $n \geq n_0$ »

devient

« si N est un très grand nombre , on a $N^N > 10.N!$ » .

Le glissement dans le langage est si insensible qu'il pourrait passer inaperçu ; la nouveauté est que l'expression « N est un très grand nombre » n'est pas une manière plus ou moins relachée de parler , mais une expression mathématique à l'emploi parfaitement codifié .

Nous pouvons maintenant donner une construction très simple des nombres réels . Introduisons le calcul des fractions a/b , avec a,b entiers , avec les règles usuelles, d'où le corps $\mathbb{Q}$ des nombres rationnels . Un nombre rationnel x s'écrit comme somme d'une partie entière n et d'une partie fractionnaire r telle que $0 \leq r < 1$; on dira que x est <u>limité</u> , ou <u>très grand</u> , selon que $|n|$ est un entier avec cette propriété . On dira qu'un nombre est <u>très petit</u> s'il est nul , ou s'il est l'inverse d'un nombre très grand ; on dira aussi qu'un nombre est <u>appéciable</u> s'il est limité , mais non très petit . Les nombres rationnels sont donc rangés en trois classes :

« très petits » « appréciables » « très grands » .

Dans le langage de l'algèbre , les nombres limités forment un sous-anneau $\mathcal{O}$ du corps $\mathbb{Q}$, et les nombres très petits forment l'unique idéal maximal $\mathfrak{m}$ de $\mathcal{O}$. On peut alors définir le corps $\mathcal{O}/\mathfrak{m}$, qui sera notre modèle des nombres réels .

De manière plus détaillée , on écrira $a \simeq b$ si a et b sont deux nombres (rationnels) tels que $a - b$ soit très petit . Si a est un nombre limité , on note a° son halo , c'est-à-dire l'ensemble (externe) formé des nombres b tels que $b \simeq a$; l'égalité approchée $a \simeq b$ équivaut donc à l'égalité stricte $a^\circ = b^\circ$. L'addition et la multplication sont définies de telle sorte que l'on ait

$$(a + b)^\circ = a^\circ + b^\circ \quad , \quad (ab)^\circ = a^\circ . b^\circ \quad .$$

Par définition , un nombre réel est donc un halo , ce qui correspond à l'idée du physicien qu'une mesure ne peut jamais être parfaite , et qu'il reste toujours une incertitude , fût-elle très petite .

L'exponentielle d'un nombre réel x se définit ainsi : choisissons un nombre rationnel a tel que $x = a^\circ$, et un entier très grand Ω ; on pose alors

$$\exp x = \left((1 + \frac{a}{\Omega})^{\Omega}\right)^\circ$$

(il faut naturellement prouver que le résultat est indépendant de a et Ω) ; c'est la définition d'Euler . Alors , si $\varepsilon_1, \ldots, \varepsilon_N$ est une suite interne de nombres très petits tous de même signe , telle que $x \simeq \varepsilon_1 + \ldots + \varepsilon_N$, on aura

$$\exp x \simeq (1 + \varepsilon_1) \ldots (1 + \varepsilon_N) \quad .$$

Pour l'ingénieur --- ou le banquier --- cette formule traduit l'essence de l'exponentielle . Pour le logarithme , on a une formule analogue

$$\log(1 + x) \simeq \sum_{i=1}^{N} \varepsilon_i/(1 + \varepsilon_1 + \ldots + \varepsilon_i) \quad .$$

7. Calcul des probabilités

Dans la théorie élémentaire des probabilités , on introduit le modèle suivant d'une expérience aléatoire : on se donne un ensemble fini Ω , dont les éléments $\omega_1,\ldots,\omega_N$ représentent les résultats possibles de l'expérience ; à chaque élément ω_i est associé un nombre positif p_i , qui représente la probabilité d'obtenir ω_i , et l'on a $p_1 + \ldots + p_N = 1$. Ceci s'exprime par le schéma

$$\begin{pmatrix} \omega_1 & \omega_2 & \cdots & \omega_N \\ p_1 & p_2 & \cdots & p_N \end{pmatrix} \quad .$$

Un événement --- autrement dit , une propriété dont la satisfaction dépend du résultat de l'expérience --- est identifié à une partie A de Ω , formée des éléments ω_i pour lesquels l'événement se produit ; la probabilité $\Pr[A]$ de A est la somme des probabilités élémentaires p_i pour tous les éléments ω_i de A . Le cas le plus simple est celui où toutes les probabilités élémentaires p_i sont égales , donc à $1/N$; on a alors $\Pr[A] = |A|/N$, où $|A|$ désigne le nombre d'éléments de l'ensemble fini A . Une variable aléatoire est un nombre dont la valeur dépend du résultat de l'expérience ; on l'identifie à l'application X de Ω dans $\underset{\sim}{R}$ qui à ω_i fait correspondre la valeur x_i prise par X lorsque ω_i est réalisé . L'espérance --- ou valeur moyenne --- de X est le nombre

$$E[X] = p_1x_1 + \ldots + p_Nx_N \quad ;$$

dans le cas homogène $p_i = 1/N$, l'espérance $E[X]$ se confond avec la moyenne arithmétique $(x_1 + \ldots + x_N)/N$.

Introduisons les probabilités conditionnelles par rapport à un événement A . Si l'on fait des prévisions conditionnellement à A , on considère que A s'est déjà produit , et l'on écarte les éléments de Ω qui n'appartiennent pas à A ; les nouvelles probabilités élémentaires seront proportionnelles aux p_i ; par exemple , si A se compose des éléments $\omega_1,\ldots,\omega_M$ de Ω, on aura le schéma

$$\begin{pmatrix} \omega_1 & \omega_2 & \cdots & \omega_M \\ q_1 & q_2 & \cdots & q_M \end{pmatrix}$$

avec $q_i = p_i/(p_1 + \ldots + p_M)$ pour $1 \leq i \leq M$. La probabilité conditionnelle d'un événement B calculée selon ce nouveau schéma sera

$$\Pr[B|A] = \Pr[B \cap A]/\Pr[A] \quad .$$

L'espérance conditionnelle d'une variable aléatoire X sera définie par

$$E[X|A] = q_1 x_1 + \ldots + q_M x_M = E[X.I_A]/Pr[A] ,$$

où la variable aléatoire I_A prend la valeur 1 lorsque A se réalise et 0 dans le cas contraire .

Le modèle précédent a été longtemps considéré comme inadéquat pour rendre compte des probabilités continues , et conduit à de nombreux paradoxes classiques . Les recherches des années 1920 , conduites par des mathématiciens tels que Borel , Wiener , Zygmund et Steinhaus , ont abouti à la création du calcul des probabilités moderne ; une axiomatisation réussie est due à Kolmogoroff (vers 1930) , et depuis 1960 environ , elle a été acceptée de manière quasi-universelle . Dans cette présentation , la théorie de la mesure et de l'intégrale de Lebesgue , généralisée aux espaces "abstraits" , est le préliminaire indispensable --- et pesant . Tout récemment (1984) , Edward Nelson a réussi à appliquer les méthodes infinitésimales , en modifiant de manière minime le modèle donné plus haut .

La nouveauté consiste à supposer que le nombre N des points de Ω est très grand (au sens axiomatique considéré au numéro précédent) , et que chacune des probabilités élémentaires p_i est très petite --- ce qui est automatique dans le cas homogène $p_i = 1/N$. On dira que l'espace probabilisé Ω est hyperfini dans ce cas . Les événements correspondent aux parties internes de Ω et , moyennant quelques précautions , à certains ensembles externes . On dira qu'un événement A est presque-sûr si l'on a $Pr[A] \simeq 1$; ceci est conforme à l'adage énoncé par Emile Borel : un événement de probabilité trop petite ne se produit pas --- ou l'on peut faire comme si c'était le cas . Si X et Y sont deux variables aléatoires , on écrit $X \simeq Y$ s'il existe un événement A tel que $Pr[A] \simeq 1$ et $X(\omega) \simeq Y(\omega)$ pour tout point ω de A ; d'après l'inégalité de Čebyčev, ce sera le cas si $E[X^2] \simeq 0$.

A titre d'illustration , considérons un jeu illimité de pile ou face . Classiquement , un modèle est donné par l'espace Ω_∞ des suites infinies

$$\omega = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \ldots)$$

formées de 0 et de 1 , avec la convention que $\varepsilon_n = 1$ signifie que l'on fait pile à la n-ième partie , alors que l'on a $\varepsilon_n = 0$ si "face" est sorti au n-ième coup . La construction de la σ-algèbre (ou tribu) des événements probabilisables , et de leur probabilité , requiert l'usage des théorèmes fondamentaux de la théorie de la mesure , par exemple le théorème de prolongement de Carathéodory . Dans la manière de Nelson , on suppose que l'on a joué un très grand nombre N de parties , et l'on considère l'en-

semble Ω_M des suites (internes) $\omega = (\varepsilon_1,\ldots,\varepsilon_M)$ formées de 0 et de 1 ; en utilisant le développement de base 2 des entiers, on énumère les éléments de Ω_M sous forme d'une suite $\omega_1,\ldots,\omega_N$ avec $N = 2^M$; il reste à poser $p_i = 1/N$ pour tout i.

Le lien entre les deux descriptions est le suivant. Etant donnés deux points distincts $\omega = (\varepsilon_1,\ldots,\varepsilon_M)$ et $\omega' = (\varepsilon'_1,\ldots,\varepsilon'_M)$ de Ω_M, leur distance $d(\omega,\omega')$ est égale à 2^{-j} où j est le plus petit entier tel que $1 \leq j \leq M$ et $\omega_j \neq \omega'_j$; on pose aussi $d(\omega,\omega') = 0$ si $\omega = \omega'$. Le halo ω° d'un point ω de Ω_M se compose alors des points ω' tels que $d(\omega,\omega') \simeq 0$. Les divers halos forment une partition de Ω_M en parties externes --- tout point de Ω_M appartient à un halo et un seul --- et sont en correspondance bijective avec les points de Ω_∞ --- c'est une conséquence de l'axiome de saturation.

Examinons rapidement la <u>loi des grands nombres</u>. On suppose donnée une variable aléatoire X, d'espérance m ; on considère une suite indépendante $(X_1,\ldots,X_N)$ de variables aléatoires, ayant toute même loi de probabilité que X, et qui correspondent au résultat de N déterminations indépendantes de X. La moyenne empirique est la <u>variable aléatoire</u>

$$\overline{X} = (X_1 + X_2 + \ldots + X_N)/N \quad ;$$

la loi des grands nombres affirme que l'on a $\overline{X} \simeq m$ pour N très grand. Quitte à remplacer X par $X - m$, ce qui remplace $\overline{X}$ par $\overline{X} - m$, on se ramène au cas $m = 0$. Un calcul immédiat --- et classique --- donne la relation

$$E[\overline{X}^2] = E[X^2]/N \quad ;$$

compte tenu de l'inégalité de Čebyčev, et revenant au cas général, on conclut qu'on a $\overline{X} \simeq m$ pourvu que le nombre $N/\mathrm{Var}[X]$ soit très grand.

Dans le raisonnement précédent, on choisit N en fonction de la loi de probabilité de X. Voici deux conditions dont la conjonction assure que l'on a $\overline{X} \simeq m$ pour <u>tout</u> nombre N très grand :

a) le nombre $E[|X|]$ est limité ;

b) on a $E[X.I_A] \simeq 0$ pour tout événement A tel que $Pr[A] \simeq 0$.

Si la propriété a) est familière, la propriété b) est plus surprenante ; elle traduit la stabilité de l'espérance de X, et elle n'est pas superflue en vertu de l'exemple suivant. Prenons un très petit nombre $p > 0$, et supposons que X prenne la valeur 0 avec probabilité $1 - p$, et la valeur $1/p$ avec la probabilité p. On a $m = 1$ et $E[|X|] = 1$, donc la condition a) est satisfaite ; par contre, si l'on considère

l'événement $A = [X \neq 0]$, on a évidemment $X.I_A = X$, d'où $E[X.I_A] = 1$ alors qu'on a $Pr[A] = p \simeq 0$, ce qui viole la condition b) . La théorie montre --- et une simulation sur micro-ordinateur le confirme --- que la loi de la variable aléatoire $\overline{X}$ dépend de la taille du nombre $c = Np$:

--- si c est très petit , on a $\overline{X} \simeq 0$;

--- si c est appréciable , on a une loi de Poisson

$$Pr[\overline{X} = k/c] \simeq e^{-c} c^k/k! \qquad \text{pour tout entier limité } k \geq 0 ;$$

--- si c est très grand , on a $\overline{X} \simeq 1$.

La loi des erreurs de Gauss se formule de manière rigoureuse comme suit . Soit X une variable aléatoire , dont la moyenne $m = E[X]$ soit limitée et la variance $\sigma^2 = Var[X]$ soit appréciable ; si X est la somme $X_1 + X_2 + \ldots + X_N$ d'une famille indépendante de variables aléatoires , et si les valeurs prises par ces variables sont très petites ($X_i(\omega) \simeq 0$ pour $1 \leq i \leq N$ et $\omega \in \Omega$) , alors X suit (approximativement)la loi de Laplace-Gauss :

$$Pr[a \leq X \leq b] \simeq \frac{1}{\sigma\sqrt{2\pi}} \int_a^b \exp\left(\frac{(x-m)^2}{2\sigma^2}\right) dx \quad .$$

8. Théorie géométrique des équations différentielles

Considérons une équation d'évolution du second ordre de la forme $\ddot{x} = F(x,\dot{x})$. Le paramètre t désigne le temps , la quantité x est une fonction à déterminer du temps , soit $x(t)$, et suivant l'usage de Newton perpétué en Mécanique , on note $\dot{x}$ la dérivée de x par rapport à t et $\ddot{x}$ la dérivée seconde . Par un artifice classique , on remplace l'équation donnée par le système équivalent de deux équations du premier ordre

$$\dot{x} = y \quad , \quad \dot{y} = F(x,y) \quad .$$

Plus généralement , considérons un système de la forme

$$(S) \qquad \dot{x} = f(x,y) \quad , \quad \dot{y} = g(x,y) \quad .$$

Exprimons d'abord le point de vue infinitésimal sur le temps . Supposons que t varie entre deux valeurs a et b , et admettons que ce soient des nombres limités . Nous effectuons une pseudo-discrétisation de l'intervalle $[a,b]$ sous la forme d'une partie interne finie T de $[a,b]$ telle que tout point de $[a,b]$ soit dans le halo d'un point de T au moins . La possibilité d'une telle construction traduit le théorème classique de la compacité de l'intervalle $[a,b]$. Puisque T est un ensem-

ble interne et fini , on peut énumérer ses éléments sous forme d'une suite croissante $(t_o, t_1, \ldots, t_N)$ et l'on aura

$$a \simeq t_o \simeq t_1 \simeq \ldots \simeq t_i \simeq t_{i+1} \simeq \ldots \simeq t_N \simeq b \quad .$$

La notation différentielle s'introduit comme suit : si t est un point de T , il existe un entier i tel que $0 \leq i \leq N$ et $t = t_i$; si t n'est pas le plus grand élément de T , on a $i < N$ et l'on pose $dt = t_{i+1} - t_i$; si h est une fonction interne définie sur T , à valeurs numériques , on pose $dh(t) = h(t_{i+1}) - h(t_i)$, d'où la relation exacte

$$h(t + dt) = h(t) + dh(t) \quad ;$$

enfin , la dérivée est définie par $\dot{h}(t) = dh(t)/dt$.

Les fonctions que nous considérerons seront à accroissements infinitésimaux , c'est-à-dire qu'on a $h(t + dt) \simeq h(t)$. Cette propriété est plus faible que la continuité qui affirme que h oscille très peu dans le halo d'un point ; on peut encore dire que la relation $t \simeq t'$ doit entraîner $h(t) \simeq h(t')$, ou encore que h applique un halo dans un halo . De ce point de vue , toute fonction admet une dérivée et des primitives , deux primitives diffèrent entre elles par une constante , toute fonction est la dérivée d'une de ses primitives , et primitive de sa dérivée . On démontre de manière entièrement élémentaire le théorème des accroissements finis , ou des valeurs intermédiaires . Si la dérivée $\dot{h}$ de la fonction h prend des valeurs limitées , alors h est continue ; si l'on a $\dot{h}(t) \simeq 0$ pour tout t dans T , alors h est presque constante , c'est-à-dire que l'on a $h(t) \simeq h(t')$ quels que soient t et t' dans T .

Donnons maintenant une description analogue du plan . Choisissons une partie interne finie $\underset{\sim}{\Phi}$ de $\underset{\sim}{Q} \times \underset{\sim}{Q}$ telle que tout point dont les deux coordonnées sont limitées soit dans le halo d'un point de $\underset{\sim}{\Phi}$; il suffit de prendre pour $\underset{\sim}{\Phi}$ l'ensemble des couples $(a/L, b/L)$, où L est un entier très grand , avec les bornes

$$-L^2 \leq a \leq L^2 \quad , \quad -L^2 \leq b \leq L^2 \quad .$$

Un point de $\underset{\sim}{\Phi}$ est dit <u>accessible</u> si ses deux coordonnées sont limitées ; deux points $\underset{\sim}{x}$ et $\underset{\sim}{y}$ de $\underset{\sim}{\Phi}$ sont <u>très proches</u> si leur distance satisfait à $d(\underset{\sim}{x}, \underset{\sim}{y}) \simeq 0$; on écrit alors $\underset{\sim}{x} \simeq \underset{\sim}{y}$. Dans l'ensemble interne et fini $\underset{\sim}{\Phi}$, les points accessibles forment une partie externe , qu'on appelle le <u>champ d'observation</u> (ou galaxie principale) ; les classes d'équivalence de la relation externe $\underset{\sim}{x} \simeq \underset{\sim}{y}$ sont les <u>halos</u> . L'interprétation est évidente : un observateur ne peut connaître que les points accessibles

et il ne peut pas distinguer les divers points à l'intérieur d'un halo donné .

De ce point de vue , une courbe C dans le plan se représente par une suite interne de points $\underset{\sim}{p}_0, \underset{\sim}{p}_1, \ldots, \underset{\sim}{p}_N$ de Φ telle que $\underset{\sim}{p}_i \simeq \underset{\sim}{p}_{i+1}$ pour $0 \leq i < N$. Du point de vue analytique , on introduira deux fonctions numériques x et y sur T , à accroissements infinitésimaux , et telle que le point $\underset{\sim}{p}_i$ ait pour coordonnées $x(t_i)$ et $y(t_i)$. Une telle vision des courbes correspond par exemple à la trajectoire d'une particule élémentaire , telle qu'elle est observée dans une chambre à bulles d'un accélérateur .

Revenons au système différentiel (S) proposé au début de ce numéro . Supposons que les fonctions internes f et g prennent des valeurs limitées en tout point accessible de Φ . Une courbe C sera une solution du système (S) si elle est paramétrée par deux fonctions x et y telles que l'on ait

$$(S') \qquad dx(t)/dt \simeq f(x(t),y(t)) \quad , \quad dy/dt \simeq g(x(t),y(t))$$

pour tout point t de T . La question de l'existence des solutions ne se pose pas , et il est facile de préciser le problème de manière à avoir une solution particulière bien déterminée ; du point de vue de l'Analyse Numérique , nous avons discrétisé le problème , et le point $\underset{\sim}{p}_{i+1}$ se construit par un procédé explicite à partir du point $\underset{\sim}{p}_i$. La vraie question n'est pas celle de l'unicité de la solution , mais celle de la stabilité : si deux courbes C et C' sont solutions du système (S) , et si l'on a initialement $\underset{\sim}{p}(t_0) \simeq \underset{\sim}{p}'(t_0)$, a-t-on $\underset{\sim}{p}(t) \simeq \underset{\sim}{p}'(t)$ pour tout t dans T ? (On a noté $\underset{\sim}{p}(t)$ le point $(x(t),y(t))$ de la courbe C et $\underset{\sim}{p}'(t)$ a une signification analogue pour la courbe C') . Il est très profitable de reprendre les contre-exemples classiques dans cette nouvelle optique .

Les propriétés qualitatives des équations différentielles ont été étudiées au moyen de ces méthodes par l'Ecole Strasbourgeoise de Reeb [4] . Un cas particulièrement intéressant est celui des équations de relaxation , dont le prototype est l'équation de van der Pol $\varepsilon\ddot{x} + (x^2 - 1)\dot{x} + x = 0$; il est commode de la mettre sous la forme de Liénard :

$$(L) \qquad \varepsilon\dot{x} = u - F(x) \quad , \quad \dot{u} = -x$$

avec $F(x) = x^3/3 - x$. Dans le plan de coordonnées (x,u) , la courbe Γ d'équation $u = F(x)$ s'appelle la courbe lente ; ce nom se justifie par les propriétés suivantes. Si ε est très petit , x est très grand en dehors du halo [5] de la courbe Γ ; un changement de paramètre temporel de la forme $t = \varepsilon T$ montre que , partant de n'importe quel point du plan , on atteint le halo de Γ en un temps très petit , et que u

ne varie que d'une quantité très petite pendant ce temps (car on a $\dot{u} = -x$) . On en déduit que , quel que soit le point initial (x_o,u_o) , il existe une courbe C_o formée d'un nombre fini de segments de droite horizontaux et d'arcs de la courbe lente, et telle que le courbe C solution du système de Liénard issue de $p_{\varepsilon o} = (x_o,u_o)$ soit contenue dans le halo de C_o . La courbe C_o est indépendante de ε , et peut se deviner facilement sur la forme du système (L) ; on peut dire que c'est l'ombre de la solution C_o ; il est ensuite facile de montrer que le système (L) admet un cycle-limite , c'est-à-dire un régime périodique au sens de l'ingénieur . La discussion de ce système a été faite en termes "intuitifs" par Jules Haag en 1943 . Il est très facile de la justifier par nos méthodes infinitésimales ; il suffira de choisir T de telle sorte que dt/ε soit très petit pour tout t dans T --- une précaution bien connue du spécialiste d'Analyse Numérique .

9. Un bilan provisoire

Ce qui précède montre que le débat sur la nature du continu n'est pas clos , contrairement à ce que l'on aurait pu croire . La pseudo-discrétisation que nous avons introduite pour le temps et l'espace au n° 8 nous ramène au modèle pythagoricien des atomes de temps et d'espace . On sait qu'un des grands paradoxes de la Physique contemporaine est que l'on imagine une matière atomique dans un cadre d'espace-temps qui est resté continu ; de là viennent de grandes difficultés dans la Théorie Quantique des Champs , et l'impossibilité --- jusqu'à présent tout au moins --- de réaliser une synthèse de la Gravitation et de la Physique Quantique . Sans espérer trop naïvement résoudre les grands problèmes , il peut être profitable de changer graduellement notre manière de voir le continu . Les méthodes infinitésimales de l'Analyse Non Standard permettent de le faire , et s'accordent assez bien à la vision des problèmes numériques développée par l'emploi des ordinateurs . Il y a beaucoup de domaines des Mathématiques où elles n'ont pas été expérimentées , et il vaut la peine de le faire .

Comme on l'a vu dans cet exposé , il règne une grande diversité dans les présentations de l'Analyse Non Standard , et aucun point de vue ne semble l'emporter . Le point de vue de Nelson que j'ai développé à la fin me semble simple et naturel , mais il faudrait avoir plus de pratique .

Pour l'enseignement du Calcul Différentiel et Intégral , je pense qu'il vaut la peine d'expérimenter avec prudence les nouvelles méthodes infinitésimales; il reste à voir si elles facilitent l'accès des étudiants à l'Analyse .

Palaiseau , le 1er mai 1985

BIBLIOGRAPHIE SOMMAIRE

[1] Lazare CARNOT , Réflexions sur la métaphysique du Calcul Infinitésimal , Librairie Scientifique et Technique Blanchard , Paris , 1970 .

[2] Pierre CARTIER , Perturbations singulières des équations différentielles ordinaires et analyse non standard , in Séminaire Bourbaki , exposé n° 580 , année 1981/82 , Astérisque vol. 92-93 , 1982 .

[3] David HILBERT , Grundlagen der Geometrie , 10e édition , Teubner , Stuttgart , 1968.

[4] K. HRBACEK , Non standard set theory , Amer. Math. Monthly , vol. 86 , 1979 , p. 659-677 .

[5] H.J. KEISLER , Elementary calculus : an approach using infinitesimals , Prindle , Weber et Schmidt , Boston , 1976 .

[6] Robert LUTZ et Michel GOZE , Nonstandard analysis : a practical guide with applications (en langue anglo-alsacienne) , Lect. Notes in Math. , vol. 881 , Springer , 1981.

[7] Edward NELSON , Internal set theory , Bull. Amer. Math. Soc. , vol. 83 , 1977 , p. 1165-1198 .

[8] A. ROBINSON , Nonstandard Analysis , North Holland , Amsterdam , 1966 .

[9] K. STROYAN et N. LUXEMBURG , Introduction to the theory of infinitesimals , Academic Press , New York , 1976 .

NOTES DE BAS DE PAGE

[1] On note P(X) l'ensemble des parties d'un ensemble X .

[2] Curieusement , ces idées n'ont pénétré que tout récemment en Union Soviétique . Peut-être l'attribut "non standard" a-t-il paru subversif à certains , qui craignaient que cette théorie ne soit pas assez "judenfrei" .

[3] Le langage choisi est tel que les ensembles internes sont des cas particuliers des ensembles externes ; il convient donc d'appeler strictement externes les ensembles externes qui ne sont pas internes .

[4] On pourra consulter mon article cité ci-dessus pour un rapport plus détaillé .

[5] Le halo d'un ensemble est la réunion des halos de ses points .