

Pourcours et détours en analyse infinitésimale

par Jacques BAIR et Valérie HENRY

Avant-propos

Tout au long de ma carrière professionnelle, j'ai donné des cours de mathématiques générales à des élèves sortant de l'enseignement secondaire et entamant, dans le supérieur, des études au sein desquelles les mathématiques occupent une place importante; il s'agissait en effet d'étudiants inscrits en première année de bachelier (anciennement en première candidature) en administration des affaires, en sciences de gestion, en ingénieur de gestion, en sciences économiques et en ingénieur industriel. Dans les programmes de telles études, l'analyse mathématique apparaît comme étant une matière à la fois incontournable en raison de ses nombreuses applications concrètes (en physique, en économie, en méthodes mathématiques de la gestion), mais également difficile à maîtriser.

Pendant plusieurs décennies, j'ai enseigné l'analyse de manière classique, selon l'approche weierstrassienne, ainsi qu'on me l'avait appris et comme le faisaient tous les enseignants du supérieur que je connaissais : il me semblait évident de proposer à mes étudiants les définitions en $\varepsilon - \eta$ des limites, avec leurs conséquences théoriques et leurs applications.

Bien entendu, je me suis vite rendu compte que cette matière paraissait souvent compliquée pour des nombreux apprenants. Néanmoins, j'insistais sur la maîtrise de la théorie de manière à ce que mes étudiants acquièrent de la rigueur, mais aussi qu'ils comprennent le sens de ce qu'ils font, qu'ils soient capables de relier la théorie à leur intuition et à leur pratique, en d'autres termes qu'ils possèdent des outils fiables et efficaces.

Vers la fin du siècle dernier, j'ai été intéressé par les travaux des logiciens belges M. BOFFA de l'Université de Mons et A. PÉTRY de l'Institut Supérieur Industriel liégeois; ils abandonnaient les définitions en $\varepsilon - \eta$ et faisaient appel à des infiniment petits ou grands. J'avais déjà entendu parler de tels nombres qui avaient été utilisés par des fondateurs de l'analyse mathématique, notamment NEWTON et LEIBNIZ; je les avais également rencontrés dans les cours de physique (et de mécanique) que j'avais reçus à l'Université, mais les physiciens nous paraissaient alors peu rigoureux et étaient d'ailleurs critiqués pour cela par leurs collègues mathématiciens. J'ai notamment été vite séduit quand j'ai entendu le professeur montois obtenir une dérivée en travaillant avec des

nombres infiniment petits, sans faire appel au concept de limite, en évitant des difficultés souvent rencontrées par mes étudiants (notamment la manipulation des quantificateurs présents dans les définitions à la WEIERSTRASS ainsi que l'exploitation de valeurs absolues), tout en garantissant un raisonnement parfaitement rigoureux grâce à la preuve de l'existence d'infiniment petits par A. ROBINSON. Je me suis donc renseigné sur les travaux de celui-ci qui débouchaient sur une branche nouvelle des mathématiques, appelée l'*analyse non standard*, ANS en abrégé. Mais cette théorie m'est rapidement apparue trop sophistiquée et abstraite pour être présentée à mes étudiants. J'ai même invité un étudiant économiste, par ailleurs fort en mathématiques et un des meilleurs de sa promotion, à réaliser son mémoire de fin d'études sur les applications économiques de l'ANS ; le sujet est vite apparu trop ardu et a dû être abandonné.

C'est alors que mon collègue liégeois A. PÉTRY m'a fait connaître les travaux de l'américain KEISLER ; ce dernier avait mis au point une présentation pédagogique de l'ANS. Cette approche était expérimentée depuis peu auprès des futurs ingénieurs industriels formés à l'ISIL ; à première vue, elle semblait pouvoir vaincre (au moins partiellement) certains obstacles rencontrés lors d'un apprentissage classique de l'analyse ; elle paraissait également a priori assez bien adaptée aux raisonnements économiques auxquels mes étudiants sont confrontés.

Ces premières impressions favorables pour l'introduction de l'ANS dans mes cours devaient toutefois faire l'objet d'une étude approfondie ; il s'agissait notamment de s'interroger sur la manière idéale d'adapter la matière au public visé, de relier au mieux la théorie aux applications, d'adapter en conséquence les exercices à proposer aux étudiants, de former les membres de l'équipe éducative,

La partie scientifique de ce travail fut réalisée par Valérie HENRY. Son travail déboucha sur une thèse doctorale intitulée *Questions de didactique soulevées par un enseignement de l'ANS à de futurs économistes*. Cette thèse en spécialité « didactique des mathématiques » fut soutenue à l'Université Paul Sabatier de Toulouse, le 17 décembre 2004 devant la commission d'examen composée du Président N. ROUCHE, du Directeur de Thèse A. ANTIBI et des examinateurs J. BAIR, A. DELEDICQ, P. ETTINGER et B. MONTERO. Le travail écrit comportait 320 pages en format A4 ; il débutait par une analyse du contexte général de l'étude en donnant une introduction historique de l'ANS, en décrivant la formation mathématique des économistes et les conceptions de ces derniers en ce qui concerne le concept d'infiniment petit ; y étaient encore décrits les divers publics auprès desquels ont été présentées les expérimentations décrites dans l'ouvrage. La deuxième partie du mémoire était consacrée à des « réflexions didactiques sur les notions de base du calcul différentiel en liaison avec l'analyse non standard ». Une troisième partie

présentait le « microscope virtuel et notions locales ». La dernière partie introduisait l'idée nouvelle de « décalage interdisciplinaire entre mathématiques et économie ». La thèse se terminait par cinq annexes et une imposante bibliographie (comprenant 179 références).

Cette étude était théorique, mais elle se basait sur de multiples expérimentations ayant pour but de tester les nouvelles idées auprès d'étudiants ; elle me conforta dans l'idée de proposer un enseignement de l'analyse de base par une approche non standard. Progressivement mes cours évoluèrent dans cette direction, en tenant compte des réactions constructives des étudiants et des divers encadrants qui devaient également s'adapter à cette présentation nouvelle.

Après quelques années de tâtonnements, donnant lieu à des améliorations successives, le cours de première année prit une tournure (quasi) définitive, ce qui se concrétisa par la publication de divers articles sur le sujet (voir la liste à la fin de cet ouvrage) ainsi que du livre *Analyse infinitésimale, Le calcul redécouvert*, écrit en collaboration avec V. HENRY, publié chez Academia Bruylant en 2008.

Le dernier épisode de cette aventure fut l'intérêt que porta Mikhail KATZ, *Full Professor at Bar Ilan University in Ramat Gan (Israël)*, aux travaux que Valérie et moi menions dans cette direction ; il insista pour que nous intégrions tous deux son équipe internationale qui vise notamment à promouvoir l'approche robinsonienne de l'analyse. Ce groupe est composé de mathématiciens mais aussi d'historiens et de philosophes, à savoir Jacques BAIR, Piotr BLASZCZYK, Robert ELY, Valérie HENRY, Vladimir KANOVEI, Karin U. KATZ, Mikhail G. KATZ, Semen S. KUTATELADZE, Thomas MCGAFFEY, David SCHAPS, David SHERRY et Steven SHNIDER ; il a rédigé un article, intitulé *Is mathematical history written by the victors ?*, paru dans les *Notices of The American Mathematical Society* (n° 60, 2013) ; il a également introduit plusieurs questions sur le site américain *MathOverflow* ainsi qu'un projet, intitulé *Toward a pluralism in foundations of mathematics and its education, history, philosophy, and practice*, dans le cadre de *The Templeton Foundation*.

Venons-en à présent au présent ouvrage.

L'élaboration du cours d'analyse infinitésimale a généré de nombreuses réflexions, parfois de nature plus philosophique ou didactique et ne faisant pas toujours partie de notre cours proprement dit. Certaines d'entre elles n'ont pas encore fait l'objet d'une publication ; d'autres enfin ont été publiées en partie dans des revues qui n'existent plus ou ne sont guère accessibles pour le public visé présentement.

A l'occasion de mon départ à la retraite, il m'a semblé opportun de rassembler ces textes qui donnent un meilleur aperçu du travail que V. HENRY

et moi avons réalisé dans le domaine ces dernières années.

Les différents chapitres de ce livre se veulent autonomes (par exemple, chacun ayant sa propre liste bibliographique) et sont le plus souvent indépendants les uns des autres; ils peuvent de ce fait être lus éventuellement dans le désordre.

Ce travail n'aurait pas pu être réalisé sans le concours, peut-être implicite et involontaire, de nombreux étudiants (plus de 10.000) qui ont suivi le cours d'analyse mathématique que j'ai dispensé aux étudiants inscrits dans des études en gestion ou en économie à l'Université de Liège. Je leur en suis reconnaissant.

Ma gratitude s'adresse aussi aux divers assistants et assistants pédagogiques qui ont encadré efficacement ce cours. J'adresse un merci tout spécial à Colette COUNSON qui a relu ce livre et a suggéré de nombreuses améliorations de la version initiale.

Je remercie vivement André PÉTRY et Mikhail KATZ; leurs expériences et compétences m'ont été d'une aide extrêmement précieuse.

J. Bair

Liège, juin 2013

Chapitre 1

Deux approches de l'analyse

Mots-clés : Infini actuel ou potentiel ; continuum archimédien ; hyperréel ; infiniment petit ou grand ; loi de transfert ; partie standard.

Introduction

Comme l'a constaté Félix KLEIN¹ dans son ouvrage *Elementary mathematics from an advanced standpoint* ([19], p. 214), il existe non pas une mais bien deux directions possibles pour développer l'analyse mathématique :

- (A) l'approche weierstrassienne
- (B) l'approche par les infinitésimaux

Nous nous proposons d'analyser d'où peuvent provenir ces deux voies possibles et quelles en sont certaines de leurs particularités. A cet effet, nous nous basons notamment sur le travail collectif [2] consacré à l'apport des principaux acteurs qui ont joué un rôle important dans l'histoire des infinitésimaux en analyse mathématique.

1.1 Infini actuel *vs* infini potentiel

Le concept d'infini occupe assurément une place de choix en mathématiques. En effet, on peut affirmer notamment que

- il est étudié depuis longtemps

1. Il s'agit d'un illustre mathématicien allemand, né en 1849 et mort en 1925, qui est connu pour ses travaux en analyse, en topologie, en théorie des groupes et en géométrie non euclidienne. Il a donné son nom à une "bouteille" (qui est une surface fermée, sans bord et non orientable, c'est-à-dire, intuitivement, une surface pour laquelle il n'est pas possible de définir un intérieur et un extérieur) et à un groupe (à savoir un groupe à 4 éléments qui est le plus petit groupe non trivial et non cyclique). On lui doit aussi le réputé programme d'Erlangen, qui unifie les différentes géométries.

- il a fasciné des générations de mathématiciens, de philosophes et de scientifiques de diverses disciplines
- il a généré des polémiques relatives à sa nature et à son exploitation
- il est fondamental en analyse mathématique.

Penchons-nous d'abord quelque peu sur ses racines.

Le mot « infini » provient du latin *infinitus* : il se compose du préfixe *in* qui possède dans ce cas un sens négatif, et du suffixe *finis* qui signifie la borne, la limite (par exemple d'un champ). Etymologiquement, ce terme, pris adjectivement, qualifie donc ce qui n'a pas de borne, pas de limite.

Depuis l'Antiquité grecque, de nombreux grands penseurs se sont penchés sur le concept d'infini ; citons, parmi les plus importants, ZÉNON avec ses fameux paradoxes du lièvre² ou de la flèche³, ou encore EUCLIDE dans son célèbre ouvrage de géométrie *Éléments*.

Toutefois, ARISTOTE signale, dans le livre III de la *Physique*, que les mathématiciens

« n'ont pas besoin et ne font pas usage de l'infini ; ils utilisent des grandeurs aussi vastes qu'ils le souhaitent, mais limitées. »

Ainsi, ARISTOTE rejetait un infini *en acte* qui se réfère à ce qui existe réellement comme entité illimitée ou sans borne ; il n'admettait qu'un infini *en puissance* en ce sens qu'il se rattache à ce qui n'est jamais réalisé effectivement (ou en acte) et n'existe donc que virtuellement, potentiellement. Cette pensée fut prépondérante jusqu'au Moyen Âge lorsqu'apparurent deux types d'infini ([21]) :

- l'infini *catégorématique* lorsque l'on considère le tout pris collectivement, par exemple dans l'expression suivante : l'ensemble des nombres entiers est infini ;
- l'infini *syncatégorématique* qui concerne, comme l'écrit le philosophe Pierre d'Espagne (1220 - 1277, qui devint le pape Jean XXI en 1276),

« non pas la quantité de la chose sujet ou prédicat, mais la manière dont le sujet se rapporte au prédicat, et ainsi il n'est pas un terme commun, mais plutôt une disposition du sujet et un signe distributif. »

En termes moins savants, on retrouve grosso modo ici la distinction entre un infini en acte (à savoir le catégorématique), lorsqu'on considère tous les éléments d'un seul coup, et l'infini potentiel (le syncatégorématique) qui correspond au cas où il est toujours possible de trouver des éléments supplémentaires. Un exemple mathématique simple illustrant cette différence est fourni par

2. Lors d'une course entre un lièvre et une tortue, celle-ci a une avance au départ, de sorte que celui-là ne la rattrapera jamais, pensait ZÉNON, puisque pendant le temps nécessaire à combler le retard, la tortue aura progressé et repris une petite avance.

3. Par un raisonnement assez similaire à celui de la course évoquée ci-avant, la flèche ne devrait jamais atteindre le mur vers lequel elle est lancée.

les nombres naturels. L'ensemble de tous les entiers, considéré dans sa totalité, contient une infinité d'éléments : son cardinal, c'est-à-dire son nombre d'éléments, est un "nombre" qui est noté \aleph_0 par le mathématicien allemand CANTOR (1845 - 1918) ; ce nombre, qualifié de *transfini*, se réfère à un infini actuel. Par ailleurs, l'énumération des entiers est sans fin : elle évoque un infini potentiel, car il est toujours possible de trouver un successeur à n'importe quel nombre entier.

Comme il sera expliqué dans la suite de ce texte, chacune de ces deux conceptions de l'infini peut donner naissance à un développement différent de l'analyse mathématique.

1.2 Axiome d'Archimède

Le concept d'infini potentiel est exploité dans ce que l'on appelle aujourd'hui la méthode d'*exhaustion*, due initialement à EUDOXE et EUCLIDE, qui permet de calculer des longueurs, des aires planes, des volumes. Ce procédé s'appuie sur un axiome de continuité, attribué à ARCHIMÈDE mais présent dans le livre X des *Éléments* d'EUCLIDE :

« si, d'une grandeur, on retranche au moins la moitié, puis du reste au moins la moitié de ce dernier, et ainsi de suite, le résultat finit [sous-entendu après un nombre fini d'itérations] par devenir plus petit que toute grandeur non nulle fixée. »

Une version plus contemporaine de cet axiome d'ARCHIMÈDE a été formulée par David HILBERT de cette manière :

« Soit A_1 un point quelconque situé sur une droite entre les points donnés A et B . (voir la figure 1.1) :

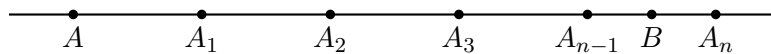


FIGURE 1.1 – Axiome d'Archimède

Construisons alors les points A_1, A_2, A_3, \dots tels que A_1 soit situé entre A et A_2 , que A_2 soit situé entre A_1 et A_3 , \dots et ainsi de suite, et tels en outre que les segments

$$AA_1, A_1A_2, A_2A_3, A_3A_4, \dots$$

soient égaux entre eux ; alors dans la série de points A_2, A_3, A_4, \dots , il existera toujours un certain point A_n tel que B soit situé entre A et A_n . » ([13], pp. 122 - 123)

Énoncée initialement dans un contexte géométrique, cette propriété s'applique désormais à d'autres structures. Dans le cadre numérique, elle peut être présentée de façon moderne comme suit :

$$(A) \quad (\forall \varepsilon > 0) (\exists n \in \mathbb{N}) \quad n\varepsilon \geq 1$$

En souvenir d'ARCHIMÈDE, on appellera *A-continuum* un ensemble de nombres vérifiant l'axiome (A). Des exemples connus sont donnés par l'ensemble \mathbb{R} des nombres réels et celui \mathbb{Q} des rationnels.

Tout axiome, étant déclaré vrai a priori, peut être nié. La négation de (A) s'écrit

$$\neg(A) \quad (\exists \varepsilon > 0) (\forall n \in \mathbb{N}) \quad n\varepsilon < 1$$

Supposé vrai $\neg(A)$ équivaut à choisir comme axiome

$$(B) \quad (\exists \varepsilon > 0) (\forall n \in \mathbb{N}) \quad \varepsilon < \frac{1}{n}$$

Dans le contexte numérique, admettre l'axiome (B) entraîne l'existence d'un nombre positif ε tel que, quel que soit le réel r positif, il existe un entier n plus grand que $\frac{1}{r}$: le nombre ε garanti par l'axiome est alors tel que $\varepsilon \leq \frac{1}{n} < r$ et est donc bien ce qu'on appelle un "infiniment petit".

C'est le suisse Johann BERNOULLI (1667 - 1748), un disciple de LEIBNIZ, qui fut un des premiers à promouvoir l'usage de tels infiniment petits ; en son honneur, on appellera *B-continuum* un ensemble de nombres vérifiant l'axiome (B). L'ensemble ${}^*\mathbb{R}$ des hyperréels et celui \mathcal{R} des superréels en sont des exemples.

Ainsi, selon que l'axiome d'ARCHIMÈDE est admis ou non, l'infini est appréhendé différemment (comme il est signalé par HAUCHART et ROUCHE dans [14], p. 351) :

- A) au sein d'un *A-continuum*, on l'approche "de biais" par une suite finie d'étapes : on est alors en présence d'un infini potentiel ;
- B) dans un *B-continuum*, on le considère "de front" grâce à l'existence de nombres infiniment petits (ou grands) et c'est alors l'idée d'infini actuel qui prévaut.

Il est encore intéressant pour notre propos de constater, à la suite de ces mêmes auteurs, que

« Archimède précisément marque une étape cruciale dans l'histoire des infinis potentiel et actuel. En effet, il a utilisé des indivisibles⁴ pour *trouver* certains de ses résultats les plus profonds, entre autres, la quadrature du segment de parabole. Mais il les a refusés pour *prouver* ces mêmes résultats : ses démonstrations

4. Les auteurs nomment ici ([14], p. 350) *indivisibles* des quantités infiniment petites non nulles. Des renseignements plus détaillés sur ce concept, étudié notamment par B. CAVALIERI, peuvent être trouvés dans [23].

recourent à la méthode d'exhaustion⁵. Ainsi, pour lui, l'infini actuel était un bon moyen d'imaginer des solutions, mais seul l'infini potentiel était acceptable dans la pensée mathématique rigoureuse. » ([14], pp. 350 - 351)

Nous allons approfondir ce point en détaillant la manière dont travaillaient les fondateurs de l'analyse mathématique, en nous attardant spécialement sur les travaux de LEIBNIZ.

1.3 Lois heuristiques de Leibniz

Souvent, dans des situations concrètes, les nombres réels représentent des quantités (ou des grandeurs). LEIBNIZ les qualifiait d'*assignables*, car elles peuvent être affectées à quelque chose de concret ; elles composent un *A*-continuum. Mais, dans sa quête d'un langage universel (qu'il appelait en latin *Characteristica Universalis*) pouvant servir pour n'importe quel type de déduction ou de découverte dans tous les champs de la connaissance, le savant allemand faisait également appel à des quantités *inassignables* : elles étaient, pour lui (comme il l'écrivait dans une lettre envoyée au Marquis de L'HOSPITAL, le 15 janvier 1696), des

« sortes de grandeurs [qui] quand elles ne seraient qu'imaginaires, peuvent servir à trouver des vérités réelles. »

LEIBNIZ introduisait ainsi des nombres « idéaux » qui sont liés d'une certaine manière au concept (mathématique et abstrait) de l'infini ; ces nouveaux nombres, qui se manipulent algébriquement comme les réels, sont de trois types :

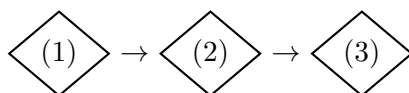
1. les *infiniment petits* qui, intuitivement, sont plus petits que toute grandeur assignable (c'est-à-dire tout ce qu'on peut "réellement" concevoir) ; plus précisément, un infiniment petit ε est un "nombre" dont la valeur absolue est inférieure à tout nombre réel positif ;
2. les *infiniment grands* qui, intuitivement, sont plus grands que tout ce qu'on peut imaginer ; de façon plus précise, un infiniment grand ω est un "nombre" non nul dont la valeur absolue est supérieure à tout nombre réel positif ; de façon équivalente, il s'agit encore de l'inverse d'un infiniment petit ;
3. les *infiniment proches* d'un réel r quelconque : ce sont les "nombres" x dont la différence $x - r$ est infiniment petite ; dans ce cas, r est le seul réel infiniment proche de x et est appelé la *partie standard* de x ; on a alors bien entendu $x = r + \varepsilon$ où $\varepsilon = x - r$ est infiniment petit.

5. Des exposés sur la méthode d'exhaustion peuvent notamment être lus dans [23] et dans [14] ; dans cette dernière référence, les auteurs signalent que « la méthode d'exhaustion emprunte clairement la voie de l'infini potentiel. Elle fixe l'attention sur des suites croissantes qui n'atteignent pas leur limite. » (p. 349)

LEIBNIZ préconisait (ainsi qu'il l'écrivait à VARIGNON dans une lettre datée du 2 février 1702) d'exploiter des grandeurs inassignables

« comme des notions idéales qui abrègent le raisonnement, semblables à ce qu'on appelle racines imaginaires dans l'analyse commune (comme par exemple $\sqrt{-2}$, lesquelles toutes imaginaires qu'on les appelle, ne laissent pas d'être utiles, et même nécessaires à exprimer analytiquement des grandeurs réelles. »

Toutes les grandeurs assignables et inassignables forment un B -continuum plus "gros" que le A -continuum habituel. En réalité, tout au long de sa carrière, LEIBNIZ travaillait dans les deux types de continuum, en procédant en trois étapes selon le schéma suivant :



où (1) désigne un problème concernant des grandeurs assignables, (2) une intervention de grandeurs inassignables et (3) un retour à des grandeurs assignables.

Le passage de l'une de ces étapes à la suivante s'effectuait selon deux lois précises, quoique non démontrées et qualifiées dès lors dans la littérature d'*heuristiques* dans le sens où elles répondaient à une sorte de "besoin intuitif" et devaient pragmatiquement conduire à la découverte de résultats nouveaux (voir notamment [2], [16]) :

- Le passage de (1) à (2) est régi par la loi de continuité (LC, en abrégé) dont une formulation simple et suggestive est la suivante :

toute propriété satisfaite par des quantités assignables l'est également par des quantités inassignables.

Ainsi, une égalité valable pour des réels l'est également pour des hyperréels quelconques, en particulier des infiniment petits (ou grands) ; par exemple, dans la formule $(x+y)^2 = x^2 + 2xy + y^2$ vraie pour tous réels x et y , on peut remplacer notamment le réel y par un infiniment petit ε et dès lors écrire $(x+\varepsilon)^2 = x^2 + 2x\varepsilon + \varepsilon^2$, donc $\Delta = (x+\varepsilon)^2 - x^2 = 2x\varepsilon + \varepsilon^2$ et partant $\frac{\Delta}{\varepsilon} = 2x + \varepsilon$.

- Le passage de (2) à (3) doit se réaliser conformément à la loi transcendante d'homogénéité (LTH, en abrégé) selon laquelle

une quantité infiniment petite peut être manipulée (algébriquement) comme toute autre grandeur⁶, mais doit être négligée⁷ par rapport à une autre lorsqu'elle est comparée à cette dernière.

6. Il s'agit de la propriété d'*homogénéité*.

7. Par une sorte de "transcendance".

En guise d'illustration, si a désigne un nombre réel et α est un infiniment petit, alors ce dernier peut être en quelque sorte éliminé dans la somme $a + \alpha$, ce qui revient en fait à remplacer la somme en question par sa partie standard a . Ainsi, en reprenant l'exemple ci-dessus, le quotient $\frac{\Delta}{\varepsilon}$ peut être assimilé simplement à sa partie standard, c'est-à-dire à $2x$; on aura reconnu ici le nombre dérivé de la fonction f définie par $f(x) = x^2$.

Les deux lois LC et LTH servaient donc à passer d'un continuum à l'autre selon la figure 1.2 (qui, au passage, suggère une façon d'imaginer chaque continuum) :

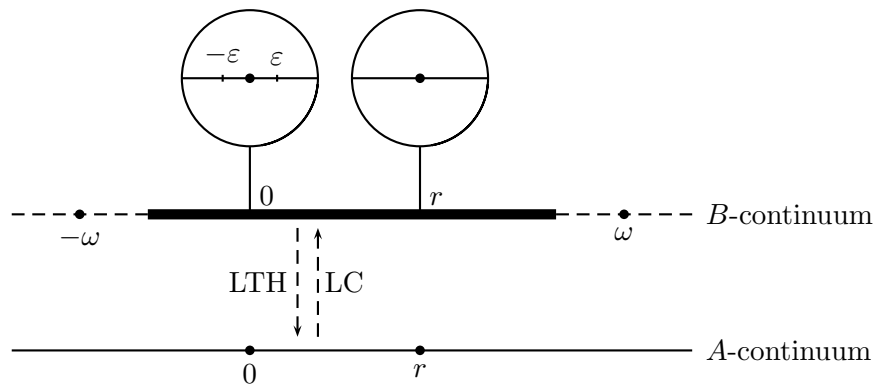


FIGURE 1.2 – Deux sortes de continuum

Il apparaît aujourd'hui que LEIBNIZ se plaçait dans un B -continuum pour inventer de nouveaux résultats ou encore pour écourter des raisonnements, mais il œuvrait dans un A -continuum pour justifier ses découvertes. En utilisant un vocabulaire dû à POLYA [25], lorsqu'il exploitait un B -continuum, le savant allemand produisait des raisonnements plausibles puisque ceux-ci reposaient sur des lois heuristiques, mais il revenait à un A -continuum pour obtenir une confirmation de ses conjectures et donc émettre un raisonnement démonstratif.

Développons une application mathématique assez spectaculaire de l'usage de ces deux lois.

Partons d'une ellipse qui peut être définie comme étant l'ensemble des points dont la somme des distances à deux points fixes, appelés les foyers, est constante. Pour fixer les idées, choisissons comme foyers les points $F_1 = (0, 0)$ et $F_2 = (a, 0)$, avec $a > 0$, et admettons que l'ellipse cherchée (voir la figure 1.3) possède le point $P = (-1, 0)$ comme sommet.

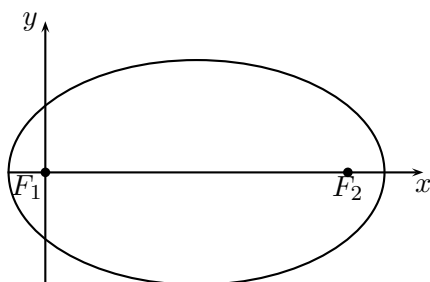


FIGURE 1.3 – Ellipse

L'équation de cette courbe est donc donnée par

$$\sqrt{x^2 + y^2} + \sqrt{(x - a)^2 + y^2} = a + 2 \quad (1)$$

En élevant au carré les deux membres de cette équation, on obtient

$$x^2 + y^2 + (x - a)^2 + y^2 + 2\sqrt{(x^2 + y^2)[(x - a)^2 + y^2]} = (a + 2)^2 \quad (2)$$

ce qui peut s'écrire sous la forme suivante

$$2\sqrt{(x^2 + y^2)[(x - a)^2 + y^2]} = a^2 + 4a + 4 - [x^2 + y^2 + (x - a)^2 + y^2] \quad (3)$$

Ceci est encore équivalent à

$$\sqrt{(x^2 + y^2)\left[\left(1 - \frac{x}{a}\right)^2 + \left(\frac{y}{a}\right)^2\right]} = 2 + x + \frac{2 - x^2 - y^2}{a} \quad (4)$$

C'est à ce moment que les deux lois heuristiques entrent en action.

La première d'entre elles est la loi de continuité affirmant grosso modo que tout ce qui est vrai dans le fini peut être étendu dans le domaine de l'infini ; cela nous permet de remplacer dans la dernière formule ci-dessus le réel a par un nombre ω infiniment grand. On peut dès lors écrire

$$\sqrt{(x^2 + y^2)\left[\left(1 - \frac{x}{\omega}\right)^2 + \left(\frac{y}{\omega}\right)^2\right]} = 2 + x + \frac{2 - x^2 - y^2}{\omega} \quad (5)$$

La deuxième loi est la loi transcendante d'homogénéité qui autorise à négliger des nombres infiniment petits lorsqu'ils sont en présence de réels ordinaires : c'est précisément le cas pour les fractions ci-dessus dont le dénominateur est l'infiniment grand ω . Cela nous conduit donc à écrire

$$\sqrt{x^2 + y^2} = 2 + x \quad (6)$$

Par une nouvelle élévation au carré des deux membres de cette égalité, on trouve

$$x^2 + y^2 = (2 + x)^2 \text{ ou encore } y^2 = 4(x + 1) \quad (7)$$

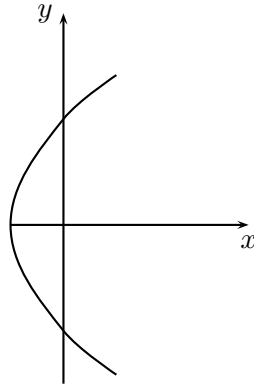


FIGURE 1.4 – Parabole

La dernière égalité est l'équation d'une parabole de sommet P et dont l'axe de symétrie est celui des abscisses (voir la figure 1.4) : c'est en effet, l'ensemble des points dont la distance à l'origine (appelée le foyer) est égale à celle de la droite verticale d'équation $x = -2$ (nommée la directrice). Pour résumer, on pourrait dire que cet exercice "de prestidigitation" se déroule en sept étapes (c'est-à-dire avec sept formules), qui sont reprises, avec leur explication, dans le tableau suivant :

Formules	Justifications
(1)	Définition d'une ellipse
(2) à (4)	Manipulations algébriques élémentaires
(5)	Loi de continuité
(6)	Loi transcendante d'homogénéité
(7)	Manipulations algébriques et définition d'une parabole

Comment interpréter mathématiquement ce "tour de magie" ?

Les calculs élémentaires effectués montrent que, lorsqu'un des foyers de l'ellipse s'éloigne très fort de l'autre, la partie visible de cette courbe située près du sommet s'apparente à une parabole. Une explication géométrique peut être donnée en considérant les courbes en question comme des intersections d'un cône avec divers plans :

- une ellipse est l'intersection d'un cône de révolution avec un plan traversant de part en part le cône ;
- une parabole est l'intersection d'un cône de révolution avec un plan parallèle à une génératrice du cône.

1.4 Analyse classique

Les infiniment petits utilisés par LEIBNIZ, BERNOULLI, NEWTON, . . . , bien que très efficaces puisqu'ils ont permis de découvrir certains résultats fonda-

mentaux en calcul différentiel ou intégral, furent très vite critiqués, notamment par BERKELEY⁸, pour des motifs de diverses natures

- a) *métaphysique* : est-il concevable d'imaginer, sans référent empirique, l'existence de quelque chose de plus petit que tout ce qui est concevable ?
- b) *méthodologique* : dans certains de leurs raisonnements, par exemple pour trouver une tangente à une courbe ou un minimum atteint par une fonction, les pionniers de l'analyse supposaient qu'un infiniment petit ε est tantôt différent de 0 ce qui les autorisait à effectuer une division par ε , tantôt égal à 0 ce qui leur permettait d'obtenir le résultat final souhaité (voir, par exemple, une telle argumentation par FERMAT dans [3]).
- c) *logique* : l'existence d'infiniment petits mène à un paradoxe ; en effet, il est aisé d'imaginer que la somme de deux infiniment petits est infiniment petite, de même que la somme de trois infiniment petits, et plus généralement de proche en proche, la somme d'un nombre quelconque n d'infiniment petits, mais ceci contredit l'axiome d'ARCHIMÈDE dès que n devient très grand !

Pour éviter de telles difficultés, d'éminents mathématiciens tels que LAGRANGE, CAUCHY, ... s'efforcèrent, dans un premier temps, à justifier les résultats de leurs illustres prédécesseurs en abandonnant l'idée de travailler non plus dans un B -continuum mais en restant constamment dans un A -continuum. Cette aventure fut longue, car il fallut notamment

- préciser ce qu'on entend exactement par fonction en mathématiques ; la définition connut des améliorations successives et ne prit la forme généralement admise de nos jours que grâce aux travaux de DIRICHLET (voir, par exemple, [6])
- définir rigoureusement les nombres réels, ce que parvint à faire DEDEKIND avec le concept de "coupure"
- introduire formellement le concept de limite, ce que réalisa définitivement WEIERSTRASS aux environs de 1870 :

« le concept de limite en (ε, N) pour les suites (et en (ε, δ) pour les fonctions) marque l'abandon⁹ des infiniment petits et un retour franc vers l'infini potentiel : il est lié à la possibilité de trouver pour tout ε fini, un N fini approprié, ε aussi petit soit-il, N aussi grand que nécessaire, mais fini lui aussi. » ([14], p. 354)

8. Il s'agit d'un évêque irlandais, né en 1685 et mort en 1753, qui était un philosophe adepte de l'empirisme. En mathématiques, il est surtout connu pour avoir publié en 1734 son ouvrage *l'Analyse* au sein duquel il critique les théories mathématiques de l'époque sur les infiniment petits.

9. Les infiniment petits furent assez vite délaissés et même parfois cloués au pilori, par exemple par Bertrand RUSSEL, un grand penseur du 20^{ème} siècle, qui écrivait : « *Infinimals as explaining continuity must be regarded as unnecessary, erroneous and self contradictory.* »

Ainsi se développa l'analyse dite *classique* (AC, en abrégé) qui est encore généralement enseignée de nos jours. Elle est construite dans un A -continuum, essentiellement dans \mathbb{R} qui peut être défini comme étant un corps commutatif, ordonné, archimédien et complet¹⁰.

Cette théorie AC est inattaquable au point de vue de la rigueur ; elle forme un ensemble cohérent et efficace, capable de résoudre bien des problèmes étudiés par les précurseurs et les fondateurs de l'analyse mathématique. Certains auteurs estiment que

« Le fondement de l'analyse par WEIERSTRASS a été considéré comme une œuvre définitive. C'est le sens des paroles de Poincaré : *on peut dire aujourd'hui que la rigueur définitive a été atteinte*. Et de même HILBERT : *principalement grâce à la sagacité de Weierstrass, ces questions ont été complètement clarifiées, et aujourd'hui l'analyse est devenue dans son domaine un guide infallible et en même temps, un instrument pratique pour manier l'infini*. » ([14], p. 354)

Mais l'apprentissage et la pratique de l'AC sont souvent, pour de nombreuses personnes, difficiles en raison d'une sophistication des raisonnements due notamment au caractère potentiel de l'infini auquel on a recours et à l'exploitation de propriétés fines des nombres réels.

Afin de donner un premier aperçu d'obstacles potentiels en AC, portons notre attention sur la définition en (ε, δ) qui est à la base de toute la théorie.

Pour définir formellement qu'un nombre, supposé connu l , est la limite d'une fonction f lorsque la variable x tend vers une valeur donnée a , il convient d'utiliser explicitement, et dans un ordre bien précis, deux quantificateurs différents : l'universel \forall portant sur le ε avec, néanmoins, une clause implicite puisque l'expression "pour tout ε " signifie dans la réalité "pour tous les ε aussi petits soient-ils", et l'existentiel \exists qui affecte le δ ; remarquons encore une différence de traitement, assez subtile, entre les abscisses et les ordonnées : effectivement, le $f(x)$ *peut être rendu arbitrairement proche* de l , tandis que x *doit être pris suffisamment proche* de a .

De plus, cette définition ne semble guère naturelle ; en effet, pour que $f(x)$ possède la caractéristique voulue, il suffit que x jouisse d'une propriété adéquate : on s'attendrait à devoir procéder en sens inverse, à savoir qu'une condition sur x entraîne nécessairement une propriété de $f(x)$.

Une autre difficulté de cette introduction d'une limite réside dans le fait que les deux parties de l'implication

$$|x - a| < \delta \Rightarrow |f(x) - l| < \varepsilon$$

n'ont aucune signification prises isolément : en effet, il est impossible de définir séparément les expressions " $f(x)$ tend vers l " et " x tend vers a ". En particulier,

¹⁰. A ce sujet, recommandons à toute personne intéressée la lecture de l'article intitulé « Pourquoi l'analyse n'est pas rationnelle », par G. NOËL, paru dans [14], pp. 95 - 107.

cette présentation tient clairement compte du fait qu'une variable dépend de l'autre (ici, y est la variable dépendante et x la variable indépendante) et ne permet pas de considérer immédiatement, par exemple, le cas où un point décrit une courbe définie implicitement.

1.5 Analyse non standard

Alors que l'infini potentiel semblait avoir définitivement pris le dessus sur l'actuel grâce aux divers succès obtenus par l'AC, l'intérêt pour ce dernier est revenu au devant de la scène lors du siècle précédent, à l'occasion principalement des travaux de deux grands mathématiciens : CANTOR avec sa théorie des nombres dits "transfinis" et ROBINSON avec celle des nombres hyperréels.

Pour notre propos, nous allons uniquement nous pencher sur ce second cas.

À la fin du siècle dernier, donc, ROBINSON a prouvé de manière incontestable (c'est-à-dire selon les normes de rigueur adoptées par les mathématiciens contemporains) l'existence des nouveaux nombres en exploitant des concepts mathématiques élaborés, notamment ceux d'« ultrafiltre » et d'« ultrapuissance ».

Parmi ces nombres, on trouve, en plus des réels, des infiniment petits, des infiniment proches d'un réel et des infiniment grands qui avaient déjà été envisagés du temps de LEIBNIZ. L'analyse mathématique peut être développée avec les hyperréels : on parle alors d'*analyse non standard*, ANS en abrégé. On y travaille au sein d'un B -continuum, noté ${}^*\mathbb{R}$, qui forme un corps ordonné non archimédien contenant \mathbb{R} comme sous-corps ; au surplus, on y utilise, de manière inattaquable au point de vue de la rigueur, les deux lois heuristiques de LEIBNIZ qui portent alors (généralement) le nom de *principe de transfert* pour la loi de continuité et *règle de la partie standard* pour la loi transcendante d'homogénéité.

Avec cette découverte de ROBINSON peuvent s'estomper les critiques historiques sur les infiniment petits, grands ou proches d'un réel. Reprenons en effet une à une les objections de BERKELEY¹¹.

- a) *Critique métaphysique*. Remarquons avant tout qu'une recommandation, faite par un empiriste, d'avoir un référent concret pour tout être mathématique n'émeut guère un mathématicien contemporain dans la mesure où ce dernier manipule fréquemment des objets abstraits¹² ; en ce sens, LEIBNIZ avait une vision très moderne en proposant ses grandeurs "idéales". En construisant rigoureusement des hyperréels infiniment petits, ROBINSON n'a fait que rassurer les éventuels sceptiques ! Ceci étant constaté, il est désormais possible de trouver des exemples concrets pou-

11. À ce propos, le lecteur intéressé peut aussi consulter [17] et [26].

12. N'est-ce pas, par exemple, le cas pour tout "nombre" non entier ou encore pour des structures mathématiques parfois très sophistiquées ?

vant suggérer une sorte de nécessité (toute platonique) pour introduire des hyperréels non réels ; mais il faut alors sortir de l’habituel cadre numérique ; par exemple, la mesure d’un angle corniculaire formé par un cercle et sa tangente est un “nombre” non nul mais plus petit que tout réel positif, puisque l’on sait depuis EUCLIDE qu’il n’existe aucune droite comprise entre le cercle et cette tangente (voir notamment [4]).

- b) *Critique méthodologique.* En préalable, notons qu’il n’est pas rare d’avoir des concepts mathématiques utilisés avec efficacité avant d’être définis rigoureusement ; un exemple typique est donné par celui de dérivée (voir par exemple [9]). Il en va de même pour les infiniment petits qui ont été aussi exploités avant d’être définis. Mais, depuis les travaux de ROBINSON, la nature des infiniment petits ne fait plus de doute : ce sont des nombres non nuls et à ce titre il est permis d’effectuer une division par eux. Leur disparition en fin de raisonnement (par exemple, pour trouver ce que l’on sait aujourd’hui être un nombre dérivé, comme vu ci-dessus) ne provient nullement d’un changement de leur statut, mais est justifiée par la règle de la partie standard qui est établie rigoureusement dans l’approche robinsonienne.
- c) *Critique logique.* L’objection devient non-fondée si on adopte non pas l’axiome d’ARCHIMÈDE, mais sa négation afin de travailler dans un B -continuum (et plus dans un A -continuum comme en AC). Remarquons que des découvertes majeures en mathématiques, comme celles relatives aux géométries non euclidiennes, sont nées grâce à une modification d’un des axiomes de base admis jusqu’alors. A ce titre, l’avènement de l’ANS constitue peut-être une avancée significative dans les mathématiques contemporaines, ainsi que le prévoit ARKERYD :

« au vu de son degré général d’“applicabilité” et sa puissance formidable, l’analyse non standard mérite assurément une place parmi les méthodes fondamentales qui seront utilisées par les mathématiciens des générations futures. » (Traduction d’un extrait de [1], p. 928)

Comme l’a écrit T. TAO, un des meilleurs mathématiciens contemporains¹³,

« l’aptitude à manipuler rigoureusement les infinitésimaux est assurément un des avantages les plus connus lorsqu’on travaille avec l’analyse non standard. » (Traduction d’un extrait de [28])

13. Terence (Chi-Shen) TAO est né le 17 juillet 1975 en Australie. Il a obtenu son doctorat en sciences mathématiques à l’âge d’à peine 20 ans. Il est actuellement professeur à l’université de Californie à Los Angeles (UCLA) ; ses travaux scientifiques ont été récompensés par de nombreux prix internationaux, dont la médaille Fields (en 2006) et le prix Crafoord (en 2012). Il est un “surdoué” des mathématiques et est encore présenté comme étant un mathématicien « sympathique, modeste, aimant travailler avec ses collègues. Il s’intéresse à l’enseignement et à la diffusion des mathématiques vers tous les publics, il est curieux. » (d’après [8])

Mais, l'ANS possède d'indéniables avantages autres que la rigueur sous-jacente. En voici quelques-uns qui ne sont pas forcément indépendants les uns des autres.

- Citons d'abord le côté intuitif de la théorie. Contentons-nous d'illustrer ce point en reprenant la définition d'une limite. Cette dernière est introduite simplement en ANS, pratiquement comme on la pense. De fait, $\lim_{x \rightarrow a} f(x) = l$ se définit par l'implication suivante : si x est infiniment proche de a , alors $f(x)$ est infiniment proche de l . Cette formulation se fait, contrairement à la présentation weierstrassienne, dans le "bon sens" et, en sus, chaque proposition de l'implication peut cette fois être utilisée isolément puisqu'on sait désormais définir quand un nombre hyperréel est infiniment proche d'un autre.
- Les études locale et asymptotique des fonctions définies implicitement ne sont guère plus difficiles que celles des fonctions explicites : dans les deux cas, elles peuvent être réalisées à l'aide de microscopes ou de télescopes infiniment puissants, et sont même particulièrement simples pour des courbes algébriques.
- Selon A. DELEDICQ,

« La manipulation de nombres en lieu et place de fonctions est un des grands avantages techniques du calcul infinitésimal sur l'analyse weierstrassienne. » ([15], p. 65)

Ainsi, le modèle non standard paraît approprié dans le cas où un nombre fort petit doit être négligé face à un autre plus grand : il suffit alors d'appliquer la loi heuristique LHT de LEIBNIZ sans être contraint d'envisager une fonction tendant vers 0 comme on le fait en AC. En guise d'application concrète, notons que cette idée peut être appliquée, par exemple, par un physicien qui, dans ses calculs, est amené à négliger la constante de PLANCK (dont l'ordre de grandeur est de 10^{-34}) sans devoir faire appel à une limite égale 0 (ce qui, dans ce cas, semble absurde d'un point de vue de la physique).

- L'ANS se révèle aussi fructueuse pour bien comprendre et expliquer la démarche suivie par un praticien lorsqu'il passe d'un modèle discret à un modèle continu comme le fait un physicien introduisant une vitesse instantanée à partir de vitesses moyennes ou encore un économiste présentant un coût marginal à partir d'une variation de coût engendrée par la production d'une unité additionnelle : les grandeurs étudiées sont prises de plus en plus petites, jusqu'à devenir "infiniment petites" auquel cas on ne conserve que la "partie observable" du résultat, c'est-à-dire la partie standard correspondante. Ce point peut être approfondi en consultant la littérature spécialisée, par exemple [4] et [12].
- Un autre intérêt de l'ANS réside dans le fait que la présentation de la théorie y est souvent allégée par comparaison à une approche classique. Une explication en est donnée par A. DELEDICQ :

« Disposant d'un vocabulaire, d'énoncés et d'outils supplémentaires, il n'est pas anormal que leur utilisation simplifie et souvent éclaire la manipulation d'objets et de concepts même exclusivement classiques. Le phénomène n'est pas nouveau : on sait bien comment le passage aux complexes rend parfois lumineux, triviaux ou excitants certains résultats et situations pourtant bien réels. » ([15], p. 73)

Dans cet ordre d'idées, mentionnons, en guise d'exemples, la réduction du nombre de quantificateurs et la disparition des valeurs absolues dans la définition robinsonienne d'une limite. Relevons encore la simplification de raisonnements grâce aux règles spécifiques de l'ANS : par exemple, la règle dite de "finitude" permet de traiter des collections contenant un nombre infiniment grand d'éléments de la même manière qu'un ensemble fini. Le lecteur intéressé par cet item peut consulter, par exemple, l'ouvrage [4].

Cette vision "optimiste" de l'ANS ne doit pas occulter le prix à payer pour maîtriser cette théorie. Comme difficultés potentielles, spécifiques et majeures, mentionnons-en quatre, dont les trois dernières explicitent en quelque sorte la première :

– Selon A. DELEDICQ,

« La difficulté essentielle tient évidemment dans l'expression "analyse non standard" et dans ce que chacun peut en connaître au regard de ce qui a été publié sur le sujet (voir la bibliographie) : ces textes entraînent en effet trop rapidement le lecteur vers des questions importantes certes, mais prématurées au regard de son accoutumance avec le sujet. » ([15], p. 55)

Signalons que ce même auteur préconise de traduire l'acronyme ANS par les mots « Analyse Nouvelle et Simple » : c'est probablement humoristique ... mais peut-être assez pertinent.

– Pour être capable de démontrer effectivement l'existence d'au moins un infiniment petit, il faut donc posséder un bagage non négligeable en logique et pouvoir manipuler des objets mathématiques abstraits (comme le concept d'ultrafiltre, la règle de transfert, ...).

Mais, dans une première approche, ceci n'est peut-être pas nécessaire, car l'on peut se contenter d'une présentation axiomatique dans laquelle l'existence d'un infiniment petit est postulée dès le départ (voir par exemple les travaux [4], [18] et [24]), quitte à suggérer cette existence par des exemples concrets tels ceux fournis par des angles corniculaires ou de contact. A ce niveau de réflexion, signalons qu'en AC, il est réclamé une profonde connaissance de la structure des nombres réels ... et que celle-ci est souvent postulée sans aucune démonstration et, généralement,

demande aussi un “acte de foi”¹⁴ initial à propos des nombres réels.

- Le vocabulaire employé est fréquemment nouveau (bien qu’assez “naturel” par rapport au langage courant). Il est également abondant, notamment parce que l’ensemble des nombres n’est pas homogène en raison de la présence de plusieurs ordres de grandeurs.
- Par expérience, il nous semble que l’obstacle le plus fréquemment rencontré auprès d’étudiants entamant l’étude de l’ANS consiste pour l’apprenant à toujours savoir dans quel continuum il se situe (dans \mathbb{R} ou dans ${}^*\mathbb{R}$), et bien entendu d’être capable de passer adéquatement d’un continuum à l’autre à l’aide des lois appropriées. Pour atténuer cette difficulté potentielle, nous recommandons à nos élèves de noter différemment les grandeurs selon qu’elles appartiennent à \mathbb{R} ou à ${}^*\mathbb{R}$, par exemple en utilisant un signe distinctif, par exemple $*$, pour désigner des variables dont les valeurs ne sont pas réelles (voir [4]).

1.6 L’analyse dans l’enseignement

Remarquons d’emblée que l’enseignement a eu, par le passé, un impact considérable sur le développement de l’AC. En effet, comme il a déjà été mentionné, il y eut à la fin du 18^{ème} siècle, surtout en France et en Allemagne, des débats passionnés sur l’opportunité ou non d’utiliser des infiniment petits (ou grands), ce qui conduisit à cette époque à leur disparition progressive. Il est probable que ce courant vers une plus grande rigueur dans la théorie est dû à la création d’Écoles supérieures de qualité, comme la prestigieuse “Ecole polytechnique”, baptisée X en abrégé et créée en 1795 à Paris. En effet, le développement de l’enseignement post-secondaire eut incontestablement une double conséquence : d’une part, l’obligation pour les mathématiciens de présenter les théories débarrassées de toutes les imprécisions de l’époque, et, d’autre part, de former de nouveaux mathématiciens, souvent de grands format. Le cours d’Analyse alors enseigné à X eut un impact décisif sur l’évolution de la discipline. Un des premiers titulaires de cette chaire fut LAGRANGE qui essaya d’éviter le recours aux infiniment petits ainsi qu’en atteste le titre d’un de ses ouvrages : « Théorie analytique des fonctions - contenant les principes du calcul différentiel dégagé de toute considération d’infiniment petits ou d’évanouissants » (1797). Un peu plus tard, A. CAUCHY, lui aussi titulaire d’une chaire d’Analyse à X , continua dans la même direction ; notons toutefois, que CAUCHY lui-même n’hésitait pas à faire appel à des infiniment petits : ainsi, pour introduire la continuité d’une fonction f en un point α , il envisageait un « accroissement infiniment petit de la variable » ce qui était une manière conceptuelle de traduire la situation obtenue, de façon plus

14. A ce sujet, signalons cette citation se trouvant dans les Notes (p. 530) du très intéressant ouvrage *Traité élémentaire de la théorie des fonctions et du calcul infinitésimal* par COURNOT (1841) : « Admettez les infiniment petits comme une hypothèse, étudiez la pratique du calcul, et la foi vous viendra. »

procédurale, lorsque « la valeur numérique de la différence $f(x + \alpha) - f(x)$ décroît indéfiniment avec celle de α . » (voir [27])

Depuis lors, l'approche classique prit de l'ampleur et est devenue (pratiquement) la seule à être enseignée dans les auditoriums de nos contrées.

Par contre, l'émergence de l'ANS ne semble pas, à notre connaissance, avoir été principalement motivée par une préoccupation de type didactique.

On pourrait se demander pourquoi l'ANS se développa bien après l'AC, alors que les problèmes fondamentaux de l'analyse ont été initialement résolus par les savants des 17^{ème} et 18^{ème} siècles en faisant appel à des infiniment petits ou grands. En d'autres termes, on pourrait chercher à expliquer pourquoi l'AC a été formalisée de façon convaincante par les mathématiciens modernes bien avant l'ANS et s'est ensuite imposée dans l'enseignement supérieur.

Une réponse à cette interrogation se trouve probablement dans la manière dont les mathématiques sont construites rigoureusement. Métaphoriquement, on peut comparer l'élaboration des mathématiques à la construction d'un mur. Celui-ci est constitué de briques placées les unes au-dessus des autres et est construit couche après couche, avec un agencement bien précis ; le premier "étage" est composé de briques disposées sur le sol, puis, quel que soit l'entier n (dans les limites fixées par la hauteur voulue), le $n + 1$ ^{ème} étage est posé sur le n ^{ème} et ne peut dès lors pas être commencé si ce dernier n'est pas terminé. Cette comparaison s'applique parfaitement à notre propos : l'ANS repose essentiellement sur l'existence de nombres hyperréels, et celle-ci nécessite celle des réels ; pour reprendre l'image du mur, l'ANS se positionne, de ce point de vue, sur le $(n + 1)$ ^{ème} niveau de briques, tandis que l'AC se trouve sur le n ^{ème}.

Il résulte de ce qui précède que l'ANS peut apparaître comme étant plus profonde que l'AC, la profondeur d'une théorie mathématique étant définie par HARDY comme suit :

« Il semble que les idées mathématiques soient disposées en strates, les idées de chaque strate étant liées à la fois entre elles et avec celles des strates inférieures et supérieures par un réseau complexe. Plus la strate est basse, plus l'idée est profonde et en général difficile. Ainsi l'idée d'un "irrationnel" est plus profonde que celle d'un entier, et pour cette raison, le théorème de PYTHAGORE est plus profond que celui d'EUCLIDE. » ([11], p. 37)

Ceci peut expliquer en partie pourquoi la théorie de ROBINSON n'est pas plus souvent enseignée dans nos universités : l'ANS est souvent jugée plus difficile que l'AC.

Il existe probablement d'autres raisons.

Il convient d'observer une certaine hésitation face à tout changement : il faut généralement du temps avant que des découvertes ne soient bien assimilées, puis enseignées à des étudiants ; de plus, toute "nouveau" bouscule

des habitudes et peut générer un certain “inconfort” pour ceux qui sont amenés à transmettre leur savoir.

Mais, une autre cause plausible semble être une certaine méconnaissance de présentations pédagogiques de l'ANS : l'exposé initial de ROBINSON peut paraître assez ardu pour des débutants en analyse et l'existence de versions de l'ANS bien adaptées à un enseignement universitaire élémentaire n'est guère connue. Pourtant, depuis quelques décennies, il est possible de trouver divers cours d'ANS qui ont été expérimentés avec un certain succès auprès de nombreux étudiants ; citons entre autres les travaux de [4], [18], [20], [24], ...

Par ailleurs, certains croient erronément que les critiques historiques, semblables à celles de BERKELEY, à propos de l'usage d'un B -continuum, subsistent encore de nos jours.

Un autre motif possible provient de ce que de nombreux mathématiciens post-weierstrassiens ont pris l'habitude de regarder les travaux fondateurs de l'analyse en faisant appel au concept de limite avec sa définition classique en $\varepsilon - \eta$, théorie qui n'existait pourtant pas du temps des découvertes initiales ; ce décryptage, que l'on pourrait qualifier d'*externe*, explique effectivement les résultats des pionniers de l'analyse, mais pas leur méthode. Une lecture *interne* est possible en recourant à l'analyse non standard (voir, par exemple, [2]) : elle rend compte à la fois des résultats et des méthodes des premiers analystes et, accessoirement, elle pourrait répondre quelque peu à la préoccupation de POLYA ([25]) lorsqu'il plaidait pour une certaine introduction des raisonnements plausibles dans l'enseignement des mathématiques.

Pour terminer par une note positive, donnons un argument en faveur d'un enseignement de l'ANS au sein de nos universités ; d'un point de vue mathématique, il est certes plus anecdotique que fondamental, mais il est peut-être assez pertinent sur le plan pédagogique. Il provient de la constatation banale suivante : les élèves belges ont reçu un enseignement de l'analyse lors de leurs deux dernières années dans le secondaire, et ceci dans une présentation classique (d'ailleurs prescrite dans les programmes officiels). Dès lors, si l'analyse de base est dispensée à l'université de la même manière qu'en humanités, c'est-à-dire selon une approche classique, certains étudiants pourraient avoir l'impression de connaître la matière et, en conséquence peut-être, ne pas faire les efforts nécessaires pour approfondir suffisamment la théorie. Par contre, un cours d'ANS surprend généralement les élèves qui ne connaissent pas les concepts nouveaux récemment introduits ; ceci peut les inviter à travailler d'emblée leur nouveau cours et, surtout, générer chez eux des réflexions de type métamathématique pour découvrir des liens entre les approches ancienne et nouvelle, ce qui devrait les aider à comprendre plus en profondeur la matière.

1.7 Conclusion

En résumé, nous pouvons affirmer, à la suite du philosophe G. GRANGER ([10], pp. 27 - 28), que :

« Nous savons aujourd’hui que deux voies devaient s’offrir pour la solution du problème

- (A) Ou bien l’on élimine du langage mathématique le terme d’infiniment petit, et l’on établit, en termes finis, le sens à donner à la notion intuitive de ‘valeur limite’ . . .
- (B) Ou bien l’on accepte de maintenir, tout au long du Calcul, la présence d’objets portant ouvertement la marque de l’infini, mais en leur conférant un statut propre qui les insère dans un système dont font aussi partie les grandeurs finies. »

On sait, depuis les travaux de ROBINSON, que les deux approches sont essentiellement équivalentes : grosso modo, tout résultat vrai en AC est également valable en ANS ; par exemple, la limite d’une suite (s_n) coïncide avec la partie standard de la valeur s_ω de l’extension naturelle pour l’indice $n = \omega$, ce qui est conforme à l’intuition selon laquelle les termes de la suite sont “très proches” de cette limite pour une valeur “très grande” de l’indice.

Ce n’est donc pas forcément au plan purement mathématique qu’il convient de se placer pour déterminer laquelle de ces méthodes doit être de préférence enseignée et pratiquée, mais peut-être bien à un niveau métamathématique. Il importe de se demander ce qui convient le mieux en envisageant divers critères tels que la facilité d’assimilation (notamment en tenant compte du bagage théorique que les étudiants devraient posséder), le caractère formateur de la théorie (pour les objectifs d’apprentissage fixés), l’utilité potentielle (en liaison avec d’autres matières), . . . C’est donc un problème éminemment multicritère (qui, de ce fait, n’a pas toujours de solution optimale unique) qu’il convient de résoudre dans ce cas !

Comme nous l’avons vu, chacune de ces approches possède des avantages et des difficultés spécifiques.

En fait, les deux présentations semblent complémentaires au niveau des raisonnements développés : les raisonnements démonstratifs sont de mise en AC, tandis que l’ANS est peut-être davantage susceptible de favoriser des raisonnements plausibles basés sur l’intuition, ainsi que nous l’ont montré les travaux des pionniers de l’analyse.

En guise de conclusion, laissons le mot de la fin à une pensée de POLYA [25] à propos de l’opportunité d’utiliser tel ou tel type de raisonnement ; cet auteur affirme en substance :

« Les raisonnements plausibles sont l’apanage de tous, y compris des mathématiciens qui doivent deviner un théorème avant

de le démontrer ou encore pressentir le principe général d'une démonstration avant de la détailler.

Les raisonnements démonstratifs devraient être pratiqués par tous, y compris les non-mathématiciens, car chacun doit, même s'il n'aura que rarement l'occasion de s'en servir directement, acquérir "un élément de comparaison qui puisse lui permettre de juger les prétendues preuves de toutes sortes qui lui seront offertes dans le monde où nous vivons actuellement" ([25], p. XI). »

Bibliographie du chapitre 1

- [1] ARKERYD L., Nonstandard Analysis, *American Mathematical Monthly*, n°112, 2005, pp. 926 - 928.
- [2] BAIR J. - BLASZCZYK P. - ELY R. - HENRY V. - KANOVEI V. - KATZ K. - KATZ M. - KUTATELADZE S. - MCGAFFEY T. - SHERRY D. - SHNIDER S., Is mathematical history written by the victors ? *Notices of the American Mathematical Society*, Volume 60, n° 7, pp. 886-904. Accessible à l'adresse électronique <http://arxiv.org/abs/1306.5973>
- [3] BAIR J. - HENRY V., Étude épistémologique sur la méthode de Fermat pour la recherche d'extrémums, *Mathématique et Pédagogie*, n° 156, 2006, pp. 49 - 61.
- [4] BAIR J. - HENRY V., *Analyse infinitésimale : Le Calculus redécouvert*, Academia-Bruylant, Louvain-La-Neuve, 2008.
- [5] BAIR J. - HENRY V., L'exponentielle : une fonction à plusieurs facettes, *Losanges*, 3, 2009, pp. 31 - 37.
- [6] BAIR J. - HENRY V., Le concept de fonction au fil du temps, *Losanges*, n° 5, 2009, pp. 11 - 20.
- [7] BAIR J. - HENRY V., Infini actuel, infini potentiel, *Mathématiques & philosophie*, Bibliothèque Tangente n° 38, Editions Pole, Paris, 2010, pp. 114 - 119.
- [8] DELAHAYE J.P., Tao : l'éducation réussie d'un surdoué, *Pour la Science*, n° 390, 2010.
- [9] GRABINER J., The changing concept of change : the derivation from Fermat to Weierstrass, *Mathematics Magazine*, Vol. 56, No. 4, 1983, pp. 195 - 206.
- [10] GRANGER G., Philosophie et mathématique leibniziennes, *Revue de Métaphysique et de Morale*, 86^{ème} année, N° 1, 1981, pp. 1 - 37.
- [11] HARDY G., *L'apologie d'un mathématicien ; Ramanujan, mathématicien indien ; Bertrand Russel et le Collège de la Trinité*, Belin, Paris, 1985.

- [12] HENRY V., *Questions de didactique soulevées par un enseignement de l'analyse non standard à de futurs économistes*, thèse présentée devant l'Université Paul Sabatier (Toulouse III) en vue de l'obtention de doctorat en didactique des mathématiques, Toulouse, 2004.
- [13] HILBERT D., Les principes fondamentaux de la géométrie, *Annales scientifiques de l'E.N.S.*, 3^{ème} série, tome 17, 1900, pp. 103 - 209, version numérisée sur le site <http://www.nundam.org/>
- [14] HAUCHART C. - ROUCHE N., *Apprivoiser l'infini, un enseignement des débuts de l'analyse*, Ciaco éditeur, Louvain-la-Neuve, 1987.
- [15] HAUCHART C. - ROUCHE N., *L'enseignement de l'analyse aux débutants*, éditions Érasme et Academia, Louvain-la-Neuve, 1992.
- [16] KATZ M. - SHERRY D., Leibniz's Laws of Continuity and Homogeneity, *Notices of the AMS*, volume 59, n° 11, 2012, pp. 1550 - 1558.
- [17] KATZ M. - SHERRY D., Leibniz'infinitesimals : Their fictionality, their modern implementations, and their foes from Berkeley to Russel and beyond, *Erkenntnis* (online first), <http://arxiv.org/abs/1205.0174>.
- [18] KEISLER J.H., *Elementary calculus : an infinitesimal approach*, Prindle, Weber & Schmidt, Boston, 1986.
- [19] KLEIN F., *Elementary Mathematics from an Advanced Standpoint*, Vol. I, Traduction par HEDRICK E.R. - NOBLE C.A., Macmillan, New York, 1932.
- [20] LUTZ R. - MAKHLOUF A. - MEYER E., *Fondement pour un enseignement de l'analyse en termes d'ordres de grandeur : les réels dévoilés*, Publication de l'A.P.M.E.P., n° 103, Paris, 1996.
- [21] MUGNAI M., LEIBNIZ, le penseur de l'universel, *Pour la Science* , n° 28, août - octobre 2006, pp. 36 - 120.
- [22] NOËL G., Pourquoi l'analyse n'est-elle pas rationnelle , dans *L'enseignement de l'analyse aux débutants*, éditions érasme et academia, Louvain-la-Neuve, 1992, pp. 95 - 107.
- [23] NOËL G. - TROMPLER S., *Vers les infiniment petites*, Publication de la Commission Pédagogique de la SBPMef, Mons, 2003.
- [24] PÉTRY A., Balade en Analyse non standard sur les traces de A. Robinson, *Non Standard Analysis*, Belgian Mathematical Society, 1996, pp. 1 - 28.
- [25] POLYA G., *Les mathématiques et le raisonnement "plausible"*, Gauthier-Villars, Paris, 1958.
- [26] SHERRY D., The wake of Berkeley's Analyst : rigor mathematicae ? *Stud. Hist. Philos. Sci.*, 18, n° 4, 1987, pp. 455 - 480.
- [27] TALL D. - KATZ M., A cognitive analysis of Cauchy's conceptions of function, continuity, limit, and infinitesimal, with implications for teaching the calculus, *Educational Studies in Mathematics*, en voie de publication.

- [28] TAO T., A cheap version of nonstandard analysis; article consultable sur le blog de l'auteur à l'adresse <http://terry-tao.wordpress.com/2012/04/02/>.

Chapitre 2

Quelques réflexions sur l'existence des infiniment petits

Une brève étude sur l'ontologie¹ des *infiniment petits* ne semble pas inutile au regard des nombreuses polémiques que ce concept a suscitées tout au long des quatre derniers siècles ; il occupe une place tout à fait centrale en analyse non standard. C'est pourquoi, il convient de rappeler de façon très synthétique les grandes étapes de son histoire, puis, après une remarque sur la relativité du concept de nombre, de donner quelques arguments susceptibles de lever (à tout le moins d'atténuer) le potentiel obstacle épistémologique relatif à son existence.

2.1 Historique succinct

De nombreuses études historiques sur les infiniment petits, en liaison avec leur intervention en analyse mathématique, se trouvent dans la littérature (voir, par exemple, [2], [8], ...). Contentons-nous de résumer très schématiquement l'évolution historique de ce concept en faisant référence à la *dialectique hégélienne*² selon laquelle la connaissance se développe par "triade"³, à savoir selon trois stades appelés respectivement "thèse, antithèse et synthèse"⁴

1. En philosophie, le mot "ontologie" désigne « une spéculation sur l'être en tant qu'être, sur l'être en soi. » (*Le petit Larousse illustré*, 1989, p. 683)

2. HEGEL : philosophe allemand, 1770-1831. La dialectique est un « mode de raisonnement qui consiste à analyser la réalité en mettant en évidence les contradictions de celles-ci et à chercher à les dépasser. » (*Le petit Larousse illustré*, 1989, p. 322)

3. Triade est un terme de didactique désignant un « groupe de trois personnes ou choses étroitement associées. » (*Le petit Larousse illustré*, 1989, p. 984)

4. Le terme synthèse est quelquefois remplacé par le terme "dépassement" (voir [4], p. 174), ce qui est probablement plus adéquat ... mais probablement moins attractif.

1. Thèse

L'existence des infiniment petits est latente depuis l'Antiquité, pressentie notamment par FERMAT, PASCAL et BARROW, admise par LEIBNIZ et NEWTON, puis exploitée par leurs successeurs aux 17^{ème} et 18^{ème} siècles pour construire les fondements de l'analyse classique.

2. Antithèse

- Critique philosophique émise notamment par BERKELEY.
- Critique de logique : caractère non archimédien.
- Développement rigoureux de l'analyse dans le cadre fonctionnel (et non plus numérique), avec un abandon systématique de l'utilisation de tout infiniment petit.

3. Synthèse

- A partir des travaux de logique concernant les modèles, naissance de l'analyse non standard par une approche constructiviste due à ROBINSON.
- Approche axiomatique de NELSON.
- Diverses approches pédagogiques par KEISLER, REEB, TALL, ...

Alors que K. GÖDEL pensait, en 1973, que

« there are good reasons to believe that nonstandard analysis, in some version or other, will be the analysis of the future »

Il convient de constater que cette prédiction ne s'est pas encore réalisée puisque la plupart des mathématiciens contemporains n'ont pas encore adopté cette théorie.

2.2 Relativité de la notion de nombre

Le mathématicien croit évidemment en l'existence des nombres, mais la croyance en des nombres peut être différente pour chaque individu, car elle dépend de

- “croyances individuelles”. Par exemple, un petit enfant croît en l'existence de 2, mais pas en celle $\sqrt{2}$;
- “conceptions philosophiques individuelles”, d'une manière générale ou d'une manière spécifique en mathématiques. Par exemple, un empiriste⁵ admet l'existence de $\sqrt{2}$ (car il peut assimiler ce nombre à (la mesure de) la diagonale d'un carré de côté unitaire), mais pas nécessairement celle de $2+i$ qu'il ne peut déduire d'aucune expérience. Comme autre illustration, mentionnons qu'un intuitionniste⁶ rejette l'existence de deux nombres

5. Un empiriste veut déduire toutes ses connaissances sur le monde de ce que ses sens lui transmettent. ([5], p. 279)

6. L'intuitionnisme est une « doctrine des logiciens néerlandais HEYTING et BROUWER selon laquelle on ne doit considérer en mathématiques que les entités qu'on peut construire par intuition. » (*Le petit Larousse illustré*, 1989, p. 534)

a et b irrationnels tels que a^b est rationnel, alors que, pour $c = \sqrt{2}^{\sqrt{2}}$, ou bien c est rationnel, ou bien $c^{\sqrt{2}}$ l'est.

Comme l'a fait remarquer C. HOUZEL ([6]) :

« on sait que les fondements ultimes des mathématiques sont inaccessibles. Si bien qu'en dernier ressort on ne sait pas ce qu'est un nombre. »

2.3 Arguments apologétiques

⁷ Nous allons donner à présent quelques arguments, relevés dans la littérature et souvent d'ordre philosophique, qui peuvent être avancés en faveur de l'existence de nombres infiniment petits. Ces arguments ne sont pas nécessairement indépendants entre eux et peuvent donc se chevaucher partiellement.

- Argument de nature "métaphysique". LEIBNIZ a "inventé" les infiniment petits parce qu'il croyait à l'infini actuel pour des raisons philosophiques (notion de "monades") ; ainsi, il écrivait cette phrase célèbre

« Je suis tellement pour l'infini actuel, qu'au lieu d'admettre que la nature l'abhorre, comme on le dit vulgairement, je tiens qu'elle l'affecte partout, pour mieux marquer les perfections de son Auteur. Ainsi je crois qu'il n'y a aucune partie de la matière qui ne soit, je ne dis pas divisible. » (cité dans [10], p. 15)

- Argument de type "nominaliste" ⁸ : on peut donner le nom d'infiniment petit à tout nombre nouveau créé de manière à être non nul et plus petit que tout réel positif, ou encore à toute suite de nombres réels jouissant de certaines propriétés algébriques. C'est d'ailleurs un peu comme cela que HAMILTON a introduit l'ensemble des complexes comme ensemble de couples muni de deux lois additive et multiplicative possédant des propriétés particulières.
- Argument de type "pragmatique" ⁹. L'histoire montre en effet que l'on peut concevoir un infiniment petit comme un artefact ¹⁰ qui se révèle efficace pour mettre en évidence des propriétés locales en analyse mathématique et qui finalement disparaît dans le résultat final. Par ailleurs, selon

7. Une apologie est un « discours ou écrit destiné à convaincre de la justesse de quelque chose, à assurer la défense de quelqu'un, de quelque chose. » (*Le petit Larousse illustré*, 1989, p. 74)

8. Le nominalisme est une « doctrine philosophique selon laquelle le concept n'est qu'un nom et n'existent effectivement que les individus auxquels renvoient le nom. » (*Le petit Larousse illustré*, 1989, p. 667)

9. L'adjectif "pragmatique" signifie « fondé sur l'action, la pratique, cautionné par la réussite. » (*Le petit Larousse illustré*, 1989, p. 775)

10. Il s'agit d'un terme de didactique qui désigne un « phénomène d'origine artificielle ou accidentelle, rencontré au cours d'une observation, d'une expérience. » (*Le petit Larousse illustré*, 1989, p. 88)

ARKERYD (2005), la droite hyperréelle et l'analyse non standard fournissent des modèles bien plus riches et plus efficaces que les classiques structures réelles, notamment pour les applications des mathématiques dans les phénomènes naturels et en économie.

- Argument de “symétrie entre les ordres extrêmes de grandeur”. PASCAL avait déjà perçu l'existence de trois ordres de grandeur : l'appréciable (à l'échelle humaine), l'infiniment petit et l'infiniment grand ; il existe clairement un lien évident entre les deux ordres extrêmes (l'inverse d'un infiniment petit est infiniment grand, et réciproquement). Or, CANTOR a démontré rigoureusement l'existence de nombres (les transfinis \aleph) infiniment grands. Donc, par un principe de “symétrie” entre les infiniment grands et les infiniment petits, il doit aussi exister des nombres infiniment petits.
- Argument de “non contradiction des axiomes”. A la suite de NELSON et de KEISLER, on peut présenter axiomatiquement l'ensemble des hyperréels contenant notamment les infiniment petits. Or, selon HILBERT,

« la non contradiction d'une famille d'axiomes assure l'existence d'une série d'objets vérifiant ces axiomes. » (cité dans [11])
- Argument d' “indécidabilité de l'axiome d'ARCHIMÈDE”. Si CANTOR, qui avait admis l'existence de nombres transfinis infiniment grands, rejetait l'existence des infiniment petits, c'est parce que ceux-ci violent le principe d'ARCHIMÈDE ($\forall x > 0, \forall y > 0, \exists n \in \mathbb{N}, n \times x > y$). Or, les travaux des logiciens, tels COHEN et GÖDEL, ont démontré que l'hypothèse du continu (il n'existe aucun nombre transfini strictement compris entre \aleph_0 et \aleph), qui se ramène au principe d'ARCHIMÈDE, est en réalité “indécidable”, c'est-à-dire qu'on ne peut ni la démontrer, ni la réfuter.
- Argument d' “indispensabilité”. Le logicien et philosophe QUINE, se basant sur les développements de la logique dite des modèles, a développé la thèse de la *relativité de l'ontologie* selon laquelle

« il n'y a pas de sens à dire ce que sont les objets d'une théorie, on peut tout juste dire comment les interpréter ou réinterpréter cette théorie dans une autre. » ([3], p. 178)

Le même auteur avance également l'idée selon laquelle les croyances aux objets mathématiques peuvent être justifiées par notre connaissance scientifique. En conséquence, on peut appeler “nombre infiniment petit” toute construction obéissant à certaines lois de l'arithmétique et vérifiant la caractéristique des infiniment petits mise en évidence par LEIBNIZ et exploitée en analyse. (voir [3], p. 178).

- Argument de la “perception sensitive”. MADDY croit au *réalisme*¹¹ en

11. En philosophie, le réalisme est « une doctrine qui affirme que la connaissance du réel

mathématiques, et plus précisément aux croyances perceptives¹² et aux inférences à partir de ces croyances. On sait déjà que tout réel peut être assimilé à un point sur la droite numérique, mais peut-on “réellement” voir tous les réels très petits ? Certainement pas directement, mais bien, en théorie, avec un “microscope agrandissant suffisamment” la région de la droite numérique aux alentours du point représentatif du nombre 0 ; de la même manière, on peut imaginer que l’on verrait sur la droite numérique un point associé à un infiniment petit si l’on disposait d’un microscope “infiniment puissant”. Cette démarche, qui fait penser à la manière que l’on peut voir des électrons à partir de données sur un écran d’ordinateur, peut être pressentie en effectuant des zooms successifs à l’aide d’un logiciel adéquat. Une autre façon d’imaginer concrètement l’existence d’infiniment petit consiste à mesurer un angle de contact¹³

- Argument d’“achèvement élémentaire”. Malgré ce titre, ce point est plus technique et ne sera qu’introduit ici ; il est dû à T. TAO. Ce dernier fait remarquer que le champ \mathbb{Q} des rationnels et celui \mathbb{R} des réels sont *algébriquement incomplets*, car il existe des équations algébriques non triviales (telles que $x^2 - 2 = 0$ dans le cas des rationnels, ou $x^2 + 1 = 0$ pour les réels) qui peuvent *potentiellement* avoir des solutions (parce qu’elles n’impliquent pas nécessairement un énoncé faux, comme $0 = 1$ par exemple, en utilisant uniquement des lois de l’algèbre), mais ne peuvent pas *actuellement* avoir des solutions dans le champ en question. Cette constatation amène TAO à présenter l’analyse non standard comme un “achèvement élémentaire” (*elementary completion* en anglais) de l’analyse standard, un peu de la même manière que, par exemple, l’algèbre complexe est un achèvement de l’algèbre réelle, ou que la géométrie réelle métrique est un achèvement de l’algèbre rationnelle métrique. Le lecteur intéressé par cette théorie peut consulter l’article original [9].

Bibliographie du chapitre 2

- [1] ARKERYD L. : Nonstandard Analysis, *Amer. Math. Monthly*, 112, 2005, pp. 925-928.

constitue le réel lui-même, que cette connaissance soit la seule réalité ou qu’à côté d’elle figure une autre réalité, l’objet auquel elle s’applique. » (*Le petit Larousse illustré*, 1989, p. 819)

12. L’adjectif “perceptif” se réfère au mot “perception” qui est polysémique ; ici, ce dernier est pris dans le sens psychologique, à savoir « représentation consciente à partir des sensations. » (*Le petit Larousse illustré*, 1989, p. 725)

13. D’autres appellations anglaises pour ce type d’angle sont les suivantes : *angle of contigence*, *curvilinear angle*, *angle of contact*, *cornicular angle*, *horn-shape angle*, *horn-like angle*, *horned angle* ; il s’agit d’un angle fait par une courbe avec une de ses tangentes. Il est clair que la mesure d’un tel angle ne peut pas être nulle (puisque l’angle ne coïncide pas avec l’angle nul), mais qu’elle doit être inférieure à tout nombre réel positif (puisque’il est impossible d’insérer un angle rectiligne entre une courbe et sa tangente).

- [2] BAIR J. - HENRY V. : De l'Analyse Classique à l'Analyse Non Standard, *Les cahiers de la mathématique appliquée*, n°1, 2003, pp. 51-74 ; n°2, 2003, pp. 75-96.
- [3] BOULEAU N. : *Philosophies des mathématiques et de la modélisation, du chercheur à l'ingénieur*, L'Harmattan, Paris, 1999.
- [4] CHÂTELET F. : *Une histoire de la raison*, Editions du Seuil, Paris, 1992.
- [5] GAARDER J. : *Le Monde de Sophie - Roman sur l'histoire de la philosophie*, Editions du Seuil, Paris, 1995.
- [6] HOUZEL C. : Qu'est-ce qu'un nombre ? , *Dossier La Recherche*, Hors-Série n° 2, 1999, pp. 6-9.
- [7] MADDY P. : *Realism in Mathematics*, Oxford University Press / Clarendon Press, Oxford, 1990.
- [8] NOËL G. - TROMPLER S. : *Vers les infiniment petits*, Brochure de la Commission Pédagogique de la SBPMef, Mons, 2003.
- [9] TAO T., Nonstandard Analysis as a completion of standard analysis, <http://terrytao.wordpress.com/2010/11/27/>.
- [10] VIDAL C. : *Georg Cantor et la découverte des infinis*, Mémoire de maîtrise en philosophie, <http://clement.vidal.philosophons.com/>.
- [11] <http://www.reunion.iufm.fr/recherche/irem/histoire/> : *construction des réels*, site de l'IUFM de la Réunion.

Chapitre 3

Infinitésimaux dans un cadre fonctionnel

Résumé. Ce texte, basé sur un travail de D. TALL [2], introduit un ensemble de nombres, nommés des *superréels*, qui inclut celui des réels mais qui contient au surplus des *infinésimaux*, c'est-à-dire des nombres inférieurs à chaque réel positif. Une particularité de cet exposé, extrait d'une annexe de [1], consiste à travailler dans le cadre fonctionnel et d'être de ce fait assez conforme à l'intuition.

Introduction

L'analyse élémentaire (des fonctions à une variable) se développe classiquement au sein de l'ensemble des nombres réels, car elle repose de façon fondamentale sur la structure algébrique de \mathbb{R} , entre autres sur le fait que $\langle \mathbb{R}, +, \times, 0, 1 \rangle$ est un champ ordonné.

Par ailleurs, il est bien connu que l'hypothèse de l'existence d'un nombre *infiment petit*, c'est-à-dire un infiniésimal non nul, au sein de l'ensemble des réels est contradictoire. Il importe dès lors de sortir du cadre numérique traditionnel pour définir un nouvel ensemble comprenant notamment les nombres réels et des infiniésimaux. Idéalement, cet ensemble devrait encore avoir la structure d'un champ ordonné et admettre comme sous-corps un ensemble isomorphe au champ $\langle \mathbb{R}, +, \times, 0, 1 \rangle$ des nombres réels.

Il existe diverses manières d'étendre le corps des réels en une nouvelle structure de corps ordonné. David TALL [2] propose une construction d'un corps \mathcal{R} des superréels qui répond aux souhaits exprimés ci-dessus et qui se réalise en deux étapes.

- Dans un premier temps est construit un corps ordonné, dont les éléments, qui sont des fonctions, forment un ensemble \mathbb{R}' , certains de ces éléments pouvant être identifiés aux nombres réels, d'autres correspondant à des

infinitésimaux. Cette construction montre la possibilité d'introduire assez concrètement ces infinitésimaux, mais ne permet pas de couvrir tous les besoins (exigeants).

- Dans un second temps, on est amené à élargir encore \mathbb{R}' en un sur-ensemble \mathcal{R} muni d'une structure de corps ordonné : les éléments de \mathcal{R} sont appelés des *superréels* et contiennent entre autres tous les réels.

3.1 Existence d'infinitésimaux

Les infinitésimaux non nuls ne pouvant pas être des nombres réels, une voie possible pour les introduire consiste à travailler dans le cadre fonctionnel, ce qui permet de visualiser les problèmes grâce aux graphes des fonctions considérées. Il est à noter que ce recours à des fonctions pour introduire des nombres infinitésimaux n'est pas neuf : elle avait déjà été suivie dès le 18^{ème} siècle, notamment par A. CAUCHY (1789 - 1857) qui percevait un infiniment petit comme une "fonction qui tend vers 0".

Il n'est pas difficile de faire correspondre univoquement à tout réel r une fonction constante et définie sur \mathbb{R} , à savoir $f : x \mapsto r$, dont le graphe est une droite horizontale composée de tous les points d'ordonnée r dans le plan. Dans ce contexte, les infinitésimaux non nuls devront être associés à des fonctions non constantes. Ainsi, pour déterminer les fonctions qui vont être retenues dans cette étude, il importe de travailler sur un ensemble comprenant les fonctions constantes, et sur lequel pourra être instaurée une algèbre ; il convient en effet de pouvoir additionner et multiplier les infinitésimaux entre eux ou avec des réels, ainsi que de les comparer entre eux et avec les réels. En termes plus techniques, il faut construire un corps ordonné de fonctions, de préférence assez élémentaires, parmi lesquelles on retrouverait entre autres les fonctions constantes.

Il s'avère opportun de se tourner vers des fonctions simples et bien connues pour réaliser une telle construction. Parmi ces dernières, les polynômes sont des fonctions qui, a priori, pourraient convenir puisqu'ils peuvent être ajoutés et multipliés entre eux ; malheureusement, ils ne forment pas un corps puisque l'existence d'un inverse pour la multiplication n'est pas assurée. C'est pourquoi, on est naturellement amené à considérer les fractions rationnelles, ou quotients de deux polynômes dont le dénominateur n'est pas nul : elles forment un corps commutatif, noté $\mathbb{Q}[x]$, lorsque sont utilisées les opérations classiques d'addition et de multiplication.

On définit alors sur ce corps $\mathbb{Q}[x]$ un ordre qui permet d'introduire des infinitésimaux en exploitant le graphe des fonctions considérées. Pour deux fonctions constantes, l'ordre est évidemment dicté par la position respective de leur graphe : la fonction constante $f_1 : x \mapsto r$ est déclarée inférieure à une autre fonction constante $f_2 : x \mapsto s$ si et seulement si la droite horizontale d'équation $y = r$ est située plus bas dans le plan que la droite d'équation $y = s$

(voir la figure 3.1).

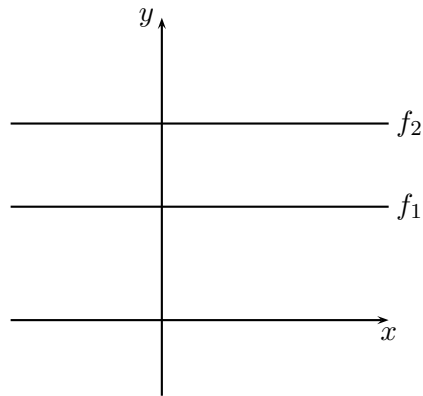


FIGURE 3.1 – Ordre pour deux réels

Par ailleurs, il semble naturel d'affirmer que la fonction $f_3 : x \mapsto 2 + x^2$ est “supérieure” à la fonction constante $f_4 : x \mapsto 1$, puisque le graphe de f_3 est partout situé au-dessus de celui de f_4 (voir la figure 3.2).

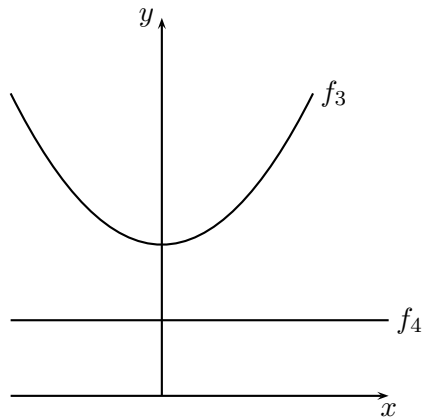


FIGURE 3.2 – Fonction supérieure à une autre

Mais comment comparer la fonction $f_5 : x \mapsto 1 + x^2$ à $f_6 : x \mapsto 2$, puisque la parabole d'équation $y = 1 + x^2$ est tantôt au-dessus, tantôt en dessous de l'horizontale d'ordonnée 2 (voir la figure 3.3) ?

L'objectif consistant à introduire des infinitésimaux, les points d'ordonnées élevées de la parabole n'ont guère d'intérêt dans ce propos. Ainsi, si l'on se restreint aux valeurs positives proches de 0 pour la variable x , la parabole

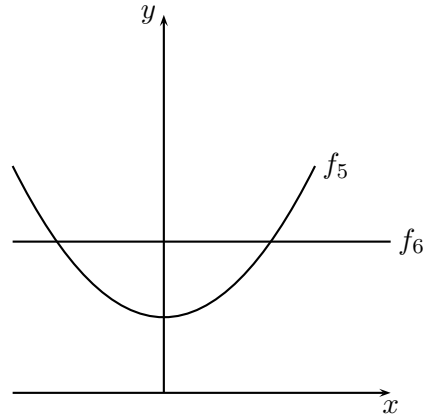


FIGURE 3.3 – Ordre local entre deux fonctions

est située en dessous de l’horizontale, de sorte qu’il semble “naturel”, dans ce contexte, de déclarer la fonction f_5 “plus petite” que f_6 .

Sur base de ces trois exemples (voir les trois figures 3.1, 3.2, 3.3), on introduit un ordre, noté $<_{\mathbb{Q}}$ sur $\mathbb{Q}[x]$. Pour deux fonctions rationnelles f et g , on définit

$$f <_{\mathbb{Q}} g \text{ si } (\exists k > 0 : f(x) < g(x) \forall x \in]0, k[) .$$

En guise d’exemple supplémentaire, considérons les fonctions $f : x \mapsto x^2$ et $g : x \mapsto x^3$. Pour $x < 1$, on a $g(x) < f(x)$, tandis que $g(x) \geq f(x)$ pour $x \geq 1$. En conséquence, si l’on choisit dans la définition ci-dessus $k = 1$, alors $g <_{\mathbb{Q}} f$ (voir la figure 3.4).

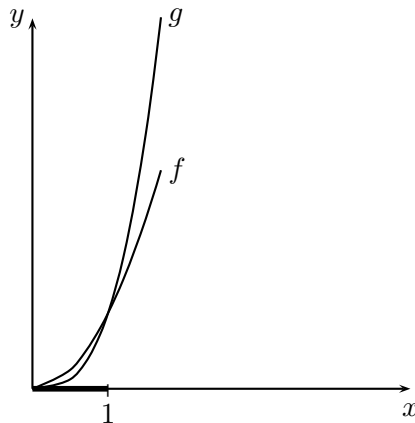


FIGURE 3.4 – Ordre local entre deux fonctions (bis)

Cet ordre répond-il aux besoins exprimés ci-dessus ? D’une part, $\mathbb{Q}[x]$ muni

de l'ordre $\leq_{\mathbb{Q}}$ est un corps ordonné ayant comme sous-corps le corps \mathbb{R} des réels. D'autre part, cet ordre garantit l'existence d'infinitésimaux. En effet, pour tout $n \in \mathbb{N}$, la fonction $f : x \mapsto x^n$, par exemple, est positive et inférieure (selon la définition de l'ordre dans $\mathbb{Q}[x]$) à toute fonction constante de graphe situé au-dessus de l'axe horizontal :

$$\forall r \in \mathbb{R}_0^+, (f_r : x \mapsto r) >_{\mathbb{Q}} (f : x \mapsto x^n) .$$

Il en va de même pour de nombreuses autres fonctions de $\mathbb{Q}[x]$.

3.2 Introduction des superréels

Ayant montré qu'il est possible de construire des infinitésimaux, il paraît intéressant de pouvoir exploiter ceux-ci en analyse. Comme ils font partie d'un corps ordonné, on peut déjà additionner, soustraire mais aussi multiplier et diviser des expressions contenant un nombre infinitésimal. Malheureusement, cela ne suffit pas car l'analyse traite aussi de fonctions non rationnelles. Par exemple, la formule suivante valable pour tout réel x :

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$$

devrait avoir un sens pour un ε infinitésimal. Avec ce développement à l'esprit, David TALL définit un nombre *superréel* comme une série formelle de puissances entières de ε dont les coefficients sont des nombres réels et dans laquelle ε est un infinitésimal non spécifié et l désigne un entier non négatif :

$$a_{-l} \frac{1}{\varepsilon^l} + \dots + a_{-1} \frac{1}{\varepsilon} + a_0 + a_1 \varepsilon + a_2 \varepsilon^2 + \dots$$

qui peut encore s'écrire sous la forme plus condensée (en posant $m = -l$) :

$$\sum_{k=m}^{\infty} a_k \varepsilon^k ,$$

où les coefficientst a_k désignent des réels arbitraires, tandis que m désigne un entier (négatif, positif ou nul) quelconque. De telles "séries" étant tout à fait formelles, il n'y a pas lieu de se préoccuper de leur éventuelle "convergence".

On désigne par \mathcal{R} l'ensemble de tous les superréels et par des lettres grecques α, β, \dots ses éléments. Il comprend en particulier le nombre superréel dont tous les coefficients a_j sont nuls et qui est naturellement noté simplement 0.

Pour tout superréel α non nul, il existe au moins un coefficient a_j non nul ; dans ce cas, nous conviendrons que l'indice m désigne le plus petit entier j tel que $a_j \neq 0$: cet entier m est encore appelé l'*ordre* de α et l'on note alors $o(\alpha) = m$. Remarquons dès à présent que tout nombre réel r peut se mettre

sous la forme d'un superréel α dont l'ordre vaut 0 et tel que $a_0 = r$ tandis que $a_j = 0$ pour tout entier j au moins égal à 1 ; dans la suite, il sera souvent noté simplement r . De la sorte, on obtient que \mathbb{R} est un sous-ensemble de \mathcal{R} .

Nous allons maintenant ériger \mathcal{R} en un corps commutatif en le munissant d'une addition et d'une multiplication, notées respectivement $+_{\mathcal{R}}$ et $\cdot_{\mathcal{R}}$, qui doivent étendre les opérations correspondantes entre nombres réels.

A cet effet, considérons deux superréels quelconques α et β définis par

$$\alpha = \sum_{j=m}^{\infty} a_j \varepsilon^j \quad , \quad \beta = \sum_{k=n}^{\infty} b_k \varepsilon^k$$

avec $m, n \in \mathbb{Z}$ et $a_j, b_k \in \mathbb{R}$. La somme de ces deux superréels se définit simplement en ajoutant entre elles, terme à terme, ces deux séries formelles selon les règles habituelles de l'algèbre ; de manière plus précise, on a

$$\alpha +_{\mathcal{R}} \beta = \sum_{i=\inf(m,n)}^{\infty} (a_i + b_i) \varepsilon^i \quad ,$$

si l'on pose $a_j = 0$ pour tout $j < m$, $b_k = 0$ pour tout $k < n$ et $\inf(m, n) = m$ (resp. n) si β (resp. α) coïncide avec 0. Le produit des deux superréels s'obtient en "multipliant de façon formelle", à l'aide de la distributivité, les deux séries, ce qui donne plus précisément

$$\alpha \cdot_{\mathcal{R}} \beta = \sum_{j=m}^{\infty} \sum_{k=n}^{\infty} a_j b_k \varepsilon^{j+k} \quad ,$$

et peut s'écrire sous la forme

$$\alpha \cdot_{\mathcal{R}} \beta = \sum_{i=m+n}^{\infty} c_i \varepsilon^i \quad .$$

Avec ces définitions, $\langle \mathcal{R}, +_{\mathcal{R}}, \cdot_{\mathcal{R}}, 0, 1 \rangle$ est un corps commutatif et possède \mathbb{R} comme sous-corps. On définit alors sur \mathcal{R} un ordre total compatible avec l'ordre habituel $<$ sur \mathbb{R} . A cet effet, on considère deux superréels α et β définis comme ci-dessus. On dira que $\alpha = \beta$ si $a_j = b_j$ pour tout $j \in \mathbb{Z}$; dès lors, si $\alpha \neq \beta$ il existe $q \in \mathbb{Z}$ tel que $a_q \neq b_q$; dans ce cas, on note p le plus petit indice tel que $a_p \neq b_p$ et on définit la relation d'ordre comme suit :

$$\alpha <_{\mathcal{R}} \beta \text{ si } a_p < b_p \quad .$$

Le champ \mathcal{R} muni de cette relation d'ordre est alors un champ ordonné au sein duquel ε est bien un nombre infiniment petit.

Bibliographie du chapitre 3

- [1] HENRY V., *Questions de didactique soulevées par un enseignement de l'analyse non standard à de futurs économistes*, thèse de doctorat en didactique des mathématiques, Université P. Sabatier, Toulouse, 2004.
- [2] TALL D., Looking at graphs through infinitesimal microscopes, windows and telescopes, *Mathematical Gazette*, 64, 1980, pp. 22 - 49.

Chapitre 4

Une vision positiviste de nombres hyperréels au moyen d'angles

4.1 Introduction

Partant des premiers développements de la géométrie classique, essentiellement des *Eléments* d'EUCLIDE ([3]), nous en arrivons à exploiter des travaux récents consacrés aux angles curvilinéaires [10]), ce qui nous fournit l'opportunité d'évoquer l'avènement de l'analyse non standard ([8]) pour laquelle des présentations simples sont désormais connues ([5], [7]). Plus précisément, nous montrons la possibilité de donner une vision positiviste¹ du concept de nombres hyperréels en concevant ceux-ci *comme le reflet d'“objets” du monde “naturel”*. *En ce sens, les nombres se doivent d'être “concrétisés” par l'un ou l'autre de ces objets* ([9], p. 61).

Cette étude est élaborée à partir d'angles plans élémentaires connus depuis l'Antiquité grecque. En effet, tout angle *corniculaire* ([1])², formé par un cercle et sa tangente en un point, peut se voir associer un “nombre” censé en donner une “mesure” ; un tel nombre est obligatoirement positif, mais inférieur à tout réel : il s'agit dès lors d'un infiniment petit. De la même manière, un angle *de demi-cercle*, formé par un cercle et son diamètre, permet d'illustrer le concept de nombre hyperréel infiniment grand. Enfin, certains angles *recti-circulaires*, dont un côté est une demi-droite tandis que l'autre est une portion de cercle et qui sont obtenus en assemblant adéquatement un angle rectiligne avec un angle corniculaire ou de demi-cercle, fournissent une représentation concrète de nombres non standards appréciables, c'est-à-dire infiniment proches d'un

1. Le positivisme est « une doctrine qui se réclame de la seule connaissance des faits, de l'expérience scientifique » (*Le Nouveau Petit Robert*, 1996, p. 1736).

2. D'autres appellations sont encore utilisées : par exemple, en français, *angle de contingence* et, en anglais, *Horn angle*.

réel non nul.

Ce travail se situe dans le plan, qui sera éventuellement rapporté à un repère orthogonal. Nous admettons connues les définitions de base données dans les *Eléments* d'EUCLIDE, à savoir les concepts de *point* ([3], Livre I. Définition 1, p. 151), de *ligne* ([3], Livre I. Définition 2, p. 152), de *(ligne) droite* ([3], Livre I. Définition 4, p. 154) et de *cercle* ([3], Livre I. Définition 15, p. 162). Au surplus, les raisonnements mathématiques exploités sont particulièrement élémentaires puisqu'ils figurent, pour l'essentiel, dans les *Eléments*.

Nous étudions uniquement des angles (plans) construits à partir de lignes, plus précisément à partir de demi-droites et de portions de cercles. Nous considérons des points $P, P' \dots$ du plan, des demi-droites $d, d' \dots$ d'extrémités $P, P' \dots$ et des portions de cercles $\mathcal{C}, \mathcal{C}' \dots$ issues également des points $P, P' \dots$.

4.2 Généralités sur les angles rectilignes et recti-circulaires

Vu la multiplicité des définitions liées à la notion d'angle ([2], p. 71) et bien que le concept d'angle soit défini dans les *Eléments*, nous précisons dans quel sens nous l'envisageons, de manière à éviter des objections³ semblables à celles soulevées, à propos de la définition initiale ([3], Livre I. Définition 8, p. 158), dans le *Commentaire*.

Définition 4.2.1 *Un couple composé d'une demi-droite d et d'une portion de ligne \mathcal{L} , toutes deux d'extrémité P , est par définition un angle de sommet P et de côtés d et \mathcal{L} ; dans un premier temps, cet angle sera noté $(d, \mathcal{L})_P$.*

Selon que la portion de ligne \mathcal{L} est une demi-droite ou une portion de cercle, l'angle est qualifié de *rectiligne* ou de *recti-circulaire*.

Définition 4.2.2 *Deux angles $(d, \mathcal{L})_P$ et $(d', \mathcal{L}')_{P'}$ sont dits égaux, ce qui se note naturellement $(d, \mathcal{L})_P = (d', \mathcal{L}')_{P'}$, lorsqu'il existe une isométrie du plan qui associe P' à P , d' à d et \mathcal{L}' à \mathcal{L} .*

Cette définition nous autorise à ne considérer que des angles dont le sommet est l'origine O du plan tandis que le côté rectiligne d est l'axe horizontal du

3. Comme le mentionne PROCLUS, « l'angle corniculaire fournit un argument à ceux qui contestent le caractère "grandeur" de l'angle [...] la considération d'angles plans mixtilignes est donc incompatible avec le caractère archimédien des grandeurs appartenant à un même genre. » ([3], p. 426)

plan h ; nous noterons alors l'angle simplement (h, \mathcal{L}) . De plus, en travaillant dans un repère orthogonal, nous avons

$$h = \{(x, y) : x \geq 0, y = 0\}.$$

En nous restreignant à des angles (h, \mathcal{L}) pour lesquels le côté \mathcal{L} est, à l'exception de l'origine, situé dans le quadrant $\{(x, y) : x > 0, y \geq 0\}$, nous définissons un ordre total :

Définition 4.2.3 *Un angle (h, \mathcal{L}) est dit inférieur à un angle (h, \mathcal{L}') , ce qui est évidemment noté $(h, \mathcal{L}) < (h, \mathcal{L}')$, lorsqu'il existe un voisinage de l'origine au sein duquel, et à l'exclusion de l'origine, le côté \mathcal{L} est situé en-dessous du côté \mathcal{L}' .*

Cette définition redonne l'ordre classique des angles rectilignes : effectivement, pour deux réels a et a' non négatifs, et deux angles rectilignes $\alpha = (h, \mathcal{L})$, $\alpha' = (h, \mathcal{L}')$, avec $\mathcal{L} = \{(x, y) : x \geq 0, y = ax\}$, $\mathcal{L}' = \{(x, y) : x \geq 0, y = a'x\}$, on voit aisément que $\alpha < \alpha'$ si et seulement si $a < a'$, ce qui équivaut à dire que l'amplitude de α est inférieure à celle de α' . Au surplus, comme nous allons le voir, elle permet aussi de comparer deux angles qui pourraient être recti-circulaires.

4.3 Infiniment petits et angles corniculaires

Considérons un angle corniculaire (h, \mathcal{C}) , désigné par la lettre grecque ε , pour lequel

$$\mathcal{C} = \{(x, y) : x \geq 0, x^2 + (y - r)^2 = r^2\} \text{ avec } r > 0$$

Il s'agit donc de l'angle formé par le demi-cercle, situé dans le premier quadrant, de rayon r et de centre le point d'abscisse nulle et d'ordonnée égale à r , et par sa tangente en l'origine.

En désignant par ν l'angle nul (h, h) , une application de la dernière définition livre $\nu < \varepsilon$.

Par ailleurs, quel que soit l'angle rectiligne aigu $\alpha = (h, d)$, avec $d = \{(x, y) : x \geq 0, y = ax\}$ et où a désigne un réel positif arbitraire, la dernière définition permet d'écrire $\varepsilon < \alpha$, puisque EUCLIDE a démontré qu'« entre le lieu compris entre la droite [la tangente] et la circonférence, une autre droite ne sera pas intercalée » ([3], Livre III, Proposition 16, p. 424).

En conséquence, sachant que les angles rectilignes peuvent être “mesurés” par un nombre réel nul pour l'angle nul ou positif (égal au coefficient directeur du côté oblique) pour un angle aigu, nous pouvons, par analogie, associer à l'angle corniculaire ε un “nombre” caractérisant en quelque sorte sa “mesure”.

Bien entendu, en accord avec l'ordre établi ci-dessus, cette "mesure" de ε doit être positive, mais plus petite que la mesure de tout angle rectiligne aigu : en d'autres termes, cette "mesure" est un nombre non réel et dès lors non standard, noté ${}^*p^4$, tel que $0 < {}^*p < a$ pour tout réel positif a ; il est dès lors permis d'imaginer que *p est un hyperréel [5], qui est, par définition, un *infinitement petit*.

En modifiant à volonté le rayon r du côté circulaire de l'angle corniculaire, il est facile de construire autant d'autres angles corniculaires dont les "mesures" sont toujours des nombres hyperréels infinitement petits ; ces derniers peuvent être ordonnés toujours en accord avec la troisième définition. De fait, soient deux réels positifs r et r' tels que $r < r'$, les angles corniculaires ε et ε' définis comme ci-dessus au moyen des cercles tangents à l'axe horizontal h en l'origine, de rayons respectifs r et r' , les "mesures" hyperréelles correspondantes étant égales à *p et ${}^*p'$: clairement, ${}^*p < {}^*p'$ si et seulement si $r > r'$, condition équivalente à $\frac{1}{r} < \frac{1}{r'}$; cela revient à dire qu'un angle corniculaire - donc le nombre hyperréel infinitement petit qui lui est associé - est plus petit qu'un autre lorsque la courbure du côté circulaire du premier est inférieure à celle du côté circulaire du second.

4.4 Infinitement grands et angles de demi-cercles

Envisageons à présent un angle de demi-cercle (h, \mathcal{C}) , désigné par la lettre grecque ω , pour lequel

$$\mathcal{C} = \left\{ (x, y) : x \geq 0, (x - r)^2 + y^2 = r^2 \right\} \text{ avec } r > 0$$

Il s'agit donc de l'angle formé par le demi-cercle, situé dans le premier quadrant, de rayon r et de centre le point d'abscisse égale à r et d'ordonnée nulle, et par son diamètre.

Reprenons les notations introduites ci-dessus. Quel que soit l'angle rectiligne aigu α défini par un côté oblique de coefficient directeur a , la définition 3 permet d'écrire $\alpha < \omega$, puisque EUCLIDE a également démontré que *l'angle du demi-cercle est plus grand [...] que tout angle rectiligne aigu* ([3], Livre III, Proposition 16, p. 424).

Dès lors, une "mesure" de l'angle ω est un nombre, forcément non standard et noté *g , qui doit être plus grand que tout réel positif : il s'agit d'un nombre hyperréel qui est, par définition, un *infinitement grand*.

4. Enseignant l'analyse non standard depuis plusieurs années, nous avons constaté qu'une difficulté potentielle pour les étudiants consiste à bien localiser le cadre dans lequel est effectué le raisonnement, à savoir l'ensemble \mathbb{R} des nombres réels, ou celui ${}^*\mathbb{R}$ des hyperréels. Afin d'atténuer cet obstacle, nous avons pris l'habitude de noter les réels par des lettres latines a, \dots et les hyperréels (non réels) par des lettres latines précédées par un astérisque ${}^*a, \dots$

Comme dans la section précédente, il est possible de construire différents nombres infiniment grands comparables en changeant le rayon r du côté circulaire de l'angle de demi-cercle. En effet, pour deux angles de demi-cercle ω et ω' définis comme ci-dessus au moyen des cercles de rayons respectifs r et r' avec $r < r'$, leurs “mesures” hyperréelles respectives $*g$ et $*g'$ sont visiblement telles que $*g < *g'$ si et seulement si $\omega < \omega'$, c'est-à-dire, conformément à la troisième définition, si et seulement si $r < r'$. Géométriquement, une telle relation résulte de ce que tout angle de demi-cercle peut être regardé comme le complément d'un angle corniculaire par rapport à un angle droit, ce qui se traduit algébriquement par le fait que l'inverse d'un nombre hyperréel infiniment grand est un infiniment petit.

4.5 Nombres non standards appréciables et angles recti-circulaires

Dorénavant, fixons arbitrairement deux réels positifs a et r , puis considérons les deux cercles tangents, en l'origine, à la demi-droite déjà considérée

$$d = \{(x, y) : x \geq 0, y = ax\} ,$$

à savoir les cercles, de rayon r et respectivement d'équation

$$\begin{aligned} \left(x - \frac{ra}{\sqrt{1+a^2}}\right)^2 + \left(y + \frac{r}{\sqrt{1+a^2}}\right)^2 &= r^2 \\ \left(x + \frac{ra}{\sqrt{1+a^2}}\right)^2 + \left(y - \frac{r}{\sqrt{1+a^2}}\right)^2 &= r^2 \end{aligned}$$

Les intersections de ces deux cercles avec le quadrant non négatif déterminent deux courbes, notées respectivement \mathcal{C} et \mathcal{C}' , qui donnent lieu aux deux angles recti-circulaires $\gamma = (h, \mathcal{C})$ et $\gamma' = (h, \mathcal{C}')$. Par rapport à l'angle rectiligne $\alpha = (h, d)$, la troisième définition livre $\gamma < \alpha < \gamma'$. Comme précédemment, les “mesures” des angles γ et γ' sont des nombres non standards, notés respectivement $*a$ et $*a'$, tels que $*a < a < *a'$. De plus, pour tout réel positif b inférieur à a , il est facile de voir que $b < *a$; de même, pour tout réel b' supérieur à a , on a $*a' < b'$. En conséquence, $*a$ est un nombre non standard inférieur à a mais qui lui est infiniment proche, tandis que $*a'$ est un nombre non standard supérieur et infiniment proche de a : $*a$ et $*a'$ sont donc des hyperréels appréciables. Ainsi, d'un point de vue géométrique, les angles γ et γ' peuvent être construits en assemblant adéquatement un angle corniculaire avec un angle rectiligne : de fait, γ (resp. γ') s'obtient en “diminuant” (resp. en “augmentant”) l'angle rectiligne α de l'angle corniculaire mené à partir de la tangente d ; cette propriété géométrique se traduit algébriquement : le nombre hyperréel $*a$ (resp. $*a'$) est la différence (resp. la somme) du réel a et d'un infiniment petit positif.

En choisissant un rayon différent de r , d'autres nombres hyperréels appréciables peuvent évidemment être obtenus. Il en va de même lorsque a varie,

c'est-à-dire lorsque la demi-droite d subit une rotation autour de l'origine ; dans ce contexte, on constate que, pour r fixé, les angles corniculaire ε et de demi-cercle ω peuvent être perçus comme étant des positions-limites des angles γ' et γ lorsque, respectivement, a tend vers ∞ ou 0.

4.6 Conclusion

Les considérations précédentes ne remplacent évidemment pas une preuve constructiviste (donnée par ROBINSON [8]) de l'existence d'hyperréels non réels, avec les trois ordres de grandeur que sont les infiniment petits, les infiniment grands et les appréciables. Elles visent uniquement à montrer, au moyen de raisonnements élémentaires, que l'introduction de tels nombres peut paraître "naturelle", car conforme à l'intuition géométrique ainsi qu'au souci d'associer une mesure à une figure géométrique, cette dernière préoccupation étant fort prégnante ainsi qu'en attestent les liens étroits qui ont toujours existé entre la théorie des nombres et la géométrie ([4]).

Bien entendu, les raisonnements effectués à partir de cercles pourraient être adaptés avec d'autres courbes, par exemple des courbes d'équation $y = x^a$, idée qui se trouve déjà sous-jacente dans l'œuvre les *Principia* de NEWTON.

Bibliographie du chapitre 4

- [1] BAIR J. - HENRY V., Les angles corniculaires et l'infiniment petit, *Tangente Sup*, n° 31, 2006, pp. 4-7.
- [2] CHEVALIER J.M. - DAVID-CHEVALIER M.C., OMNI : objet mathématique non identifié. Un outil pédagogique au service de l'apprentissage de la notion d'angle et de sa mesure, *Repères - IREM*, **63**, 2006, pp. 69-93.
- [3] EUCLIDE d'Alexandrie, *Les Eléments*, volume 1, Livres I à IV, Presses Universitaires de France, Paris, 1990.
- [4] FRIEDELMEYER J. P., De l'angle de contingence au rayon de courbure : comment penser, comparer, mesurer le courbe, *La figure et l'espace*, IREM de Lyon Villeurbanne, 1993, pp. 365-396.
- [5] HENRY V. , Les hyperréels en analyse, *Mathématique et Pédagogie*, **141**, 2003, pp. 47-58.
- [6] HENRY V., *Questions de didactique soulevées par un enseignement de l'analyse non standard à de futurs économistes*, Thèse doctorale présentée devant l'université Paul Sabatier (Toulouse III) en vue de l'obtention du doctorat en didactique des mathématiques, Toulouse, 2004.
- [7] KEISLER H.J., *Elementary Calculus*, Prindle, Weber & Schmidt, Boston, 1976.

- [8] ROBINSON A., *Non-Standard Analysis*, Revisited edition, North-Holland Pub. Comp., Amsterdam-London, 1974.
- [9] ROSSEEL H. - SCHNEIDER M., Ces nombres que l'on dit "imaginaires" I, *Petit x*, **63**, 2003, pp. 53-71.
- [10] VAN ASCH A.G. - VAN DER BLIJ F., Measurements of curvilinear angles, *Bull. Belgian Math. Soc. - Simon Stevin*, **2**, n° 5, 1995, pp. 573-588.

Chapitre 5

Nombres superréels et angles corniculaires

Résumé. En nous appuyant sur des travaux de VAN ASCH et VAN DER BLIJ, nous montrons que les angles corniculaires, qui ont été introduits par EUCLIDE, peuvent se voir attribuer une mesure numérique donnée par un nombre superréel infiniment petit au sens de TALL. Nous en déduisons la possibilité d’estimer d’une part le rapport entre les mesures d’un angle corniculaire et d’un angle mixtiligne formé par une spirale d’ARCHIMÈDE, et d’autre part le rapport des mesures de deux angles corniculaires.

5.1 Introduction

La notion d’angle, fondamentale en mathématiques, est délicate et a fait l’objet de multiples études, parfois contradictoires. Ainsi en est-il de la théorie des *angles corniculaires*, dont l’histoire peut être schématiquement décrite en trois étapes conformément à la dialectique hégélienne.

1. **Thèse.** Dans le Premier Livre de ses *Éléments* [2], EUCLIDE (III^e siècle av. J.-C.) définit la notion d’angle plan, en ne se limitant pas seulement aux angles rectilignes : *un angle plan est l’inclinaison, l’une sur l’autre, dans un plan, de deux lignes qui se touchent l’une l’autre . . . et quand les lignes contenant l’angle sont droites, l’angle est appelé rectiligne* ([2], p. 158). Il y considère des angles *mixtilignes*, notamment l’*“angle corniculaire”* - *inclinaison mutuelle de la circonférence d’un cercle et de sa tangente en un point* ([2], p. 158) ; de fait, dans le Troisième Livre de ses *Éléments*, il démontre que *la droite menée à angles droits avec le diamètre du cercle à partir d’une extrémité tombera à l’extérieur du cercle, et dans le lieu compris entre la droite et la circonférence, une autre droite ne sera pas intercalée ; en outre, l’angle restant [c’est-à-dire l’angle corniculaire ainsi construit] est plus petit que tout angle rectiligne*

aigu ([2], pp. 423-424).

2. **Antithèse.** Dans le *Commentaire* au Premier Livre des *Éléments*, PROCLUS (412 - 485) s'interroge sur la nature des angles et met notamment en évidence le paradoxe suivant : d'une part, les angles apparaissent comme étant des quantités qui peuvent être comparées et divisées, mais, d'autre part, EUCLIDE a démontré qu'un angle corniculaire est inférieur à tout angle rectiligne aigu, alors qu'il devrait donc être une quantité non nulle. Après bien des débats, J. WALLIS (1616 - 1703) affirmera qu'un angle corniculaire n'est pas un angle : *Je dis que cette déflexion (qui éloigne la courbe de sa tangente ...) n'est pas un angle ou une déclinaison, (pas plus que dans un mouvement une accélération n'est une vitesse). Mais c'est le commencement d'une déclinaison, qui montre le degré de courbure* ([7], p. 655, cité par [4], p. 162).
3. **Synthèse.** En sortant du cadre numérique pour travailler dans le cadre fonctionnel, les mathématiciens hollandais VAN ASCH et VAN DER BLIJ parviennent, en 1995, à définir une « mesure non triviale pour l'angle situé entre deux courbes tangentes » ([6], p. 573).

Cela permet de lever l'objection de PROCLUS rappelée ci-dessus.

5.2 Estimation fonctionnelle des angles corniculaires ou mixtilignes

Dans un premier temps, appliquons les résultats des deux hollandais aux angles corniculaires.

Sans nuire à la généralité, nous allons étudier un angle corniculaire γ formé par l'axe horizontal H des abscisses et la portion C située dans le premier quadrant d'un cercle qui vient toucher tangentiellement H en l'origine du plan.

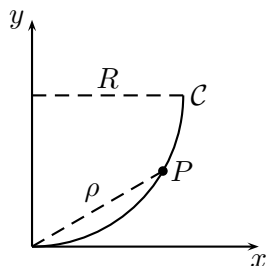


FIGURE 5.1 – Angle corniculaire

Désignons par R le rayon du cercle considéré, de sorte que C est d'équation cartésienne :

$$y = R - \sqrt{R^2 - x^2} .$$

En remplaçant les coordonnées cartésiennes x et y par leurs homologues en coordonnées polaires, on peut encore écrire

$$(\rho \cos \theta)^2 + (\rho \sin \theta - R)^2 = R^2 .$$

Après quelques calculs algébriques et simplifications, on peut encore écrire

$$\sin \theta = \frac{\rho}{2R} .$$

Ainsi, l'angle polaire θ d'un point P de \mathcal{C} , situé à une distance ρ de l'origine, est donné par

$$\theta = \arcsin \left(\frac{\rho}{2R} \right) .$$

D'autre part, la valeur de l'angle polaire θ peut également être mise sous la forme suivante

$$\theta = \frac{l(\rho)}{\rho} ,$$

où $l(\rho)$ désigne la longueur de l'arc, intercepté par l'angle au centre θ , du cercle centré à l'origine et passant par le point considéré.

Comme l'angle polaire θ d'un point courant de \mathcal{C} varie avec la distance ρ de ce point à l'origine, l'angle corniculaire γ peut être en quelque sorte "estimé" par sa *fonction d'estimation* f_γ définie comme suit :

$$f_\gamma : \rho \mapsto \theta = \arcsin \left(\frac{\rho}{2R} \right) = \frac{l(\rho)}{\rho} = f_\gamma(\rho) .$$

On a visiblement

$$\lim_{\rho \rightarrow 0} f_\gamma(\rho) = 0 \text{ et } \lim_{\rho \rightarrow 0} \frac{f_\gamma(\rho)}{\rho} = \frac{1}{2R} .$$

En tenant compte de ces deux limites, on peut assimiler la fonction f_γ à un infiniment petit d'ordre 1 relativement à la variable ρ tendant vers 0.

A la suite de VAN ASCH et VAN DER BLIJ [6], nous pouvons généraliser ces résultats à une large classe d'angles mixtilignes. De fait, toute fonction analytique F au voisinage de 0 détermine, dans le plan, un angle mixtiligne dont un des côtés est l'axe horizontal des abscisses, tandis que le second côté est la courbe dont l'équation en coordonnées polaires est donnée par

$$\theta = F(\rho) .$$

En reprenant le raisonnement réalisé pour un angle corniculaire et avec les mêmes notations, on peut "évaluer" un tel angle mixtiligne α par sa *fonction d'estimation* F_α définie comme suit :

$$F_\alpha : \rho \mapsto \theta = \frac{l(\rho)}{\rho} = F_\alpha(\rho) .$$

Parmi les angles mixtilignes considérés figurent notamment, outre les angles corniculaires, les angles rectilignes dont la fonction d'estimation est une fonction constante, ou encore les angles "spirales" dont le côté curviligne est une spirale et dont la fonction d'estimation est une fonction linéaire.

En adaptant une idée de D. TALL [5], on peut instaurer un ordre, noté $<_T$, relatif à ces angles mixtilignes : pour deux tels angles α et β , dont la fonction d'estimation est respectivement F_α et F_β , on a la définition suivante

$$\alpha <_T \beta \Leftrightarrow \exists k > 0, \quad F_\alpha(\rho) < F_\beta(\rho) \text{ pour tout } \rho \in]0, k[.$$

En particulier, pour deux angles corniculaires γ_1 et γ_2 , définis par des cercles de rayon R_1 et R_2 respectivement, on a visiblement

$$\gamma_1 <_T \gamma_2 \Leftrightarrow R_1 > R_2.$$

Par ailleurs, il est clair que tout angle corniculaire γ est un "infinitement petit positif" ; en effet, pour tout angle rectiligne δ possédant une fonction d'estimation positive et pour l'angle nul ν dont les deux côtés rectilignes sont superposés, on dispose de ces inégalités :

$$\nu <_T \gamma <_T \delta.$$

L'objectif qui va être poursuivi dans la suite de ce chapitre consiste à sortir du cadre fonctionnel et à définir une mesure numérique d'un angle corniculaire. D'après ce qui précède, une telle mesure devrait être un "nombre" positif, mais inférieur à tout nombre réel positif. Il ne peut donc pas s'agir d'un nombre réel, mais bien d'un "infinitement petit" au sens considéré déjà par LEIBNIZ puis par ses successeurs.

Nous nous proposons d'associer à chaque angle corniculaire, et plus généralement à chaque angle mixtiligne considéré ci-dessus, une mesure qui sera un nombre *superréel* ([5]). Mais, auparavant, nous allons présenter ces nouveaux nombres d'une manière nouvelle.

5.3 Introduction des nombres superréels

L'idée poursuivie consiste à construire une extension de \mathbb{R} dans laquelle les nouveaux nombres sont introduits au moyen des mêmes règles du calcul algébrique que les réels et qui contienne un nombre, forcément non réel, qui soit infinitement petit, c'est-à-dire non nul mais inférieur à tout réel positif. A cet effet, nous allons procéder par analogie avec la construction hamiltonienne des nombres complexes sous forme de couples de réels soumis à des règles algébriques bien définies ([3]).

De manière à bien percevoir cette analogie, rappelons succinctement comment peut être présenté l'ensemble \mathbb{C} des nombres complexes.

Rappel sur les nombres complexes

Dans le corps \mathbb{R} des réels, l'équation $x^2 + 1 = 0$ ne possède pas de solution. Pour remédier à pareille lacune, on construit une extension de \mathbb{R} , que l'on appelle l'*ensemble des complexes*,

$$\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{(a, b) : a, b \in \mathbb{R}\},$$

sur laquelle on définit deux opérations algébriques comme suit :

$$+_C : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C} : ((a, b), (c, d)) \mapsto (a + c, b + d)$$

$$\cdot_C : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C} : ((a, b), (c, d)) \mapsto (ac - bd, ad + bc).$$

Soit $a \in \mathbb{R}$; on identifie a avec $(a, 0)$ et on a ainsi $\mathbb{R} \subset \mathbb{C}$. De plus, en posant $i = (0, 1)$, on a $i^2 + 1 = 0$ et tout nombre complexe $z = (a, b)$ peut s'écrire sous la forme $za + ib$.

Construction des superréels analogue à celle des complexes

Dans le corps \mathbb{R} des réels, il est impossible de trouver un nombre infiniment petit positif, c'est-à-dire un nombre x tel que $0 < x < r$ pour tout réel r positif; en effet, si on suppose qu'un tel nombre x existe et est réel, alors $\frac{x}{2}$ appartient aussi à \mathbb{R} et on a donc $x < \frac{x}{2}$, ce qui est absurde.

Pour disposer d'un nombre infiniment petit, on est donc conduit à construire un ensemble "plus grand" que \mathbb{R} , ce qui peut être réalisé en définissant l'*ensemble des superréels*

$$\mathcal{R} = \left\{ (a_j)_{j \in \mathbb{Z}} : a_j \in \mathbb{R}, \exists m \in \mathbb{Z} : a_j = 0, \forall j < m \right\}$$

sur lequel sont introduites deux opérations :

$$+_R : \mathcal{R} \times \mathcal{R} \rightarrow \mathcal{R} : \left((a_j)_{j \in \mathbb{Z}}, (b_j)_{j \in \mathbb{Z}} \right) \mapsto (a_j + b_j)_{j \in \mathbb{Z}}$$

$$\cdot_R : \mathcal{R} \times \mathcal{R} \rightarrow \mathcal{R} : \left((a_j)_{j \in \mathbb{Z}}, (b_j)_{j \in \mathbb{Z}} \right) \mapsto \left(\sum_{k=i+j} a_i b_j \right)_{i \in \mathbb{Z}}$$

Moyennant l'introduction de ces deux opérations, on vérifie aisément que la structure $(\mathcal{R}, +_R, 0, \cdot_R, 1)$ est un corps commutatif si on pose

$$0 = (a_j)_{j \in \mathbb{Z}} \text{ tel que } a_j = 0 \forall j \in \mathbb{Z}$$

et

$$1 = (a_j)_{j \in \mathbb{Z}} \text{ tel que } a_j = 0 \forall j \neq 0, a_0 = 1.$$

Tout réel r est un élément de \mathcal{R} si on écrit $r = (a_j)_{j \in \mathbb{Z}}$ tel que $a_j = 0$ pour tout indice j non nul, avec $a_0 = r$; on a donc $\mathbb{R} \subset \mathcal{R}$.

Ordre sur \mathcal{R} et existence d'un infiniment petit

Définition 5.3.1 Ordre d'un élément $\alpha \neq 0$ de \mathcal{R}

Soit $\alpha \in \mathcal{R}$, $\alpha = (a_j)_{j \in \mathbb{Z}}$, il existe $m \in \mathbb{Z}$ tel que $a_j = 0$ pour tout $j < m$ et $a_m \neq 0$. On appelle m l'ordre de α , ce que l'on note $m = o(\alpha)$.

Définition 5.3.2 Pour $\alpha = (a_j)_{j \in \mathbb{Z}}$ et $\beta = (b_j)_{j \in \mathbb{Z}}$ avec $\alpha \neq \beta$, soit p le plus petit entier tel que $a_p \neq b_p$, la relation d'ordre sur \mathcal{R} est définie comme suit :

$$\alpha <_{\mathcal{R}} \beta \text{ si } a_p < b_p .$$

Particularisons à présent un élément de \mathcal{R} , à savoir

$$\varepsilon = (a_j)_{j \in \mathbb{Z}} : a_1 = 1 \text{ et } a_j = 0, \forall j \neq 1 .$$

Il est clair que ε n'appartient pas à \mathbb{R} . De plus, pour un réel positif r arbitraire, on a visiblement au vu des définitions précédentes

$$0 <_{\mathcal{R}} \varepsilon \text{ et } \varepsilon <_{\mathcal{R}} r$$

Ce superréel ε est donc ce qui est appelé un infiniment petit positif.

Superréels et développement en série

A l'aide du superréel particulier ε , nous allons pouvoir proposer une autre écriture pour les éléments de \mathcal{R} : tout élément α de \mathcal{R} peut s'écrire sous la forme :

$$\alpha = (a_j)_{j \in \mathbb{Z}} = \sum_{k=-\infty}^{+\infty} a_k \varepsilon^k .$$

Les définitions du produit et de la somme dans \mathcal{R} se justifient aisément par les propriétés de la somme et du produit dans \mathbb{R} puisque, dans le cas du produit par exemple, on a

$$\begin{aligned} \alpha \cdot_{\mathcal{R}} \beta &= \left(\sum_{j=-\infty}^{+\infty} a_j \varepsilon^j \right) \cdot \left(\sum_{k=-\infty}^{+\infty} b_k \varepsilon^k \right) \\ &= \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} a_j b_k \varepsilon^{j+k} \\ &= \sum_{i=-\infty}^{+\infty} \left(\sum_{j+k=i} a_j b_k \right) \varepsilon^i \end{aligned}$$

On définit ensuite des ordres de grandeur au sein de l'ensemble \mathcal{R} .

Définition 5.3.3 Un élément α non nul de \mathcal{R} est

- appréciable si $o(\alpha) = 0$;
- infiniment petit si $o(\alpha) > 0$;
- infiniment grand si $o(\alpha) < 0$;
- limité si $o(\alpha) \geq 0$.

Définition 5.3.4 Soit $\alpha \in \mathcal{R}$ limité, la partie standard de α , notée $\text{st}(\alpha)$, est le terme indépendant dans le développement de α :

$$\alpha = \sum_{j=m}^{+\infty} a_j \varepsilon^j \implies \text{st}(\alpha) = a_0 .$$

De cette définition découlent immédiatement ces propriétés de la partie standard :

Proposition 5.3.5 - La partie standard de tout superréel limité est unique.

- Soient α et β des limités de \mathcal{R}
 - a) $\text{st}(\alpha +_{\mathcal{R}} \beta) = \text{st}(\alpha) + \text{st}(\beta)$;
 - b) $\text{st}(\alpha \cdot_{\mathcal{R}} \beta) = \text{st}(\alpha) \cdot \text{st}(\beta)$;
 - c) $\alpha <_{\mathcal{R}} \beta \implies \text{st}(\alpha) \leq \text{st}(\beta)$;
 - d) $\text{st}(\alpha) < \text{st}(\beta) \implies \alpha <_{\mathcal{R}} \beta$;

5.4 Mesure numérique d'un angle corniculaire

Pour "mesurer" un angle corniculaire γ déterminé par un cercle de rayon R , construisons le développement de MAC LAURIN de la fonction d'estimation f_γ introduite ci-dessus. On peut écrire, pour ρ suffisamment petit

$$\begin{aligned} \theta &= \frac{\rho}{2R} + \frac{1}{2 \times 3} \left(\frac{\rho}{2R}\right)^3 + \frac{1 \times 3}{2 \times 4 \times 5} \left(\frac{\rho}{2R}\right)^5 + \dots \\ &+ \frac{1 \times 3 \times \dots \times (2p-1)}{2 \times 4 \times \dots \times (2p) \times (2p+1)} \left(\frac{\rho}{2R}\right)^{2p+1} + o(\rho) \rho^{2p+2} \end{aligned}$$

Lorsque ρ tend vers 0, cette égalité est d'application : on constate alors que les coefficients de la puissance de la variable ρ restent invariants (pour un même cercle de rayon R). En d'autres termes, l'angle corniculaire γ considéré est parfaitement déterminé par la suite des coefficients c_j des puissances de ρ , à savoir pour tout entier positif n ,

$$c_{2n-1} = \frac{(2(n-1))!}{2^{4n-3} ((n-1)!)^2 (2n-1)R^{2n-1}} \text{ et } c_{2n} = 0.$$

On peut dès lors associer à cet angle corniculaire γ un nombre superréel d'ordre 1, à savoir

$$\mu_\gamma = \left(\frac{1}{2R}, 0, \frac{3!}{(2 \times 3)^2} \frac{1}{(2R)^3}, 0, \frac{5!}{(4 \times 5 \times 2)^2} \frac{1}{(2R)^5}, 0, \dots \right) :$$

ce nombre superréel est défini comme étant une *mesure* de l'angle corniculaire considéré. A la suite de TALL ([5]), on peut encore écrire ce nombre superréel μ_γ sous la forme d'une série formelle, à savoir, avec les nombres réels c_j introduits plus haut

$$\mu_\gamma = \sum_{j=1}^{+\infty} c_j \varepsilon^j,$$

où ε désigne un symbole "non spécifié" ([5], p. 27).

Plus généralement, pour un angle mixtiligne α dont la fonction d'estimation F_α admet comme développement de MAC LAURIN

$$F_\alpha(\rho) = \sum_{j=0}^{+\infty} a_j \rho^j,$$

on définit la mesure μ_α comme étant le nombre superréel écrit de façon formelle sous la forme suivante :

$$\mu_\alpha = \sum_{j=0}^{+\infty} a_j \varepsilon^j.$$

Dans ce cas, la partie standard du superréel fini ([5], p. 28) μ_α représente la mesure classique de l'angle obtenu à partir de l'angle mixtiligne lorsque le côté curviligne est remplacé par la tangente menée à la courbe depuis le sommet de l'angle.

En comparant la présentation formelle d'un angle mixtiligne avec le développement de MAC LAURIN de sa fonction d'estimation, le superréel ε peut être assimilé à une variable ρ tendant vers zéro, ce qui correspond, lorsque le côté curviligne est tangent au côté rectiligne, à la conception d'un infiniment petit donnée par CAUCHY ([1], p. 3).

Une interprétation géométrique de ε se réfère à la mesure de l'angle mixtiligne formé, dans le plan, par l'axe horizontal des abscisses et la spirale d'ARCHIMÈDE, puisque cette dernière est d'équation polaire :

$$\theta = \rho.$$

La structure algébrique définie sur l'ensemble des superréels \mathcal{R} peut être appliquée aux mesures ainsi introduites pour les angles corniculaires. Nous allons montrer que l'ordre $<_{\mathcal{R}}$ et l'addition $+_{\mathcal{R}}$ sont conformes à l'intuition géométrique, la multiplication $\times_{\mathcal{R}}$ ne semblant pas avoir d'interprétation évidente dans ce contexte.

Ainsi, un angle corniculaire γ_1 défini par un cercle de rayon R_1 avait été déclaré ci-dessus "plus grand" qu'un angle corniculaire γ_2 correspondant à un rayon R_2 , ce que nous avons noté $\gamma_2 <_T \gamma_1$, si et seulement si $R_1 < R_2$; cette condition équivaut bien à l'inégalité correspondante sur leurs mesures superréelles; en effet,

$$\gamma_2 <_T \gamma_1 \Leftrightarrow \mu_{\gamma_2} < \mu_{\gamma_1}.$$

Par ailleurs, pour additionner deux angles corniculaires γ_1 et γ_2 , il suffit, intuitivement, d'en mettre un au-dessus de l'axe des abscisses et l'autre en-dessous, puis de définir l'angle formé par ces deux portions de cercles ; ce n'est plus un angle corniculaire ([6]), mais bien un angle curviligne dont les deux côtés sont des portions de cercle et pour lequel on dispose de cette égalité :

$$\frac{l(\rho)}{\rho} = \frac{l_1(\rho)}{\rho} + \frac{l_2(\rho)}{\rho}.$$

Algébriquement, la mesure de la somme des deux angles s'obtient simplement en additionnant terme à terme au sens de TALL les deux séries définissant γ_1 et γ_2 .

5.5 Rapport des mesures de deux angles corniculaires

Soit un angle corniculaire γ engendré par un cercle de rayon R . Sa mesure μ_γ vaut le nombre superréel infiniment petit écrit de façon formelle

$$\mu_\gamma = \sum_{j=1}^{+\infty} c_j \varepsilon^j.$$

Or, le superréel ε peut être adopté comme mesure de l'angle mixtiligne σ défini par l'axe horizontal et la spirale d'ARCHIMÈDE. En notant μ_σ la mesure de σ , on peut estimer le rapport de μ_γ et de μ_σ : il s'agit du quotient de deux superréels infiniment petits, mais ce quotient est un superréel appréciable, égal à

$$\frac{\mu_\gamma}{\mu_\sigma} = \frac{\sum_{j=1}^{+\infty} c_j \varepsilon^j}{\varepsilon} = \frac{1}{2R} + \sum_{j=2}^{+\infty} c_j \varepsilon^{j-1} ;$$

la partie standard de ce rapport vaut donc le nombre réel

$$\text{st} \left(\frac{\mu_\gamma}{\mu_\sigma} \right) = \frac{1}{2R}.$$

Dans le même ordre d'idées, on peut comparer entre elles les mesures de deux angles corniculaires quelconques, ce qui rend obsolète la remarque, faite par PROCLUS dans son *Commentaire* des célèbres *Eléments* d'EUCLIDE, selon laquelle les angles corniculaires ne peuvent pas être des *grandeurs qui ont un rapport l'une par rapport à l'autre* ([2], p. 426).

En effet, considérons deux angles corniculaires, notés respectivement γ_1 et γ_2 , ayant un même sommet P , un même côté rectiligne H , l'autre côté étant un cercle, tangent à H au point P , de rayon R_1 et R_2 . Avec des notations

conformes avec ce qui précède, les mesures de ces deux angles sont données par

$$\mu_{\gamma_1} = \sum_{j=1}^{+\infty} c_j \varepsilon^j \text{ et } \mu_{\gamma_2} = \sum_{j=1}^{+\infty} c'_j \varepsilon^j.$$

On en déduit donc

$$\frac{\mu_{\gamma_1}}{\mu_{\gamma_2}} = \frac{\frac{\mu_{\gamma_1}}{\varepsilon}}{\frac{\mu_{\gamma_2}}{\varepsilon}} = \frac{\frac{1}{2R_1} + \dots}{\frac{1}{2R_2} + \dots},$$

où les signes ... remplacent chaque fois un superréel infiniment petit. En conséquence, le quotient étudié est appréciable et sa partie standard égale à

$$\text{st} \left(\frac{\mu_{\gamma_1}}{\mu_{\gamma_2}} \right) = \frac{R_2}{R_1}.$$

Bibliographie du chapitre 5

- [1] ARTIGUE M., GAUTHERON V., ISAMBERT E., *Analyse non standard et enseignement*, Cahier de didactique des mathématiques, n° 15, I.R.E.M. - Université de Paris VII, 1985.
- [2] EUCLIDE, *Les Eléments - Volume 1*, Traduits du texte de HEIBERG, *Introduction générale* par CAVEING M. - *Livres I à IV*, Presses Universitaires de France, 1990.
- [3] HENRY V., *Questions de didactique soulevées par un enseignement de l'analyse non standard à de futurs économistes*, thèse de doctorat en didactique des mathématiques, Université P. Sabatier, Toulouse, 2004.
- [4] RADELET - DE GRAVE P., La mesure de la courbure et la pratique du calcul différentiel du second ordre, Actes du Colloque de Peyresq sur *La pensée numérique*, ALVAREZ C. - DHOMBRES J. - PONT J.C. éd., 1999, pp. 159 - 178.
- [5] TALL D., Looking at graphs through infinitesimal microscopes, windows and telescopes, *The Mathematical Gazette*, vol. 64, 1980, pp. 24 - 49.
- [6] VAN ASCH A.G. - VAN DER BLIJ F., Measurements of curvilinear angles, *Bull. Belg. Math. Soc.*, 2 (1995), pp. 573 - 588.
- [7] WALLIS J., *Operum mathematicorum pars altera, qua continentur : de Angulo contactus et semicirculi disquisitio geometrica*, Oxonii, 1654.

Chapitre 6

Vitesses, angles et infinitésimaux

Résumé. Nous proposons une approche géométrique pour introduire le concept physique de vitesse instantanée dans le cas d'un mobile ayant un déplacement unidimensionnel. Nous montrons que cette vitesse correspond à la mesure d'un angle (non nécessairement rectiligne) au sein du graphe décrivant l'évolution temporelle du mouvement. Cette méthode nous fournit une occasion pour donner une interprétation concrète de nombres non standards. Nous concluons ce texte en donnant quelques réflexions relatives à l'enseignement de l'analyse mathématique de base.

Mots clés : vitesse moyenne, vitesse instantanée, angle de contact, analyse non standard, nombre infinitésimal.

Introduction

Nous nous intéressons à un problème fondamental de la cinématique, à savoir le déplacement d'un point mobile sur une ligne droite.

Nous nous plaçons dans le cas où le point se meut sur un axe, sa position au temps t y étant repérée par l'abscisse $x(t)$; les nombres t et $x(t)$ sont exprimés dans des unités adéquates, par exemple en secondes pour le temps et en mètres pour la distance parcourue. Nous allons donc considérer la *fonction cinétique* $x : t \mapsto x(t)$, qui est supposée dérivable sur $]0, +\infty[$; nous noterons \mathcal{C} la représentation graphique correspondante : elle est supposée passer par l'origine du plan et est tout naturellement appelée *courbe cinétique* (voir la figure 6.1).

Pour estimer dans quelle mesure le mouvement étudié est rapide ou non, on évalue la *vitesse*. Nous allons nous intéresser à la vitesse en un temps donné t_0 : elle est indiquée à ce moment par l'aiguille d'un cadran (étalonné selon les unités adéquates) des vitesses dont est (virtuellement) muni le point en question.

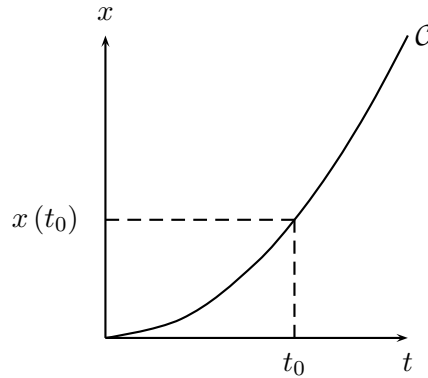


FIGURE 6.1 – Courbe cinétique

Pour un physicien expérimentateur, il est généralement malaisé de déterminer concrètement cette vitesse précise. En effet, le mobile n'est pas forcément équipé avec un cadran indiquant les vitesses ; et, dans le cas où existe un tel appareil, son aiguille a le plus souvent une position variable, difficile à localiser en un instant donné : en effet, le temps de consulter le cadran et la position de l'aiguille peut avoir changé !

Dans ce texte, nous nous proposons d'introduire ladite vitesse, qualifiée d'*instantanée*, sans faire appel aux concepts de dérivée et de limite comme on le fait habituellement, mais en travaillant géométriquement avec des angles. Nous présentons la *vitesse géométrique* au temps t_0 : elle dépend de la forme qu'a la courbe cinétique \mathcal{C} au point $P_0 = (t_0, x(t_0))$. Cette idée intuitive a toutefois un prix à payer : il faut œuvrer dans le contexte de nombres non standards.

Nous avons conçu ce travail en deux étapes. Nous avons d'abord développé, d'une manière intuitive, les idées principales (dans la section 6.1), puis avons rejeté à plus tard (dans la section 6.2) des explications qui paraissent plus techniques, notamment une introduction aux théories mathématiques assez récentes qui sont exploitées.

6.1 Introduction d'une vitesse géométrique

6.1.1 Mouvement avec vitesse constante

Considérons tout d'abord le cas d'un point se déplaçant de façon uniforme (on parlera alors de mouvement uniforme, en abrégé MU). Géométriquement, \mathcal{C} est une droite d'équation

$$x = vt$$

où v désigne la distance parcourue après une unité de temps ou encore la pente de la droite en question.

Appelons *angle cinétique*, noté $\hat{\alpha}$, l'angle rectiligne aigu $\widehat{P_0^*P_0P_1}$ où $P_0 = (t_0, vt_0)$, $P_0^* = (t_0 + \Delta t, vt_0)$ et $P_1 = (t_0 + \Delta t, v(t_0 + \Delta t))$; la mesure (voir la sous-section 6.2.4) de cet angle est en radians.

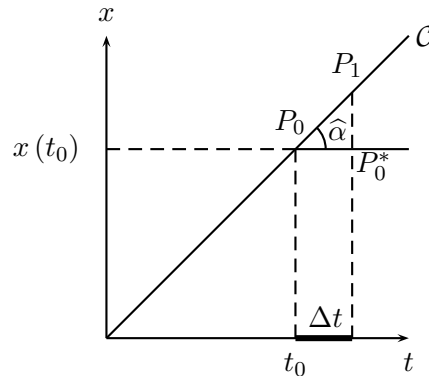


FIGURE 6.2 – Espace parcouru en fonction du temps (cas d'un MU)

Un observateur peut repérer les positions du point sur l'axe à tout moment ; ainsi, entre les instants t_0 et $t_0 + \Delta t$, avec $\Delta t > 0$, il peut calculer la distance parcourue : elle est égale à $\Delta x = x(t_0 + \Delta t) - x(t_0)$. En conséquence, il peut aisément évaluer la *vitesse moyenne* du mobile entre les deux instants considérés : elle est donnée par

$$v_{t_0;\Delta t} = \frac{\Delta x}{\Delta t} = \operatorname{tg} \left(\widehat{P_0^*P_0P_1} \right) = \frac{v(t_0 + \Delta t) - vt_0}{\Delta t} = v$$

Comme on le voit, la vitesse moyenne ne dépend pas alors de la durée Δt du mouvement observé. Il est dès lors naturel de choisir ici pour vitesse instantanée v_{t_0} au moment t_0 la valeur de chacune des vitesses moyennes correspondantes, c'est-à-dire la constante v :

$$v_{t_0} = v = \operatorname{tg} (\hat{\alpha})$$

Dans ce cas particulier, l'angle cinétique $\hat{\alpha}$ décrit parfaitement aussi bien la vitesse instantanée en t_0 que toutes les vitesses moyennes mesurées à partir de t_0 .

6.1.2 Mouvement avec une accélération constante

Considérons à présent le cas d'un mouvement possédant une vitesse variable, mais une accélération constante a (on parlera alors de mouvement uniformément accéléré, en abrégé MUA) : la courbe cinétique \mathcal{C} est une parabole d'équation

$$x = \frac{1}{2}at^2$$

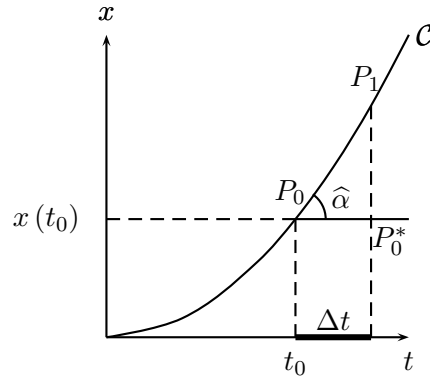


FIGURE 6.3 – Espace parcouru en fonction du temps (cas d'un MUA)

Lorsque le mobile n'est pas muni d'un cadran adéquat, la vitesse instantanée en un moment t_0 ne peut pas être observée directement ; toutefois, un expérimentateur est encore capable de calculer une vitesse moyenne sur un intervalle de temps $[t_0, t_0 + \Delta t]$. A cet effet, il considère les points $P_0 = (t_0, \frac{1}{2}at_0^2)$ et $P_1 = (t_0 + \Delta t, \frac{1}{2}a(t_0 + \Delta t)^2)$ sur la courbe cinétique, ainsi que le point $P_0^* = (t_0 + \Delta t, \frac{1}{2}at_0^2)$ situé sur l'horizontale passant par P_0 et peut calculer

$$v_{t_0;\Delta t} = \frac{\Delta x}{\Delta t} = at_0 + \frac{1}{2}a\Delta t = \text{tg} \left(\widehat{P_0^*P_0P_1} \right)$$

A présent, la vitesse moyenne change donc en fonction du temps écoulé puisqu'elle dépend de Δt .

Quant à la vitesse instantanée v_{t_0} , on sait qu'elle est ici proportionnelle à t_0 , ce qui donne la figure 6.4 :

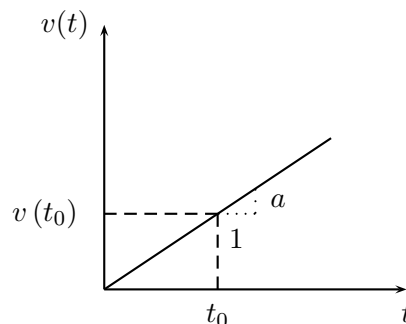


FIGURE 6.4 – Vitesse en fonction du temps (cas d'un MUA)

En fait, on a, pour tout Δt positif,

$$v_{t_0;\Delta t} = v_{t_0} + \frac{1}{2}a\Delta t$$

D'un autre point de vue, en se référant aux notations de la figure 6.5, la vitesse v_{t_0} peut être décrite géométriquement par l'angle rectiligne $\hat{\gamma}$ fait par l'horizontale passant par P_0 par la tangente \mathcal{T} en P_0 à la parabole \mathcal{C} ; en effet,

$$v_{t_0} = at_0 = \text{tg} (\hat{\gamma})$$

où $\hat{\gamma}$ est encore mesuré en radians.

Il est par ailleurs évident que cet angle $\hat{\gamma}$ est le plus grand angle rectiligne aigu de sommet P_0 ayant un côté horizontal et qui est situé sous la parabole (voir la sous-section 6.2.6).

Géométriquement, une comparaison avec la situation rencontrée pour un MU peut nous amener à considérer, dans le cas d'un MUA, l'angle $\hat{\alpha}$ de sommet P_0 , dont les côtés sont d'une part l'horizontale $P_0P_0^*$ et d'autre par la courbe \mathcal{C} (voir la figure 6.3) : cet angle n'est évidemment plus rectiligne, mais il est curvilinéaire puisque ses deux côtés sont une droite et une courbe (voir la sous-section 6.2.1).

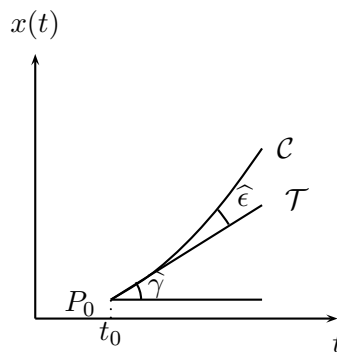


FIGURE 6.5 – Angle de contact

Pour cet angle $\hat{\alpha}$, qui semble donc caractériser une “vitesse”, nous appelons sa mesure la *vitesse géométrique* du mobile au temps t_0 : elle est notée σ_{t_0} .

Or, on a visiblement :

$$\hat{\alpha} = \hat{\gamma} + \hat{\epsilon}$$

où $\hat{\epsilon}$ est, conformément à la terminologie de NEWTON, un *angle de contact*, c'est-à-dire l'angle fait par la parabole \mathcal{C} et sa tangente \mathcal{T} en P_0 (voir la figure 6.5). La mesure de $\hat{\epsilon}$ est évidemment positive, mais ne peut pas être un nombre réel, puisqu'elle doit être inférieure à tout nombre réel du fait que l'angle de contact est plus petit que tout angle rectiligne non nul (voir la sous-section 6.2.4). En conséquence, la mesure de $\hat{\epsilon}$ est un nombre non-standard qui est bien un *infinitésimal* (voir la sous-section 6.2.5).

Dès lors, la vitesse géométrique σ_{t_0} est plus grande que le nombre réel qui est la mesure de $\hat{\gamma}$, mais plus proche de celui-ci que la mesure d'un quelconque angle rectiligne (autre que l'angle nul). Il est équivalent d'affirmer que la mesure de $\hat{\gamma}$ est l'unique réel qui est infiniment proche de σ_{t_0} , de sorte qu'il en

est sa partie observable, c'est-à-dire sa partie standard (voir la sous-section 6.2.5). On peut définir la vitesse instantanée en t_0 comme le nombre réel

$$v_{t_0} = \text{st}(\sigma_{t_0})$$

Géométriquement, v_{t_0} est la pente de la tangente en P_0 à la parabole cinétique; ce nombre peut donc être défini sans faire appel au concept de dérivée.

D'un point de vue physique, remplacer la parabole par sa tangente en P_0 revient à considérer que le mouvement devient uniforme après le temps t_0 . La vitesse instantanée est donc la vitesse (moyenne ou instantanée) qu'aurait le mobile s'il se déplaçait d'un mouvement uniforme à partir du temps t_0 . Cette présentation est conforme à des lois physiques classiques et anciennes; c'est le cas du principe d'inertie donné par GALILÉE selon lequel tout corps possède une certaine "inertie" qui l'oblige à conserver sa vitesse, à moins qu'une force extérieure, une force de frottement par exemple, ne l'oblige à arrêter ce mouvement ou encore de la première loi du mouvement de NEWTON affirmant que « tout corps persévère dans l'état de repos ou de mouvement uniforme en ligne droite dans lequel il se trouve, à moins que quelque force n'agisse sur lui, et ne le contraigne à changer d'état. »

6.1.3 Quelques compléments

Considérons à présent le cas d'un mouvement défini par une fonction cinétique quelconque x .

En nous appuyant sur l'exemple d'un mouvement avec une accélération constante, on peut introduire une *vitesse géométrique* d'un mobile au temps t_0 : c'est la mesure de l'angle mixte $\hat{\alpha}$ de sommet P_0 , avec la droite horizontale passant par P_0 comme côté rectiligne, et la courbe cinétique \mathcal{C} comme côté curviligne. Si on trace la tangente en P_0 à \mathcal{C} , on peut décomposer l'angle $\hat{\alpha}$ comme suit, avec les mêmes notations qu'à la section précédente (voir la figure 6.5) :

$$\hat{\alpha} = \hat{\gamma} + \hat{\epsilon}$$

En conséquence, la mesure de $\hat{\gamma}$ est la partie standard de la vitesse géométrique σ_{t_0} : le nombre égal à sa tangente $\text{tg}(\hat{\gamma})$ peut être pris comme étant la vitesse instantanée v_{t_0} en t_0 .

Signalons au passage la possibilité de trouver plusieurs courbes cinétiques passant par un même point où elles ont même tangente mais une convexité différente; elles correspondent donc à des mouvements pour lesquels le mobile occupe la même position en un moment donné t_0 , avec une même vitesse instantanée, mais avec des vitesses géométriques différentes. C'est le cas notamment pour les courbes cinétiques définies par

$$x = v_0 t + (t - t_0)^n$$

où n désigne un entier supérieur à 1 : la vitesse instantanée v_{t_0} est égale à v_0 quel que soit n , tandis que la vitesse géométrique dépend de l'exposant n : effectivement, les angles de contact en P_0 deviennent de plus en plus grands au fur et à mesure que n grandit ; NEWTON avait déjà remarqué cette propriété dans ses *Principia* (Book 1, Section 1, Scholium).

Ces considérations nous amènent à ces constatations concernant un enseignement de base en analyse mathématique :

- le passage de vitesses moyennes à une vitesse instantanée en cinématique nous offre une occasion concrète de rencontrer des situations fondamentales étudiées par l'analyse mathématique ;
- le concept de mesure d'un angle curvilinéaire nous fournit concrètement une opportunité d'introduire des nombres infinitésimaux.

En conclusion, l'existence de nombre infiniment petits peut être illustrée par des exemples issus de la vie courante, notamment en décrivant des situations rencontrées classiquement en physique.

Nous pensons qu'une présentation axiomatique (voir, par exemple, [1] et [2]) des nombres hyperréels peut être utilement accompagnée par un exposé semblable au contenu de notre texte, de manière à illustrer concrètement l'existence et la manipulation de nombres hyperréels.

6.2 Outils mathématiques

6.2.1 Un concept général d'angle planaire

EUCLIDE a donné, dans ses *Éléments* (Book 1, définition 8), une définition générale d'angle en géométrie plane : *un angle plan est l'inclinaison, l'une sur l'autre, dans un plan, de deux lignes qui se touchent l'une l'autre.*

Un version plus moderne et plus précise peut être trouvée dans un travail de VAN ASCH et VAN DER BLIJ [4].

Par définition, un *angle* (sous-entendu orienté) est une paire ordonnée (F_1, F_2) de courbes F_1 et F_2 issues d'un point P_0 qui est appelé le *sommet*, tandis que F_1 et F_2 sont les *côtés* de cet angle.

Nous ne considérerons de tels angles que *localement*, c'est-à-dire dans un voisinage circulaire du sommet P_0 .

Pour deux angles (F_1, F_2) et (G_1, G_2) , nous écrirons $(F_1, F_2) = (G_1, G_2)$, et nous dirons alors que les angles sont *égaux* (ou *congruents*) lorsqu'il existe une translation et / ou une rotation du plan qui envoie le sommet du premier angle sur le sommet du second, et qui transforme F_1 et F_2 localement en G_1 et G_2 respectivement ([4], p. 576).

Ainsi, il n'est pas restrictif de supposer que le sommet coïncide avec l'origine O du plan (rapporté à des axes cartésiens rectangulaires, à savoir, en recourant à des notations classiques, Ox pour l'axe horizontal et Oy pour le vertical) et que l'angle est situé dans le demi-plan de droite.

Nous admettrons également que les courbes considérées sont les graphes de fonctions différentiables, non négatives et croissantes dans un voisinage de l'origine par où elles passent. Par souci de commodité, nous utiliserons les mêmes notations pour les fonctions que pour leurs graphes.

En particulier, nous considérerons des angles dont un côté est défini par la fonction H telle que

$$H(x) = 0 \text{ si } x \geq 0$$

L'angle (H, F) est appelé *nul* si localement $H = F$; il est *rectiligne* si F est une fonction linéaire et il est *mixte* (ou *curvilinéaire*) si F n'est pas linéaire.

6.2.2 Addition d'angles

Deux angles peuvent être ajoutés lorsque le second côté du premier angle coïncide localement avec le premier côté du second angle ; par définition

$$(F_1, F_2) + (F_2, F_3) = (F_1, F_3)$$

Il est clair que l'on a dans ces conditions

$$(F_1, F_2) = (F_1, H) + (H, F_2)$$

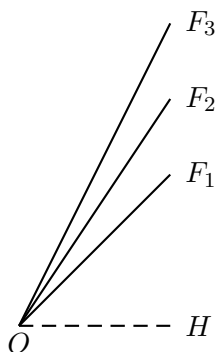


FIGURE 6.6 – Somme de deux angles

6.2.3 Un ordre pour les angles

Nous introduisons un ordre au sein de l'ensemble des angles (H, F) , où F est une courbe située dans le premier quadrant. Par définition :

$$(H, F) < (H, G)$$

lorsque le côté G de (H, G) est localement au-dessus du côté F de (H, F) , c'est-à-dire lorsqu'il existe un nombre réel positif a tel que $F(x) < G(x)$ pour tout $x \in]0, a[$.

Par exemple, sur la figure de gauche ci-dessous, l'angle (H, F) est visiblement inférieur à (H, G) puisque la droite F est située en-dessous de G . Sur la figure de droite, l'angle mixte (H, C) fait par une portion de cercle et sa tangente (cet angle étant nommé *corniculaire*), est plus petit qu'un angle rectiligne arbitraire (H, F) , car dans un voisinage du sommet O (voir les pointillés sur la figure 6.7), la droite est située au-dessus du cercle.

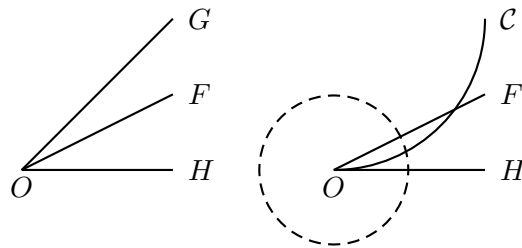


FIGURE 6.7 – Ordre sur les angles

6.2.4 Mesures d'angles

Il est toujours possible de “mesurer” n'importe quel angle (plan).

Considérons en premier lieu le cas d'un angle rectiligne aigu $\hat{\alpha} = (H, F)$ et notons C_ε le cercle ayant pour centre l'origine et pour rayon le nombre réel positif ε . Une *mesure* de $\hat{\alpha}$ est définie par la formule

$$\mu(\hat{\alpha}) = \frac{l(\varepsilon)}{\varepsilon} \quad (^\circ)$$

où $l(\varepsilon)$ est la longueur de l'arc du cercle C_ε intercepté par les deux côtés H et F de $\hat{\alpha}$; clairement, $\mu(\hat{\alpha})$ ne dépend pas de ε et est la classique mesure de cet angle exprimée en radians. Si le côté F est défini par $F(t) = mt$ où m désigne un nombre réel positif, alors ([4], p. 578)

$$\mu(\hat{\alpha}) = \text{arctg}(m)$$

Une telle mesure μ peut être étendue à n'importe quel angle (F_1, F_2) ; en d'autres termes, la formule $(^\circ)$ est valable pour F_1 et F_2 quelconques. Néanmoins, dans le cas général, $\frac{l(\varepsilon)}{\varepsilon}$ peut dépendre de ε . Par exemple, si $F_1 = H$ et F_2 est la parabole d'équation

$$x = \frac{1}{2}at^2$$

alors on a ([4], p. 578)

$$\mu(H, F) = \arcsin\left(\frac{a\varepsilon}{1 + \sqrt{1 + a^2\varepsilon^2}}\right)$$

Il est facile de vérifier qu'un angle (H, F) est inférieur à un autre angle (H, F') quand, pour tout ε suffisamment petit, la longueur $l(\varepsilon)$ de l'arc porté par \mathcal{C}_ε et situé entre les deux côtés de (H, F) est plus petite que la longueur $l'(\varepsilon)$ de \mathcal{C}_ε comprise entre les deux côtés de (H, F') ; formellement, $\mu(H, F) < \mu(H, F')$ quand il existe un nombre réel positif p tel que $l(\varepsilon) < l'(\varepsilon)$ pour tout $\varepsilon \in]0, p[$. De plus, si la fonction F est localement positive, alors $l(\varepsilon) < \frac{\pi}{2}\varepsilon$ pour n'importe quel ε assez petit, et donc $\mu(H, F) < \frac{\pi}{2}$.

Moyennant ces définitions, cette mesure μ vérifie les propriétés suivantes (voir, par exemple, dans [4], p. 577) :

- si $(F_1, F_2) = (G_1, G_2)$, alors $\mu(F_1, F_2) = \mu(G_1, G_2)$;
- μ est additive, en ce sens que $\mu(F_1, F_2) + \mu(F_2, F_3) = \mu(F_1, F_3)$;
- $\mu(F, F) = 0$;
- $\mu(F_1, F_2) = -\mu(F_2, F_1)$.
- si $(H, F) < (H, G)$, alors $\mu(H, F) < \mu(H, G)$.

6.2.5 Nombres hyperréels

Au sein de l'ensemble \mathbb{R} comprenant tous les nombres réels, il est impossible de trouver un élément non nul qui est inférieur à tous les réels positifs.

C'est pourquoi, A. ROBINSON [3] a introduit une extension de \mathbb{R} : il s'agit de l'ensemble ${}^*\mathbb{R}$ des nombres *hyperréels* (qui sont utilisés en analyse non standard).

${}^*\mathbb{R}$ est un corps algébrique totalement ordonné qui possède le corps des réels comme sous-corps, mais qui contient en plus au moins un nombre non nul ε dont la valeur absolue est inférieure à n'importe quel nombre réel positif. Un exemple concret de nombre infinitésimal est fourni par la mesure d'un angle de contact formé par une courbe et sa tangente : elle en effet visiblement non nulle, mais doit être inférieure à tout réel positif puisqu'aucune droite n'est localement située entre la courbe et sa tangente.

Il est clair que si ε est infinitésimal, alors $-\varepsilon, 2\varepsilon, \varepsilon^2, \dots$ sont également infinitésimaux. De plus, l'inverse d'un infinitésimal non nul est un nombre *infini* : c'est un nombre hyperréel qui est, en valeur absolue, supérieur à tout réel positif.

Pour un nombre hyperréel x qui n'est pas infini, il existe un seul et un seul réel r qui lui est *infinitement proche*, c'est-à-dire tel que la différence $x - r$ est un infinitésimal ; r est usuellement appelé la *partie standard* de x .

6.2.6 Le plus grand angle rectiligne sous une parabole

Plaçons-nous dans le cas d'un mouvement avec une accélération a constante, de sorte que la courbe \mathcal{C} a pour équation $x = \frac{1}{2}at^2$. La tangente \mathcal{T} à \mathcal{C} en $P_0 = (t_0, \frac{1}{2}at_0^2)$ a pour équation $x(t) = \frac{1}{2}at_0^2 + m(t - t_0)$, la pente m de cette droite étant égale à $m = at_0$. En vertu de la convexité de la fonction cinétique, la tangente \mathcal{T} est située sous la courbe \mathcal{C} ; de plus, toute droite passant par P_0

et de pente $m' = m + r$, avec $r > 0$, est située au-dessus de la parabole \mathcal{C} pour tout t appartenant à l'intervalle $]t_0, t_0 + \frac{2r}{a}[$. Il en résulte que le plus grand angle rectiligne sous la parabole est bien celui dont le côté non horizontal est porté par la tangente \mathcal{T} .

Bibliographie du chapitre 6

- [1] BAIR J. - HENRY V., *Analyse mathématique - le calculus redécouvert*, Ed. Academia - Bruylant, Louvain-la-Neuve, 2008.
- [2] KEISLER H.J., *Elementary Calculus*, Prindle, Weber & Schmidt Inc., Boston, 1976.
- [3] ROBINSON A., *Non-standard Analysis*, Revised edition, North-Holland Pub. Comp., Amsterdam - London, 1974.
- [4] VAN ASCH A.G. - VAN DER BLIJ F., Measurements of curvilinear angles, *Bull. Belg. Math. Soc.*, 2, 1995, pp. 573 - 588.

Chapitre 7

Construction des hyperréels selon Robinson

Résumé. Nous avons expérimenté, en deuxième année du programme d'ingénierat de gestion à l'Université de Liège, un enseignement d'une construction des nombres hyperréels qui est inspirée par les travaux de A. PÉTRY. Après avoir rappelé la méthode, nous analysons quelques réactions des étudiants face à cette matière.

Introduction

Bien que les nombres infiniment petits aient été utilisés avec efficacité par des savants dès le 16^{ème} siècle, il a fallu attendre la seconde moitié du 20^{ème} siècle, avec les travaux de ROBINSON [3], pour que leur existence soit indiscutable d'un point de vue mathématique.

La construction robinsonienne des hyperréels utilise de façon fondamentale la notion mathématique de filtre. Plus précisément, il s'agit d'une ultrapuissance du champ des réels modulo un ultrafiltre non principal sur l'ensemble \mathbb{N} des entiers naturels, l'existence d'un tel ultrafiltre étant garantie par l'axiome du choix.

Des versions pédagogiques de cette construction peuvent être trouvées dans la littérature. Toutefois, elles restent fort abstraites et les notions qu'elles abordent sont inconnues des étudiants sortant de l'enseignement secondaire et d'une première année à l'Université.

C'est pourquoi, il nous a semblé opportun de proposer à des étudiants de deuxième année en ingénierat de gestion une présentation rigoureuse des hyperréels qui devrait a priori leur sembler plus abordable. A cet effet, nous avons opté pour la version pédagogique donnée par PÉTRY ([1], [2]) qui fait appel à une sorte de "mesure", un peu semblable à une probabilité, de manière à éviter le recours à des filtres et ultrafiltres.

7.1 Une \mathbb{N} -mesure

Travaillant sur les parties de l'ensemble \mathbb{N} des naturels $0, 1, 2, 3, \dots$, on introduit une "mesure μ sur \mathbb{N} ", ou plus simplement d'une \mathbb{N} -mesure μ de la manière suivante :

Définition 7.1.1 μ est une \mathbb{N} -mesure lorsque μ est une fonction dont l'ensemble de définition est l'ensemble des parties de \mathbb{N} , dont les valeurs sont 0 ou 1 et telle que

1. $\mu(\mathbb{N}) = 1$,
2. $\mu(X) = 0$ pour toute partie finie X de \mathbb{N} ,
3. $\mu(X \cup Y) = \mu(X) + \mu(Y)$ pour toutes parties X, Y de \mathbb{N} telles que $X \cap Y = \emptyset$.

Ces trois conditions ont été reliées à la présentation axiomatique d'une probabilité; signalons toutefois qu'une telle fonction μ n'est pas σ -additive (mais seulement finiment additive) sur l'ensemble des parties de \mathbb{N} , de sorte qu'il ne s'agit pas d'une mesure habituelle au sens de la théorie de la mesure. Par ailleurs, l'existence de cette \mathbb{N} -mesure est postulée dans notre présentation bien qu'elle découle de l'axiome du choix, ce dernier étant supposé admis par nos étudiants; à l'instar de A. PÉTRY, nous considérerons ultérieurement une \mathbb{N} -mesure μ fixée une fois pour toutes ([1], p. 83).

De cette définition ont été déduites et démontrées les propriétés suivantes de cette \mathbb{N} -mesure μ^1 :

1. $X \subseteq Y$ implique $\mu(X) \leq \mu(Y)$;
2. $\mu(X) = 0$ si et seulement si $\mu(\mathbb{N} \setminus X) = 1$, $\mathbb{N} \setminus X$ désignant le complémentaire de X dans \mathbb{N} ;
3. $\mu(X \cup Y) = \mu(X) + \mu(Y) - \mu(X \cap Y)$;
4. $\mu(X) = 1$ et $\mu(Y) = 1$ impliquent $\mu(X \cap Y) = 1$;
5. $\mu(X \cup Y) = 1$ implique $\mu(X) = 1$ ou $\mu(Y) = 1$.

7.2 Définition et propriétés de ${}^*\mathbb{R}$

La construction de ROBINSON se réalise en travaillant au sein de l'ensemble S composé de toutes les suites de réels. On définit sur S une relation, notée \sim , comme suit : si (x_i) et (y_i) sont des suites de réels,

$$(x_i) \sim (y_i) \iff \mu(\{i : x_i = y_i\}) = 1.$$

1. D'après ces propriétés, l'existence de μ équivaut à l'existence d'un ultrafiltre sur \mathbb{N} : effectivement, $\mu(X) = 1$ si et seulement si X est un élément de l'ultrafiltre considéré.

Les propriétés de la mesure μ permettent de démontrer aisément que \sim est une relation d'équivalence sur l'ensemble S ([1], p. 83). Il est dès lors permis de considérer les classes d'équivalence de \sim ; pour toute suite de réels (x_i) , sa classe d'équivalence pour \sim est notée $\langle x_i \rangle$; cela conduit à cette définition des hyperréels :

Définition 7.2.1 ${}^*\mathbb{R}$ est l'ensemble de toutes les classes d'équivalence de \sim sur S . Les éléments de ${}^*\mathbb{R}$ sont appelés **hyperréels**.

On définit ensuite sur ${}^*\mathbb{R}$ une addition, une multiplication, une fonction module et une relation d'ordre \leq de la manière suivante :

Définition 7.2.2

- $\langle x_i \rangle + \langle y_i \rangle = \langle x_i + y_i \rangle$,
- $\langle x_i \rangle \times \langle y_i \rangle = \langle x_i \times y_i \rangle$,
- $|\langle x_i \rangle| = \langle |x_i| \rangle$,
- $\langle x_i \rangle \leq \langle y_i \rangle \iff \mu(\{i : x_i \leq y_i\}) = 1$.

Ces définitions sont cohérentes, car les notions définies sont indépendantes du choix des suites (x_i) et (y_i) sélectionnées dans les classes d'équivalence considérées. De plus, elles conduisent à ce théorème ([1], pp. 83-84) :

Théorème 7.2.3 $\langle {}^*\mathbb{R}, +, \times, \leq, \langle 0 \rangle, \langle 1 \rangle \rangle$ est un champ ordonné.

Il convient ensuite de montrer que les réels peuvent être retrouvés dans ${}^*\mathbb{R}$. Pour chaque réel r , désignons par (r) la suite dont tous les termes sont égaux à r et par $\langle r \rangle$ la classe d'équivalence de cette suite (r) . La fonction définie sur \mathbb{R} par

$$x \in \mathbb{R} \mapsto \langle x \rangle$$

définit une bijection de \mathbb{R} sur $\{\langle u \rangle : u \in \mathbb{R}\}$. Au surplus, il est équivalent d'effectuer les opérations arithmétiques sur les réels a, b, c, \dots ou sur leur interprétation $\langle a \rangle, \langle b \rangle, \langle c \rangle, \dots$ dans ${}^*\mathbb{R}$: par exemple, on a $\langle a \rangle + \langle b \rangle = \langle a + b \rangle$. En conséquence, on identifie chaque réel r avec ${}^*r = \langle r \rangle$, de sorte que le champ associé à ${}^*\mathbb{R}$ apparaît comme une extension du champ des réels ([1], p. 84).

7.3 Existence d'infiniment petits ou grands

Nous sommes à présent en mesure de trouver des nombres non nuls dont la valeur absolue est inférieure à tout réel positif; il convient néanmoins de sortir du cadre traditionnel des nombres réels et de travailler avec des hyperréels. Nous nommerons *infiniment petit* tout nombre hyperréel non réel qui est inférieur en valeur absolue à tout réel positif.

Voici un énoncé garantissant l'existence d'hyperréels infiniment petits ([1], p. 85).

Théorème 7.3.1 *Si (x_i) est une suite de réels de limite 0, l'hyperréel $\langle x_i \rangle$ est un infiniment petit. En particulier, $\langle \frac{1}{1+i} \rangle$ est un infiniment petit.*

En corollaire, il est possible d'exhiber des infiniment grands positifs, à savoir des hyperréels supérieurs à tout réel positif : il suffit bien entendu de considérer l'inverse d'un hyperréel infiniment petit positif. On vérifie d'ailleurs directement que l'hyperréel $\langle 1+i \rangle$ est infiniment grand positif. De même, l'opposé d'un hyperréel infiniment grand positif est un infiniment grand négatif, c'est-à-dire un hyperréel inférieur à tout réel négatif ; c'est le cas notamment pour $\langle -i \rangle$.

7.4 Une expérimentation

Nous avons présenté la théorie ci-dessus à des étudiants de deuxième année à l'Université ; en première année, ils avaient dû admettre l'axiome de l'existence d'au moins un nombre hyperréel infinitésimal non nul.

Quarante étudiants de la seconde candidature en ingénierat de gestion ont participé à une leçon sur le sujet. A l'issue de celle-ci, nous leur avons posé une question très ouverte pour connaître leur impression générale sur ce cours ; la question était formulée comme suit :

« Comment avez-vous perçu le cours sur la construction de ${}^*\mathbb{R}$ selon la méthode de ROBINSON ? (compréhension, difficulté, niveau d'abstraction, utilité, ...). »

Voici une analyse sommaire des réponses fournies.

- Plus de la moitié des élèves (soit exactement 25) ont avoué ne guère avoir compris ce cours, surtout avec l'introduction de la \mathbb{N} -mesure μ (8 mentions) et le recours à des classes d'équivalence : il s'agit de notions nouvelles, difficiles à imaginer et à manipuler (10 mentions) ; de plus, les analogies et différences entre la mesure μ et une probabilité n'ont pas été bien perçues.
- De nombreux étudiants ont trouvé cette leçon assez, voire extrêmement difficile (14 mentions) et fort abstraite (8 mentions) ; certains ont été surpris (7 mentions), car ils « ne s'étaient pas imaginés que les hyperréels pouvaient être introduits de cette façon » ; l'un d'entre eux s'est même dit « déstabilisé. »
- D'aucuns, en petit nombre, estimaient avoir déjà assimilé les notions de l'analyse non standard (nombres infiniment petits, notions d'infiniment proche ou de halo, ...) ; c'est pourquoi, cette présentation, plus compliquée que celle vue l'année précédente, ne leur a paru guère utile (2 mentions) ; d'ailleurs, comme l'a écrit un autre, « heureusement que l'approche pédagogique² avait précédé ce cours. »

2. Il s'agit de l'approche inspirée de KEISLER et vue en première année.

- Pour d'autres élèves, ce cours a paru intéressant (10 mentions), mais avec des justifications assez variées. Ainsi, une étudiante estime que « cette nouvelle approche les habitue à sortir du cadre traditionnel et les fait réfléchir, ce qui devrait s'avérer utile à long terme. »
- Deux autres personnes ont été impressionnées par la diversité des raisonnements en mathématiques, ce qui « prouve qu'il ne s'agit pas d'une science fermée. » Dans le même ordre d'idées, un étudiant fait remarquer que la méthode de ROBINSON met bien en évidence le fait qu'il « n'existe pas qu'une et une seule manière de présenter une même matière. »
- Trois personnes estiment que cette présentation leur permettra d'admettre plus facilement que les hyperréels existent et de « mieux comprendre d'où ils viennent » ; pour l'un d'eux, « le déclic s'est d'ailleurs fait plus rapidement » que l'an passé. Enfin, un étudiant a écrit qu'il se posait jusqu'alors des questions à propos « de la continuité de la droite des réels, car si on y rajoute des nombres, c'est qu'il y avait des trous » ; l'approche de ROBINSON l'a complètement rassuré sur ce point et il termine sa copie, avec peut-être un brin d'humour, par ces mots : « dommage qu'il n'y manquait que la compréhension. »

Même si un étudiant a trouvé cette matière « amusante quand on la connaît », la conclusion de cette expérience semble pouvoir être donnée par deux étudiants qui estiment que « pour des gestionnaires, il suffit de savoir que les hyperréels existent » ; pour eux, cette théorie de ROBINSON devrait être « réservée à un public averti de mathématiciens. »

Bibliographie du chapitre 7

- [1] PÉTRY A., Introduction à l'Analyse non standard ou la redécouverte des infiniment petits, in *Les bases scientifiques de l'Etude du Milieu*, CSIP-WIC, Commissariat Général aux Relations Internationales de la Communauté française de Belgique, Liège, 1995, pp. 82 - 91.
- [2] PÉTRY A., Balade en Analyse non standard sur les traces de A. ROBINSON, *Non Standard Analysis*, Belgian Mathematical Society, 1996, pp. 1 - 28.
- [3] ROBINSON A., *Non-standard Analysis*, Koninkl. Ned. Akad. Wetesch. Proc. Ser. A 64, 1961, pp. 432 - 440.

Chapitre 8

Algorithme de Sluse : une présentation contemporaine

Résumé. Nous nous penchons sur un ancien algorithme, mis au point par Sluse au 17^{ème} siècle pour trouver une tangente à une courbe. Nous comparons la présentation originale à une version moderne développée dans le cadre des nombres hyperréels et exploitant un microscope infiniment puissant.

8.1 Un savant liégeois

8.1.1 Une courte biographie

René-François Walter de SLUSE (parfois appelé René de SLUZE) est né le 2 juillet 1622 à Visé, une petite ville de la région liégeoise tout près de la frontière entre la Belgique et la Hollande, au sein d'une famille aisée et cultivée. Après avoir étudié le droit à l'Université Catholique de Louvain de 1638 à 1642, il partit à Rome où il séjourna pendant une petite dizaine d'années : il y décrocha en 1643 le titre de docteur en droit de l'université de Rome *La Sapienza*, puis eut l'occasion d'étudier en Italie la théologie, mais aussi la philosophie, l'histoire, les mathématiques, la physique, l'astronomie ainsi que diverses langues (on prétend qu'il était compétent en grec, hébreu, arabe et syriaque) ; il eut aussi l'occasion de rencontrer des savants réputés de l'époque, notamment CAVALIERI et TORRICELLI.

Ensuite, il revint dans son pays natal pour y exercer diverses fonctions au sein de l'Eglise de la principauté de Liège : il fut nommé chanoine de la cathédrale Saint-Lambert en 1650, membre du conseil privé de la cathédrale de Liège en 1659 et enfin abbé de la Collégiale d'Amay en 1666 ; il fut encore conseiller privé du prince-évêque Maximilien-Henry de Bavière.

A cause de ses lourdes charges ecclésiastiques, SLUSE n'eut plus l'occasion de voyager autant qu'il l'aurait souhaité. Pour satisfaire sa grande curiosité intellectuelle, il entretint toutefois une correspondance soutenue avec quelques

savants européens de son époque : le grand savant français PASCAL, le renommé scientifique hollandais HUYGENS, le chanoine et mathématicien italien RICCI, le mathématicien anglais WALLIS, le secrétaire de la Royal Society de Londres (dont il devint membre en 1674) OLDENBOURG, ... ; c'est principalement grâce à sa correspondance, publiée en 1884 par le mathématicien liégeois LE PAIGE [8], que ses travaux scientifiques furent connus.

Il mourut à Liège le 19 mars 1685.

8.1.2 L'œuvre mathématique

SLUSE était, d'après ROSENFELD [10], « un homme d'une érudition formidable, doué pour les mathématiques, d'un esprit d'invention et de généralisation très remarquable. »

Les sujets mathématiques abordés par ce savant étaient bien évidemment liés aux préoccupations scientifiques de l'époque. D'une part, le 16^{ème} siècle remit à l'honneur des préoccupations géométriques datant de l'Antiquité (avec, entre autres, les travaux d'EUCLIDE et d'ARCHIMÈDE) : notamment, CAVALLIERI développa sa « théorie des indivisibles », qui eut une influence certaine sur le développement de l'analyse mathématique. D'autre part, c'est à cette époque que l'algèbre se développa au départ de la France avec VIÈTE, qui introduisit le langage algébrique, puis, surtout, DESCARTES avec sa « géométrie analytique ». La conjonction de ces deux courants a conduit de nombreux savants contemporains de SLUSE à s'intéresser à des propriétés géométriques de courbes définies algébriquement, telles que les spirales et la cycloïde qui étaient des thèmes à la mode à cette époque ; notamment, le problème de tangentes devint un des principaux sujets étudiés par les mathématiciens de la deuxième moitié du 17^{ème} siècle.

Dans plusieurs lettres adressées à HUYGENS en 1657 et 1658 et à PASCAL en 1658, il étudia des courbes définies par une équation générale du type

$$y^n = k(a - x)^p x^m$$

où a et k désignent des constantes réelles tandis que les exposants n , p et m sont des entiers positifs : le savant liégeois attribua l'invention de ces courbes à PASCAL mais ce dernier leur donna le suggestif et joli nom de « perles de Sluse » sous lequel elles sont encore connues de nos jours ; ces perles sont ainsi nommées en raison de leur forme pour certaines valeurs des paramètres dans leur équation.

SLUSE reprit les notations de VIÈTE, et proposa dans son ouvrage intitulé *Mesolabum* (première édition en 1659 et deuxième édition en 1668) une méthode de résolution géométrique des équations qui s'avéra « bien plus souple et plus élégante que celle de Descartes ».

A la suite de FERMAT et de BARROW notamment, il se pencha sur le problème des tangentes. Selon l'historien BOYER [4], il fut peut-être le premier

mathématicien à fournir un algorithme général pour trouver la tangente à des courbes telles que ses perles. Sa méthode, probablement découverte en 1652 mais publiée seulement en 1673, fournit en fait la valeur de la “sous-tangente” et la tangente en un point P_0 d’une courbe \mathcal{C} : dans un repère orthonormal, si la tangente à \mathcal{C} en $P_0 = (x_0, y_0)$ et la droite verticale menée par ce même point P_0 coupent l’axe des abscisses aux points T et M respectivement, la sous-tangente en question est définie par la mesure algébrique, notée σ , du segment TM , tandis que la pente m de la tangente correspondante vaut alors $\frac{y_0}{\sigma}$.

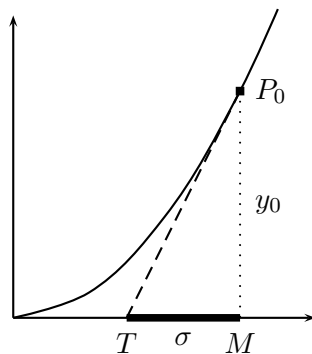


FIGURE 8.1 – Sous-tangente

Nous allons présenter cette méthode, appelée désormais l’algorithme de SLUSE, sur un exemple élémentaire, puis nous la justifierons en donnant des explications inspirées par des travaux de ROSENFELD [10] et de BOSCKSTAELE [3]; ensuite, nous la comparerons à une méthode moderne recourant à l’usage de microscopes infiniment puissants utilisés dans le cadre des nombres hyperréels.

8.2 L’algorithme de Sluse

8.2.1 Un exemple

Nous allons en premier lieu expliquer comment appliquer l’algorithme de SLUSE dans le cas particulier d’une courbe appelée une cubique de SLUSE d’équation

$$a(x - a)(x^2 + y^2) = b^2x^2$$

Soit $f : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ la fonction définie par

$$f(x, y) = ax^3 - (a^2 + b^2)x^2 + axy^2 - a^2y^2$$

Afin de rechercher la pente m de la tangente en un point $P_0 = (x_0, y_0)$, supposé d’ordonnée non nulle, nous pouvons procéder comme suit :

Étape 1. On considère la partie de $f(x, y)$ composée des termes contenant la variable x , soit $ax^3 - (a^2 + b^2)x^2 + axy^2$; on multiplie chaque terme de cette expression par l'exposant de x , puis on divise le résultat par x : on obtient de la sorte le résultat, noté ϕ_1 , suivant

$$\phi_1 = 3ax^2 - 2(a^2 + b^2)x + ay^2$$

Étape 2. On retient de $f(x, y)$ les termes contenant la variable y , à savoir $axy^2 - a^2y^2$; chaque terme de cette expression est multiplié par l'exposant de y , mais, contrairement à ce qui a été réalisé ci-dessus, on ne divise pas cette expression par la variable (ici y); on obtient de la sorte le résultat suivant

$$\phi_2 = 2axy^2 - 2a^2y^2$$

Étape 3. On divise ϕ_2 par $-\phi_1$ et on remplace x par x_0 et y par y_0 ; on trouve

$$\sigma = -\frac{2ax_0y_0^2 - 2a^2y_0^2}{3ax_0^2 - 2(a^2 + b^2)x_0 + ay_0^2}$$

Étape 4. On divise y_0 par σ et on trouve

$$m = \frac{y_0}{\sigma} = -\frac{3ax_0^2 - 2(a^2 + b^2)x_0 + ay_0^2}{(2ax_0 - 2a^2)y_0}$$

8.2.2 Justification de l'algorithme

Nous allons expliquer cet algorithme en travaillant sur une courbe algébrique \mathcal{C} définie par l'équation

$$f(x, y) = \sum_{i=0}^p \sum_{j=0}^q a_{ij} x^i y^j = 0$$

(où p et q désignent des entiers positifs supérieurs à 1, tandis que tous les a_{ij} sont des nombres réels arbitraires).

Nous considérons donc un point $P_0 = (x_0, y_0)$ sur \mathcal{C} et nous nous proposons de chercher la tangente \mathcal{T} (qui est supposée exister et être unique) à la courbe \mathcal{C} au point P_0 .

Soit $P_1 = (x_1, y_1)$ un point de \mathcal{C} avec $x_1 \neq x_0$. Dès lors, nous pouvons écrire

$$f(x_0, y_0) - f(x_1, y_1) = \sum_{i=0}^p \sum_{j=0}^q a_{ij} (x_0^i y_0^j - x_1^i y_1^j) = 0$$

ou de façon équivalente

$$f(x_0, y_0) - f(x_0, y_1) + f(x_0, y_1) - f(x_1, y_1) = 0$$

et après quelques calculs algébriques élémentaires (utilisant notamment une formule donnant la différence de puissances) :

$$(y_0 - y_1) \sum_{i=0}^p x_0^i \sum_{j=1}^q a_{ij} S_j + (x_0 - x_1) \sum_{j=0}^q y_1^j \sum_{i=1}^p a_{ij} T_i = 0 \quad (*)$$

où S_j (resp. T_i) est la somme des j (resp. i) termes de la forme $y_0^{j-1-k} y_1^k$ pour $k = 0, \dots, j-1$ (resp. égaux à $x_0^{i-1-k} x_1^k$ pour $k = 0, \dots, i-1$).

Désignons par $P^* = (x^*, 0)$ l'intersection de la droite $P_0 P_1$ avec l'axe horizontal : nous avons

$$x^* = \frac{x_0 y_1 - x_1 y_0}{y_1 - y_0}$$

Considérons ensuite les points $P_2 = (x_0, y_1)$ et $P_3 = (x_0, 0)$: les triangles $P_0 P_1 P_2$ et $P_0 P_3 P^*$ sont semblables, de sorte que

$$\frac{y_0 - y_1}{y_0} = \frac{x_0 - x_1}{x_0 - x^*}$$

En conséquence, (*) devient

$$y_0 \sum_{i=0}^p x_0^i \sum_{j=1}^q a_{ij} S_j + (x_0 - x^*) \sum_{j=0}^q y_1^j \sum_{i=1}^p a_{ij} T_i = 0$$

A présent, si nous faisons tendre P_1 vers P_0 , nous allons prendre la limite de tous les termes de la dernière égalité lorsque x_1 et y_1 tendent respectivement vers x_0 et y_0 . La limite de S_j est égale à $j y_0^{j-1}$, celle de T_i à $i x_0^{i-1}$, tandis que la limite de $x_0 - x^*$ vaut σ . En conséquence, nous avons (pour autant que l'opération algébrique en question puisse être réalisée)

$$\sigma = - \frac{y_0 \sum_{i=0}^p \sum_{j=1}^q a_{ij} x_0^i j y_0^{j-1}}{\sum_{i=1}^p \sum_{j=0}^q a_{ij} i x_0^{i-1} y_0^j}$$

Dès lors, la pente m de la tangente \mathcal{T} est donnée par

$$m = \frac{y_0}{\sigma} = - \frac{\sum_{i=1}^p \sum_{j=0}^q i a_{ij} x_0^{i-1} y_0^j}{\sum_{i=0}^p \sum_{j=1}^q j a_{ij} x_0^i y_0^{j-1}}$$

8.3 Une approche nouvelle au moyen de microscopes

8.3.1 Le cadre de la méthode

Nous allons travailler avec des nombres hyperréels qui ont été introduits rigoureusement par A. ROBINSON en 1961 [9] ; nous travaillerons donc dans le cadre de l'analyse non - standard et utiliserons essentiellement les définitions et notations adoptées par KEISLER [7].

Pour la bonne compréhension de ce texte, rappelons que les nombres hyper-réels forment une extension des nombres réels avec les mêmes lois algébriques ; techniquement, l'ensemble ${}^*\mathbb{R}$ des nombres hyperréels est un corps, qui n'est pas archimédien, dont le corps \mathbb{R} des réels, muni des lois d'addition et de multiplication traditionnelles, est un sous-corps (qui lui est achimédien). De plus, ${}^*\mathbb{R}$ contient au moins un (et dès lors une infinité) nombre *infinitésimal* non nul : il s'agit d'un nombre ε tel que sa valeur absolue est positive mais inférieure à n'importe quel nombre réel positif ; l'inverse d'un tel nombre, soit $\frac{1}{\varepsilon}$, est *infini*, encore appelé *infinitement grand* : c'est un nombre dont la valeur absolue est supérieure à tout nombre réel. Evidemment, un nombre infinitésimal non nul ou un nombre infini n'est pas réel. Un nombre hyperréel x qui est fini, c'est-à-dire qui n'est pas infinitement grand, se voit associer un et un seul nombre réel r qui est *infinitement proche* de x , en ce sens que la différence $x - r$ est un infinitésimal : r est appelé la *partie standard* de x et est notée $r = \text{st}(x)$; formellement, $\text{st} : {}^*R \mapsto \mathbb{R} : x \mapsto \text{st}(x)$ est un homomorphisme d'anneau de l'ensemble des nombres hyperréels finis dans \mathbb{R} et dont le noyau est l'ensemble des infinitésimaux.

Le concept de microscope (virtuel) est bien connu (voir, par exemple, [1], [5], [11]). Néanmoins, rappelons que, pour un point $P = (a, b)$ du plan hyperréel ${}^*\mathbb{R}^2$ et un nombre infinitement grand H , le microscope *pointé* sur P et de *puissance* H agit comme une application, notée par \mathcal{M}_H^P , définie sur ${}^*\mathbb{R}^2$ comme suit

$$\mathcal{M}_H^P : (x, y) \mapsto (X, Y) \text{ avec } X = H(x - a) \text{ et } Y = H(y - b)$$

ce qui équivaut à écrire

$$x = a + \frac{X}{H} \text{ et } y = b + \frac{Y}{H}$$

8.3.2 La méthode proprement dite

Soient $P_0 = (x_0, y_0)$ un point situé sur une courbe algébrique \mathcal{C} à nouveau définie par

$$f(x, y) = \sum_{i=0}^p \sum_{j=0}^q a_{ij} x^i y^j = 0$$

et H un hyperréel infinitement grand positif arbitraire. Nous choisissons encore un point $P_1 = (x_1, y_1)$ sur \mathcal{C} tel que $x_1 = x_0 + \frac{X}{H}$ et $y_1 = y_0 + \frac{Y}{H}$ pour des nombres finis X et Y : P_0 et P_1 sont dès lors deux points de la courbe qui sont infinitement proches l'un de l'autre ; X et Y sont les coordonnées de l'image de P_1 dans l'oculaire du microscope $\mathcal{M}_H^{P_0}$.

Comme P_1 est situé sur \mathcal{C} , nous avons

$$\sum_{i=0}^p \sum_{j=0}^q a_{ij} \left(x_0 + \frac{X}{H}\right)^i \left(y_0 + \frac{Y}{H}\right)^j = 0$$

ou, de façon équivalente

$$\sum_{i=0}^p \sum_{j=0}^q a_{ij} \left[x_0^i + \frac{1}{H} (i x_0^{i-1} X + \dots) \right] \times \left[y_0^j + \frac{1}{H} (j y_0^{j-1} Y + \dots) \right] = 0$$

sachant que les expressions notées par le symbole \dots sont infinitésimales puisqu'elles sont des sommes dont tous les termes contiennent une puissance de H en dénominateur. Quelques calculs algébriques élémentaires conduisent à

$$\sum_{i=0}^p \sum_{j=0}^q a_{ij} x_0^i y_0^j + \frac{1}{H} \left(Y \sum_{i=0}^p \sum_{j=1}^q j a_{ij} x_0^i y_0^{j-1} + X \sum_{i=1}^p \sum_{j=0}^q i a_{ij} x_0^{i-1} y_0^j + \dots \right) = 0$$

En tenant compte de l'appartenance de P_0 à la courbe \mathcal{C} et en multipliant tous les termes par H , nous trouvons

$$Y \sum_{i=0}^p \sum_{j=1}^q j a_{ij} x_0^i y_0^{j-1} + X \sum_{i=1}^p \sum_{j=0}^q i a_{ij} x_0^{i-1} y_0^j + \dots = 0$$

Un passage par les parties standards dans les deux membres de cette dernière égalité mène à

$$Y \sum_{i=0}^p \sum_{j=1}^q j a_{ij} x_0^i y_0^{j-1} + X \sum_{i=1}^p \sum_{j=0}^q i a_{ij} x_0^{i-1} y_0^j = 0$$

Cette équation est celle d'une droite au sein du repère XOY . En fait, cette image linéaire de la courbe \mathcal{C} dans l'oculaire du microscope est la parallèle à la tangente \mathcal{T} cherchée et la pente m de \mathcal{T} est donc donnée par

$$m = - \frac{\sum_{i=1}^p \sum_{j=0}^q i a_{ij} x_0^{i-1} y_0^j}{\sum_{i=0}^p \sum_{j=1}^q j a_{ij} x_0^i y_0^{j-1}}$$

8.3.3 Comparaison avec l'algorithme de Sluse

Nous pouvons observer que les deux expressions $x_0 - x_1$ et $\frac{X}{H}$ qui interviennent respectivement dans la méthode de SLUSE et dans celle du microscope jouent des rôles tout à fait similaires dans les calculs. Dans le cas des courbes algébriques régulières, les deux méthodes sont équivalentes sachant que le passage à la limite dans l'algorithme de SLUSE correspond au recours aux parties standards dans la méthode du microscope.

Mais, la méthode du microscope peut être appliquée même lorsque l'algorithme de SLUSE ne peut pas fonctionner, ce qui est notamment le cas quand la courbe \mathcal{C} ne possède pas, au point P_0 considéré, une unique tangente, ni lorsque la tangente est verticale (car alors $x_0 = x_1$), ni enfin lorsque la tangente est horizontale (car le dénominateur de σ est alors nul).

En guise d'exemples, considérons la perle de SLUSE définie par

$$y^4 = x^4(1 - x)$$

Voici trois cas particuliers pour lesquels l'algorithme de SLUSE ne peut pas être appliqué.

1. L'origine du plan en est un de ses points, mais il est singulier (en ce sens qu'il n'y existe pas une tangente unique). Si nous appliquons un microscope de puissance infiniment grande H pointé vers l'origine, c'est-à-dire pour $x_0 = 0$ et $y_0 = 0$, alors nous obtenons

$$\frac{Y^4}{H^4} = \frac{X^4}{H^4} - \frac{X^5}{H^5} \iff Y^4 = X^4 - \frac{X^5}{H}$$

finalemt, après un passage par les parties standards des deux membres, nous trouvons

$$Y^4 = X^4 \iff Y = \pm X$$

En conséquence, en l'origine, la courbe considérée possède non pas une, mais bien deux tangentes, à savoir les droites $y = x$ et $y = -x$.

2. Si nous dirigeons un microscope de puissance infinie H vers le point $(1, 0)$, alors nous avons

$$\frac{Y^4}{H^4} = \left(1 + \frac{X}{H}\right)^4 - \left(1 + \frac{X}{H}\right)^5$$

Nous trouvons finalement $X = 0$: il existe donc en ce point une tangente verticale.

3. Si nous pointons un microscope sur le point $\left(\frac{4}{5}, \frac{4}{5\sqrt[4]{5}}\right)$, nous obtenons

$$\left(\frac{4}{5\sqrt[4]{5}} + \frac{Y}{H}\right)^5 = \left(\frac{4}{5} + \frac{X}{H}\right)^4 - \left(\frac{4}{5} + \frac{X}{H}\right)^5$$

puis enfin $Y = 0$: la tangente est cette fois horizontale.

En conclusion, on peut affirmer que la méthode nouvelle est très performante en ce qui concerne des courbes algébriques et que les microscopes jouent bien le rôle de *médiateurs épistémiques* [5].

Bibliographie du chapitre 8

- [1] BAIR J. - HENRY V., *Analyse infinitésimale : le calculus redécouvert*, Academia - Bruylant, Louvain-la-Neuve, 2008.
- [2] BAIR J. - HENRY V., Sluse, ses perles et son algorithme, *Losanges*, n° 14, 2011, pp. 14 - 18.
- [3] BOCKSTAELE P., La théorie des tangentes aux courbes algébriques dans l'œuvre de René-François de Sluse, *Bull. Soc. Roy. Sc. Liège*, 545 (1), 1986, pp. 135 - 144.
- [4] BOYER *The History of the Calculus and its Conceptual Development*, Courier Dover Publications, 1959.
- [5] DOSSENA R. - MAGNANI L., Mathematics through Diagrams : Microscopes in Non-Standard and Smooth Analysis, *Studies in Computational Intelligence*, 64, 2007, pp. 193 - 213.
- [6] JONGMANS F. - HALLEUX R. - LEFEBVRE P. - BERNES A.C., *Les Sluse et leur temps : une famille, une ville, un savant au XVII^e siècle*, Edition du Crédit Communal, Bruxelles, 1985.
- [7] KEISLER H.J., *Elementary Calculus*, Prindle, Weber & Schmidt Inc., Boston, 1976.
- [8] LE PAIGE C., Correspondance de René-François de Sluse, *Bullettino di Bibliografia e di Storia delle Scienze matematiche e fisiche*, t. XVII, Rome, 1814.
- [9] ROBINSON A., Non-Standard Analysis, *Koninkl. Wetensch. Proc., Series A*, 64, 1961, pp. 22 - 49.
- [10] ROSENFELD L., René-François de Sluse et le problème des tangentes, *ISIS*, 10, 1928, pp. 416 - 434.
- [11] TALL D., Looking at graphs through infinitesimal microscopes, windows and telescopes, *Mathematical Gazette*, 64, 1980, pp. 22 - 49.

Chapitre 9

Introduction du cercle osculateur au moyen de microscopes

Pour l'essentiel, ce texte est une traduction libre de l'article récemment paru en anglais, sous le titre « *Osculating Circle With Microscopes Within Microscopes* » [5].

Résumé. Le cercle osculateur en un point d'une courbe plane est classiquement présenté de manière assez technique, souvent au moyen de formules donnant les coordonnées du centre et le rayon dudit cercle. Dans cet article est proposée une présentation nouvelle et intuitive de ce concept, développée dans une approche non standard de l'analyse : parmi tous les cercles qui ont, au point considéré, la même tangente que la courbe étudiée et qui dès lors possèdent la même image que la courbe dans l'oculaire d'un microscope infiniment puissant dirigé vers le point considéré, le cercle osculateur est celui qui coïncide avec la courbe dans l'oculaire d'un microscope de microscope au sens de MAGNANI et DOSSONA.

Mots-clés : cercle osculateur, microscope de microscope, analyse non standard.

9.1 Introduction

Depuis très longtemps, les courbes ont joué un rôle capital dans la construction des mathématiques, notamment parce qu'elles décrivent le mouvement que peut avoir un corps, ramené à un point géométrique (par exemple, son centre de gravité), soumis à diverses forces. Leur étude fut en effet décisive dans la création de concepts mathématiques aussi importants que celui de

fonction et, plus généralement, dans le développement du calcul différentiel et intégral. Nous restreindrons la présente étude au cas des courbes planes.

Les plus grands mathématiciens se sont efforcés de caractériser comment une courbe diffère d'une droite qui l'approxime au mieux (i.e. sa tangente au point considéré) ; à la suite des travaux du grec APOLLONIUS (262 - 180 BC), ils ont à cet effet introduit le concept de *courbure*. Citons trois apports majeurs dans cette direction :

- Le mathématicien et philosophe allemand Gottfried Wilhelm LEIBNIZ¹ (1646 - 1716) est le créateur de nombreux concepts fondamentaux de l'analyse. Il a ainsi défini la courbure du graphe d'une fonction à partir de la dérivée seconde tout en tenant compte de la longueur de l'arc. Il a également distingué, parmi les cercles "touchant" une courbe donnée (à savoir des cercles qui possèdent même tangente que la courbe et qui sont appelés en latin *circulo curvam propositam*), le cercle "baisant" (en latin *circulum osculans*²) : parmi tous les cercles touchant la courbe, il s'agit de celui qui épouse celle-ci au mieux ; ce cercle est désormais connu sous le nom de "cercle osculateur".
- L'anglais NEWTON (1642-1727) fut l'un des plus grands scientifiques que le monde a connu³. Dans son œuvre *Principia* (1687), il s'intéressa à la courbure ; il proposa notamment un test pour reconnaître quand deux courbes possèdent même courbure, puis étudia la courbure de courbes particulières, notamment un cercle, une section conique, une cardioïde, ... (*Section I, Lemma XI*). Il donna également une construction géométrique de ce qui est désormais appelé le centre et le rayon du cercle osculateur ; en fait, il résolut le problème suivant : « étant connue la vitesse que possède un corps en trois points d'une orbite connue, décrite par la vitesse sous l'action de forces convergeant vers un centre commun, trouver ce centre » (*Section II, Proposition V, Problem I*).
- Le mathématicien et baron français Augustin Louis CAUCHY (1789 - 1857) peut être considéré comme un des premiers à présenter rigoureusement l'analyse mathématique⁴, dans son cours d'analyse [7] dispensé à l'École polytechnique de Paris. Il caractérisa également le centre du cercle osculateur en un point d'une courbe comme étant l'intersection de deux normales infiniment proches (cité dans [6]).

Cette note est divisée en deux parties.

1. Il est quelquefois désigné sous le nom de VON LEIBNIZ ou encore de LEIBNITZ.

2. Ce terme apparaît chez LEIBNIZ, *Meditatio nova de natura anguli contactus et osculi, Acta Eruditorum*, Juin 1686, in Gerhardt, *Mathematische*, tome VII, p. 326-329.

3. NEWTON fut à la fois alchimiste, astronome, mathématicien, philosophe, physicien et théologien. Sur le plan mathématique, il se disputa avec LEIBNIZ la paternité du Calcul différentiel et intégral. Les historiens contemporains s'accordent à penser que les deux savants peuvent être considérés comme les pères spirituels de l'analyse mathématique, leurs approches des infinitésimaux étant assez différentes l'une de l'autre.

4. En particulier, il donna une définition rigoureuse de la convergence d'une série.

Tout d'abord sont rappelés des définitions et des résultats classiques relatifs aux concepts de courbure et de cercle osculateur.

Ensuite est présentée une version originale du cercle osculateur, recourant à l'usage d'un microscope, et plus précisément d'un microscope de microscope. Cette méthode permet de repérer des détails infinitésimaux au voisinage d'un point et ainsi de "voir" un cercle osculateur.

9.2 Présentations classiques

On considère une courbe régulière plane

$$\gamma :]a, b[\mapsto \mathbb{R}^2 \text{ avec } \gamma(t) = (x(t), y(t))$$

9.2.1 Courbure

De manière à mesurer en quelque sorte de combien γ diffère, au voisinage d'un des ses points (correspondant au paramètre t), de sa tangente est introduite la *courbure* (orientée) : elle est notée $\kappa_2[\gamma](t)$ et est définie comme suit ([9], p. 14) :

$$\kappa_2[\gamma](t) = \frac{\gamma''(t) \cdot J\gamma'(t)}{\|\gamma'(t)\|^3}$$

où J désigne la rotation de $\frac{\pi}{2}$ dans le sens anti-horlogique ; en termes équivalents, $J : \mathbb{R}^2 \mapsto \mathbb{R}^2$ est la transformation linéaire définie par $J(p_1, p_2) = (-p_2, p_1)$.

Une des premières formules caractérisant la courbure fut celle-ci donnée par NEWTON :

$$\kappa_2[\gamma](t) = \frac{x'(t)y''(t) - x''(t)y'(t)}{\left[(x'(t))^2 + (y'(t))^2\right]^{\frac{3}{2}}}$$

En corollaire, si γ est le graphe d'une fonction réelle f , alors

$$\kappa_2[\gamma](t) = \frac{f''(x)}{\left[1 + (f'(x))^2\right]^{\frac{3}{2}}}$$

9.2.2 Une première définition

Le concept de courbure permet de trouver le cercle qui approche "au mieux" la courbe au voisinage d'un de ses points. Ce cercle, appelé *cercle osculateur*, est défini comme suit au point $\gamma(t)$ tel que $a < t < b$ et $\kappa_2[\gamma](t) \neq 0$ ([9], p. 111) : son rayon est égal à

$$R(t) = \frac{1}{|\kappa_2[\gamma](t)|}$$

et son centre est le point

$$\gamma(t) + R(t) \frac{J\gamma'(t)}{\|\gamma'(t)\|}$$

Il s'agit donc du cercle dont le centre se trouve sur la normale (c'est-à-dire la perpendiculaire à la tangente) à une distance égale à l'inverse, en valeur absolue, de la courbure.

9.2.3 Une deuxième définition

La définition précédente est assez technique et abstraite. Il est possible de donner une version plus géométrique du concept en prouvant que le cercle osculateur est la limite de cercles passant par trois points de la courbe considérée lorsque ces points tendent vers le point de contact envisagé. Plus formellement, soit $C(t_1, t_2, t_3)$ le cercle passant par trois points $\gamma(t_1)$, $\gamma(t_2)$ et $\gamma(t_3)$ pour autant que ceux-ci soient distincts et ne soient pas situés en ligne droite; si $\kappa_2[\gamma](t_0) \neq 0$, alors le cercle osculateur γ au point $\gamma(t_0)$ est le cercle

$$C = \lim_{\substack{t_1 \rightarrow t_0 \\ t_2 \rightarrow t_0 \\ t_3 \rightarrow t_0}} C(t_1, t_2, t_3)$$

([9], p. 112).

De fait, soit f une fonction au moins deux fois dérivable en un point a et telle que $f''(a) \neq 0$. La limite des cercles \mathcal{C}_n passant par les trois points

$$A = (a, f(a)) \quad , \quad A_n = \left(a - \frac{1}{n}, f\left(a - \frac{1}{n}\right) \right) \quad , \quad B_n = \left(a + \frac{1}{n}, f\left(a + \frac{1}{n}\right) \right)$$

est le cercle dont le centre est le point

$$M_0 = \left(a - \frac{f'(a)}{f''(a)} [1 + (f'(a))^2], f(a) + \frac{1 + (f'(a))^2}{f''(a)} \right)$$

et dont le rayon est égal à

$$\frac{(1 + (f'(a))^2)^{\frac{3}{2}}}{|f''(a)|}$$

Il s'agit du cercle osculateur au graphe de f en a .

Pour prouver ce résultat, posons

$$b = f(a) \quad , \quad c_n = f\left(a - \frac{1}{n}\right) \quad \text{et} \quad d_n = f\left(a + \frac{1}{n}\right)$$

Le cercle \mathcal{C}_n passant par A , A_n et B_n possède pour équation

$$(x - x_n)^2 + (y - y_n)^2 = R_n^2$$

où x_n et y_n désignent respectivement l'abscisse et l'ordonnée du centre, tandis que R_n est le rayon.

En prenant la limite de la suite des \mathcal{C}_n lorsque $n \rightarrow +\infty$, il vient en désignant par $\bar{x} = \lim_{n \rightarrow +\infty} x_n$, $\bar{y} = \lim_{n \rightarrow +\infty} y_n$ et $\bar{R} = \lim_{n \rightarrow +\infty} R_n$:

$$(x - \bar{x})^2 + (y - \bar{y})^2 = \bar{R}^2$$

Or, le point (x_n, y_n) se trouve à l'intersection des médiatrices des segments $[A, A_n]$ et $[A, B_n]$; en d'autres termes, x_n et y_n vérifient ces deux égalités

$$\begin{cases} y_n - \frac{b+c_n}{2} = \frac{1}{n(c_n-b)} \left(x_n - a + \frac{1}{2n} \right) \\ y_n - \frac{b+d_n}{2} = \frac{1}{n(b-d_n)} \left(x_n - a - \frac{1}{2n} \right) \end{cases}$$

Des calculs élémentaires livrent

$$x_n = a + \frac{1}{2b - d_n - c_n} \left[n \left(\frac{d_n - c_n}{2} \right) (c_n - b) (b - d_n) - \frac{1}{2n} (c_n - d_n) \right]$$

Un passage à la limite pour $n \rightarrow +\infty$ dans l'expression ci-dessus donne

$$\bar{x} = a - \frac{f'(a)}{f''(a)} \left(1 + (f'(a))^2 \right)$$

d'où l'on déduit

$$\bar{y} = f(a) + \frac{1 + (f'(a))^2}{f''(a)} \quad \text{et} \quad \bar{R}^2 = \frac{(1 + (f'(a))^2)^3}{(f''(a))^2}$$

ce qui permet de conclure.

9.2.4 Une troisième définition

Il est encore possible de prouver que le cercle osculateur de γ , avec $a < t_0 < b$ et $\kappa_2[\gamma](t_0) \neq 0$, est l'unique cercle qui a au moins un ordre 2 de contact avec γ en $\gamma(t_0)$ ([9], p. 228), sachant que le concept de contact peut être défini comme suit ([9], p. 225). Soit $F : \mathbb{R}^2 \mapsto \mathbb{R}$ une fonction différentiable; on dit que γ et la courbe définie implicitement⁵ par F ont un contact d'ordre n en $\gamma(t_0)$ quand on a

$$(F \circ \gamma)(t_0) = (F \circ \gamma)'(t_0) = \dots = (F \circ \gamma)^{(n)}(t_0) = 0, \quad (F \circ \gamma)^{(n+1)}(t_0) \neq 0$$

5. On appelle *courbe définie implicitement* par une fonction $F : \mathbb{R}^2 \mapsto \mathbb{R}$ l'ensemble des zéros de F ; on note parfois cet ensemble de zéros $F^{-1}(0)$ ([9], p. 225).

9.3 Une autre approche : pourquoi, comment ?

Les présentations classiques sont généralement assez techniques ; ainsi, on peut utiliser des formules pour trouver le rayon et le centre d'un cercle osculateur, mais ces calculs ne permettent pas réellement de “voir” ce que l'on cherche.

Nous allons proposer une méthode qui permet de distinguer, parmi tous les cercles touchant la courbe (c'est-à-dire les cercles qui possèdent au point considéré la même tangente que la courbe) celui qui “épouse” le mieux cette courbe initiale.

Illustrons intuitivement cette méthode sur un exemple élémentaire.

9.3.1 Un premier exemple

Dans le plan, on considère la parabole \mathcal{P} définie par $y = x^2$ et son sommet $O = (0, 0)$; sa tangente \mathcal{T} en O est l'axe des x . Le cercle osculateur correspondant est caractérisé par son rayon égal à $\frac{1}{2}$ et son centre qui est le point $(0, \frac{1}{2})$.

La figure 9.1 ci-dessous montre comment sont perçues, à l'œil nu, la parabole \mathcal{P} , le cercle osculateur \mathcal{S} en caractères gras, ainsi que deux autres cercles tangents \mathcal{C}' et \mathcal{C}'' en pointillés.

Si l'on regarde toutes ces figures au moyen d'un zoom suffisamment puissant orienté vers le point O , la parabole et les trois cercles considérés ci-dessus semblent indiscernables.

Par contre, si l'on utilise un zoom encore plus puissant pointé sur un point voisin du point de contact, on semble voir la parabole et les trois cercles comme des segments de droite horizontaux, donc parallèles à la tangente \mathcal{T} , le cercle osculateur paraissant confondu avec \mathcal{P} , alors que les deux cercles \mathcal{C}' et \mathcal{C}'' sont situés de part et d'autre de la courbe ainsi qu'en atteste la figure 9.2.

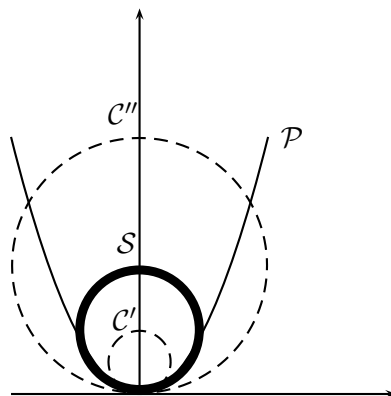


FIGURE 9.1 – Parabole et cercles tangents à l'œil nu

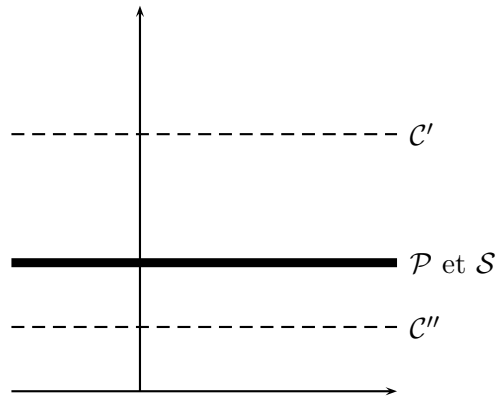


FIGURE 9.2 – Vision avec un microscope

Dans la section suivante, nous allons formaliser la méthode suggérée par cet exemple. Auparavant, il nous faut rappeler le cadre au sein duquel nous allons travailler (à savoir celui des nombres hyperréels), ainsi que les outils auxquels nous allons recourir : il s’agit respectivement de l’ensemble des nombres hyperréels et de leur algèbre, ainsi que des concepts de microscope infiniment puissant et de microscope de microscope.

9.3.2 Nombres hyperréels

Bien que des savants des siècles précédents, comme FERMAT, NEWTON, LEIBNIZ, ..., aient déjà travaillé avec des “infinitésimaux”, ces nombres ont été introduits rigoureusement en 1961 par A. ROBINSON, qui en a fait le cadre de l’approche désormais qualifiée de “non standard” pour l’analyse mathématique. Pour fixer le vocabulaire et les notations adoptés ultérieurement, rappelons brièvement les premiers éléments de cette théorie dans une présentation à la fois simple et didactique, initiée notamment par KEISLER ([3], [10]).

Tout d’abord, nous précisons que les nombres hyperréels forment une extension de l’ensemble des réels, avec les mêmes lois algébriques ; techniquement, l’ensemble ${}^*\mathbb{R}$ des nombres hyperréels est un corps totalement ordonné et non archimédien qui est une extension algébrique du corps \mathbb{R} des réels. Au surplus, ${}^*\mathbb{R}$ contient au moins un *infinitésimal*, c’est-à-dire un nombre ε non nul dont la valeur absolue est plus petite que tout réel positif (ou, formellement, $0 < |\varepsilon| < r, \forall r \in \mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$) ; on convient de plus que 0 est infinitésimal. Il existe une infinité d’infinitésimaux non nuls ; l’inverse de chacun d’eux, soit $H = \frac{1}{\varepsilon}$, est *infinitement grand* ou, simplement, *infini*, c’est-à-dire supérieur en valeur absolue à tout nombre réel positif (ou, formellement, $|H| > r, \forall r \in \mathbb{R}$)⁶. Il est clair que les infinitésimaux non nuls et les infiniment

6. Le symbole utilisé par KEISLER pour désigner un infiniment grand est la première lettre H du mot anglais « huge ».

grands ne sont pas des réels. Un nombre hyperréel x qui n'est pas infini est naturellement appelé *fini*; dans ce cas, il existe un et un seul nombre réel r qui est *infinitement proche* de x , en ce sens que la différence $x - r$ est un infinitésimal; r est appelé la *partie standard* de x : celle-ci est désignée par $r = \text{st}(x)$; formellement, st est un homomorphisme d'anneau (de l'ensemble des nombres hyperréels finis dans l'ensemble \mathbb{R} des réels) dont le noyau est l'ensemble des infinitésimaux. De plus, chaque fonction à une ou plusieurs variables réelles possède une *extension naturelle* au sein des nombres hyperréels, avec la même définition et les mêmes propriétés que celles de la fonction initiale : ultérieurement, une fonction et son extension naturelle seront notées de la même manière.

9.3.3 Microscope

Le concept de microscope (virtuel) est relativement bien connu (voir, par exemple, [1], [3], [8], [14]).

Pour un point $P = (a, b)$ du plan hyperréel ${}^*\mathbb{R}^2$ et un nombre infiniment grand positif H , un H -microscope *pointé* vers P et avec H comme *puissance* agrandit les distances depuis P par un facteur H ; un tel microscope sera noté dans la suite \mathcal{M}_H^P : techniquement, il agit comme une application μ définie sur ${}^*\mathbb{R}^2$ comme suit

$$\mu : (x, y) \mapsto \mu(x, y) = (H(x - a), H(y - b))$$

Soit f une fonction à deux variables réelles x et y . Nous allons considérer d'abord dans le plan usuel \mathbb{R}^2 , la courbe \mathcal{C} définie par

$$f(x, y) = 0$$

On suppose que $P = (r, s)$ est un point situé sur \mathcal{C} et que f est de classe C^p , pour p suffisamment grand, dans un voisinage de P . Pour alléger les notations, nous désignons par d_1, d_2, d_{11}, d_{12} et d_{22} les dérivées partielles correspondantes de f en P , i.e. $d_1 = f_x(r, s)$, $d_2 = f_y(r, s)$, $d_{11} = f_{xx}(r, s)$, $d_{12} = f_{xy}(r, s)$ et $d_{22} = f_{yy}(r, s)$. De plus, nous admettons que $d_2 \neq 0$.

Quand on regarde \mathcal{C} dans l'oculaire d'un H -microscope \mathcal{M}_H^P , où H désigne un hyperréel infiniment grand positif arbitraire, on fait appel à l'application $\mu_1 : (x, y) \mapsto ({}^*X, {}^*Y)$ avec

$${}^*X = H(x - r) \Leftrightarrow x = r + \frac{{}^*X}{H} \text{ et } {}^*Y = H(y - s) \Leftrightarrow y = s + \frac{{}^*Y}{H}$$

De la sorte, la courbe observée est, avec les nouvelles coordonnées *X et *Y , d'équation

$$f\left(r + \frac{{}^*X}{H}, s + \frac{{}^*Y}{H}\right) = 0$$

En appliquant le développement de TAYLOR de f en P , on obtient moyennant des calculs élémentaires

$$f(r, s) + d_1 \frac{{}^*X}{H} + d_2 \frac{{}^*Y}{H} + \frac{\varepsilon}{H} = 0$$

où ε désigne un nombre infinitésimal. Après multiplication par H , on prend la partie standard des deux membres de la nouvelle égalité pour obtenir l'image réellement observée dans l'oculaire du microscope ; en notant X et Y au lieu de $\text{st}({}^*X)$ et $\text{st}({}^*Y)$ respectivement, on trouve

$$d_1 X + d_2 Y = 0 \iff Y = -\frac{d_1}{d_2} X$$

Le coefficient $m = -\frac{d_1}{d_2}$ est la pente de la tangente \mathcal{T} à \mathcal{C} en P . En faisant appel aux variables de départ x et y , une équation de \mathcal{T} est donc donnée par

$$y - s = m(x - r)$$

9.3.4 Microscope de microscope

Nous souhaitons distinguer la courbe \mathcal{C} de sa tangente \mathcal{T} . A cet effet, il convient de faire appel à un microscope beaucoup plus puissant que celui utilisé jusqu'à présent, par exemple de puissance H^2 ; nous allons donc utiliser ce que l'on appelle un "microscope de microscope" ([8], [12]). Mais il nous faut diriger ce H^2 -microscope non plus vers P (sinon on y verrait \mathcal{C} et \mathcal{T} toujours égales), mais vers un autre point : nous choisirons non pas un point de la courbe comme on le trouve dans la littérature ([8], [12]), mais bien vers le $P_1 = (r + \frac{1}{H}, s + \frac{m}{H})$ qui est infiniment proche de P et qui appartient à \mathcal{T} ; de la sorte, nous travaillons avec un microscope $\mathcal{M}_{H^2}^{P_1}$; nous pourrions également suivre un même raisonnement avec un point $P_2 = (r - \frac{1}{H}, s - \frac{m}{H})$ et faire appel au microscope $\mathcal{M}_{H^2}^{P_2}$, mais nous ne développons ici que le premier de ces deux cas. Ainsi, nous recourons à une nouvelle application μ_2 de ${}^*R^2$ vers ${}^*R^2$ définie par

$$\mu_2 : (x, y) \mapsto ({}^*\mathcal{X}, {}^*\mathcal{Y})$$

avec

$${}^*\mathcal{X} = H^2 \left(x - r - \frac{1}{H} \right) \text{ et } {}^*\mathcal{Y} = H^2 \left(y - s - \frac{m}{H} \right)$$

Nous travaillons alors avec les nouvelles coordonnées ${}^*\mathcal{X}$ and ${}^*\mathcal{Y}$. L'image de \mathcal{T} est toujours la droite passant par l'origine et de pente égale à m ; pour la courbe \mathcal{C} , on obtient

$$\begin{aligned} 0 &= f\left(r + \frac{1}{H} + \frac{{}^*\mathcal{X}}{H^2}, s + \frac{m}{H} + \frac{{}^*\mathcal{Y}}{H^2}\right) \\ &= f(r, s) + d_1 \left(\frac{1}{H} + \frac{{}^*\mathcal{X}}{H^2}\right) + d_2 \left(\frac{m}{H} + \frac{{}^*\mathcal{Y}}{H^2}\right) + \\ &\quad \frac{1}{2} \left[d_{11} \left(\frac{1}{H} + \frac{{}^*\mathcal{X}}{H^2}\right)^2 + 2d_{12} \left(\frac{1}{H} + \frac{{}^*\mathcal{X}}{H^2}\right) \left(\frac{m}{H} + \frac{{}^*\mathcal{Y}}{H^2}\right) + d_{22} \left(\frac{m}{H} + \frac{{}^*\mathcal{Y}}{H^2}\right)^2 \right] + \frac{\varepsilon_1}{H^2} \end{aligned}$$

où ε_1 désigne à nouveau un infinitésimal.

Au prix de calculs algébriques élémentaires, en appliquant les règles relatives aux parties standards et en écrivant \mathcal{X} , \mathcal{Y} au lieu de $\text{st}(*\mathcal{X})$, $\text{st}(*\mathcal{Y})$ respectivement, nous obtenons

$$d_1 \mathcal{X} + d_2 \mathcal{Y} + \frac{1}{2} [d_{11} + 2 m d_{12} + m^2 d_{22}] = 0$$

ou encore, de façon équivalente

$$\mathcal{Y} = m \mathcal{X} - \frac{1}{2d_2} [d_{11} + 2 m d_{12} + m^2 d_{22}]$$

Il s'agit de la droite parallèle, mais distincte, de l'image de la tangente.

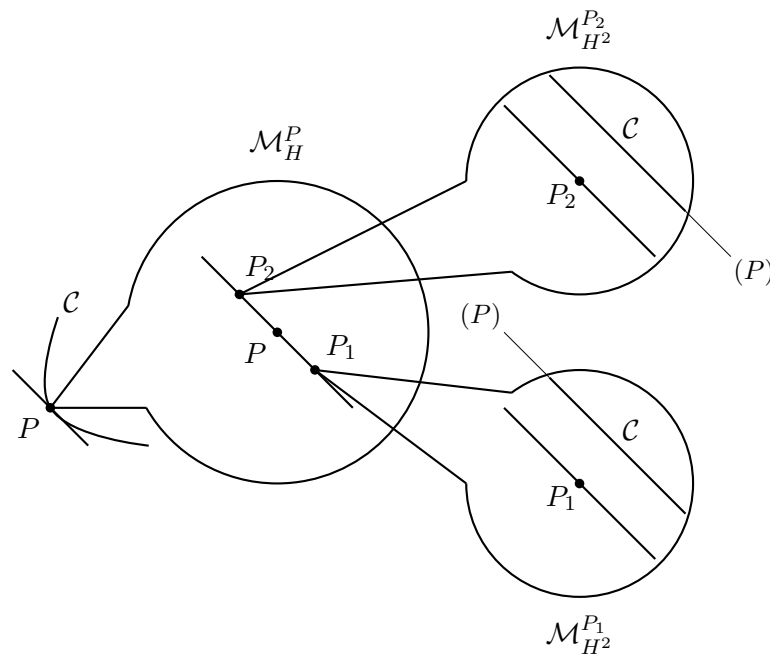


FIGURE 9.3 – Microscope de microscope

Par exemple, pour le cas traité par la figure 9.1, la parabole \mathcal{P} est la droite horizontale $\mathcal{Y} = 1$, alors que la tangente \mathcal{T} au point $P = (0; 0)$ coïncide avec l'axe des abscisses $\mathcal{Y} = 0$.

9.3.5 Cercles tangents

Nous considérons à présent un cercle qui est tangent au point P à la droite \mathcal{T} (celle-ci étant la tangente à la courbe en P); pour plus de simplicité, nous

appelons un tel cercle un *cercle tangent* (sous-entendu à la courbe \mathcal{C} au point P). Son centre est situé sur la normale intérieure définie par

$$y - s = \frac{-1}{m} (x - r)$$

si $m \neq 0$, sur la droite verticale $x = r$ si $m = 0$. Dans la suite, nous allons exclusivement traiter la première éventualité $m \neq 0$, car le cas particulier $m = 0$ est trivial.

Si son rayon vaut R et son centre est le point (a, b) (avec $b \neq s$), alors nous avons

$$R^2 = (a - r)^2 + (b - s)^2 \text{ et } b - s = \frac{-1}{m} (a - r)$$

De plus, son image dans l'oculaire du microscope $\mathcal{M}_{H^2}^{P_1}$ (et le résultat est le même avec $\mathcal{M}_{H^2}^{P_2}$), c'est-à-dire en recourant à l'application μ_2 introduite plus haut, est donnée par

$$\left(r + \frac{1}{H} + \frac{*X}{H^2} - a \right)^2 + \left(s + \frac{m}{H} + \frac{*Y}{H^2} - b \right)^2 = (a - r)^2 + (b - s)^2$$

A l'aide de calculs algébriques élémentaires et des règles sur les parties standards, on obtient aisément

$$2(r - a)X + 2(s - b)Y + 1 + m^2 = 0 \iff Y = -\frac{r - a}{s - b}X - \frac{1 + m^2}{2(s - b)}$$

De façon équivalente, on peut encore écrire

$$Y = mX - \frac{1 + m^2}{2(s - b)}$$

En appliquant cette théorie à la première illustration (voir la figure 9.2), les cercles \mathcal{C}' , \mathcal{S} et \mathcal{C}'' et \mathcal{S} ont respectivement pour équations

$$Y = 2, \quad Y = 1 \quad \text{et} \quad Y = \frac{1}{2}$$

9.4 Nouvelle présentation du cercle osculateur

Avec les outils de la section précédente et à la lumière de notre première illustration, nous pouvons donner cette nouvelle définition.

9.4.1 Une quatrième définition

Par définition, nous appelons le *cercle osculateur* au point P à la courbe \mathcal{C} le seul des cercles tangents qui possède la même image observable que \mathcal{C} par l'application μ_2 ; autrement dit, quand nous regardons la courbe \mathcal{C} et

un cercle tangent dans l'oculaire de $\mathcal{M}_{H^2}^{P_1}$, ce dernier est le cercle osculateur cherché lorsqu'est vue une seule droite.

En comparant deux formules données plus haut, il vient

$$-\frac{1}{2d_2} [d_{11} + 2 m d_{12} + m^2 d_{22}] = -\frac{1 + m^2}{2(s - b)}$$

En conséquence les coordonnées du centre du cercle osculateur sont données par ces égalités

$$b - s = -\frac{(1 + m^2) d_2}{d_{11} + 2 m d_{12} + m^2 d_{22}} \text{ et } a - r = -m (b - s) = m \frac{(1 + m^2) d_2}{d_{11} + 2 m d_{12} + m^2 d_{22}}$$

Le carré du rayon R du cercle osculateur est égal à $(1 + m^2) (s - b)^2$. On trouve donc

$$R = (1 + m^2)^{\frac{3}{2}} \left| \frac{d_2}{d_{11} + 2 m d_{12} + m^2 d_{22}} \right| = \frac{(d_1^2 + d_2^2)^{\frac{3}{2}}}{|d_2^2 d_{11} - 2 d_1 d_2 d_{12} + d_1^2 d_{22}|}$$

Dans le cas particulier, la courbe \mathcal{C} est le graphe d'une fonction f , c'est-à-dire quand on a $y - f(x) = 0$, on obtient alors

$$R = \frac{\left[1 + (f'(x))^2 \right]^{\frac{3}{2}}}{f''(x)}$$

ce qui est équivalent avec la formule livrée dans la première sous-section ci-dessus.

9.4.2 Un deuxième exemple

Voici une seconde illustration très simple et fort parlante.

On considère le point $P = (0, -1)$ sur l'ellipse \mathcal{E} définie par

$$x^2 + 4y^2 = 4$$

La tangente \mathcal{T} à \mathcal{E} en P est la droite horizontale d'équation $y = -1$. En recourant aux notations introduites plus haut, il vient

$$f(x, y) = x^2 + 4y^2 - 4, \quad d_1 = d_{12} = m = 0, \quad d_2 = -8$$

$$d_{11} = 2, \quad d_{22} = 8, \quad R = \frac{|-8|}{2} = 4$$

Une application du microscope de microscope $\mathcal{M}_{H^2}^{P_1}$ avec $P_1 = \left(\frac{1}{H}, -1\right)$ fournit comme image de l'ellipse \mathcal{E} la droite horizontale d'ordonnée $\frac{1}{8}$.

Au surplus, les cercles tangents à \mathcal{E} en P ont leur centre sur l'axe vertical du plan. Plus précisément, notons \mathcal{C}_a le cercle de centre $(0, -1 + a)$ et de rayon a (avec $a > 0$) ; il admet pour équation

$$x^2 + (y + 1 - a)^2 = a^2 \iff x^2 + y^2 + 2(1 - a)y + 1 - 2a = 0$$

L'image de \mathcal{C}_a dans l'oculaire de $\mathcal{M}_{H^2}^{P_1}$ est la droite horizontale d'ordonnée $\frac{1}{2a}$. Donc, le cercle osculateur de \mathcal{E} en P est \mathcal{C}_4 , puisque $2a = 8$.

La situation réelle est visible sur la figure 9.4 ci-dessous : il est clair que, dans un voisinage de P , \mathcal{C}_4 est une bien meilleure approximation circulaire de \mathcal{E} que (notamment) \mathcal{C}_3 ou \mathcal{C}_5 .

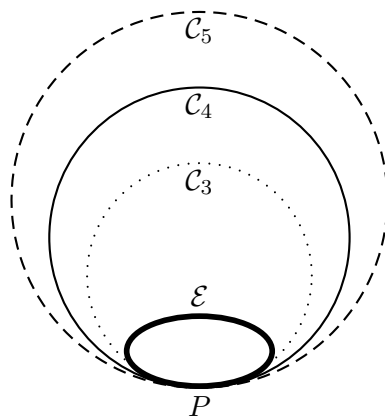


FIGURE 9.4 – Approximation circulaire d'une ellipse

Conclusion

Il est bien connu que les microscopes et les microscopes de microscopes sont d'intéressants “médiateurs épistémiques” ([11], [12]) : ils permettent de découvrir par la pratique (en anglais *discover through doing*) plusieurs concepts importants relatifs aux courbes, tels que ceux de tangente, de concavité (ou convexité), de point d'inflexion ([3]). Nous pouvons désormais ajouter à cette liste les concepts de cercle osculateur et de courbure.

Bibliographie du chapitre 9

- [1] ANTIBI - BAIR J. - HENRY V., Une modélisation d'un zoom au moyen de microscopes virtuels, *Teach. Math. Comput. Sc.*, 2/2, 2004, pp. 319 - 335.
- [2] ANTIBI - BAIR J. - HENRY V., Limites de courbes : théorie et applications en analyse, *Mathématique et Pédagogie*, n° 147, 2004, pp. 65 - 87.
- [3] BAIR J. - HENRY V., *Analyse infinitésimale : le calculus redécouvert*, édition Academia Bruylant, 2008, Louvain-la-Neuve.

- [4] BAIR J. - HENRY V., Implicit Differentiation with microscopes, *The Mathematical Intelligencer*, 31 (1), 2010, pp. 53 - 55.
- [5] BAIR J. - HENRY V., Osculating Circle With Microscopes Within Microscopes, *Foundations of Science*, Volume 18, Number 2, 2013, pp. 319 - 325.
- [6] BOROVNIK A - KATZ M., Who gave you the Cauchy–Weierstrass tale? The dual history of rigorous calculus, *Foundations of Science*, 17 (3), 2012, pp. 245-276.
- [7] CAUCHY A. L., *Leçons sur les applications du calcul infinitésimal à la géométrie*, Imprimerie Royale, Paris, 1826.
- [8] DOSSONA R. - MAGNANI L., Mathematics through Diagrams : Microscopes in Non-Standard and Smooth Analysis, *Studies in Computational Intelligence*, 64, 2007, pp. 193–213.
- [9] GRAY A., *Modern Differential Geometry of Curves and Surfaces with Mathematica*, Second Edition, CRC Press LLC, Boca Raton, Florida, 1998.
- [10] KEISLER H.J., *Elementary Calculus : An Infinitesimal Approach*, On-line Edition, 2000 revised February 2012, <http://www.math.wisc.edu/~keisler/calc.html>.
- [11] MAGNANI L., Epistemic mediators and model-based discovery in sciences, *Model-Based Reasoning : Science, Technology, Values*, L. Magnani and N.J. Nersanian eds, Kluwer Academic / Plenum Publishers, 2002, New York, pp. 305 - 329.
- [12] MAGNANI L. - DOSSONA R., Perceiving the Infinite and the Infinitesimal World Unveiling and Optical Diagrams in Mathematics, *Foundations of Science*, Volume 10, n° 1, 2005, pp. 7 - 23.
- [13] ROBINSON A., *Non-Standard Analysis*, North-Holland, 1996, Amsterdam.
- [14] TALL D., Looking at graphs through infinitesimal microscopes, windows and telescopes, *The Math. Gazette*, 64, 1980, pp. 22–46.

Chapitre 10

Fondement de l'analyse multivariée par une approche non standard

L'analyse mathématique multivariée traite *calculus* (c'est-à-dire des calculs différentiel et intégral) concernant des fonctions à plusieurs variables ; elle est logiquement enseignée après l'analyse univariée qui ne concerne que les fonctions à une seule variable. Cette dernière peut être abordée par une approche non standard en travaillant avec des nombres hyperréels ; une approche pédagogique, inspirée par le livre *Elementary Calculus*, par H. Jerome KEISLER [8], en est donnée dans notre ouvrage *Analyse infinitésimale - le calculus redécouvert* ([3]).

Le texte qui suit constitue une suite au dernier ouvrage cité, dont les définitions et notations sont ici reprises ; il complète un syllabus rédigé naguère dans une présentation classique [2] ; il correspond, de manière très synthétique, à ce qui a été enseigné pendant de nombreuses années lors du cours de mathématiques dispensé en deuxième année d'ingénieur de Gestion à HEC - Ecole de Gestion de l'Université de Liège.

10.1 Espace numérique hyperréel à plusieurs dimensions

On sait que la droite numérique réelle \mathbb{R} peut être prolongée en la droite numérique hyperréelle ${}^*\mathbb{R}$. Les nouveaux nombres hyperréels sont de deux types :

- les *limités* : chaque réel r se voit “entouré” d'hyperréels *infinitement proches* formant le *halo* $H(r)$, ces nouveaux nombres ne pouvant être “vus” qu'à l'aide d'un *microscope virtuel infinitement puissant* pointé sur

le nombre réel r ;

- les *non limités*, qui sont soit des hyperréels *infinitement grands positifs*, soit des hyperréels *infinitement grands négatifs*, ces nouveaux nombres n'étant "observables" qu'au moyen d'un *télescope virtuel infinitement puissant* pointé vers la droite (pour les positifs) ou vers la gauche (pour les négatifs).

Plus généralement, l'espace numérique réel à n dimensions \mathbb{R}^n peut être prolongé en *l'espace numérique hyperréel à n dimensions*. Ce nouvel espace, noté ${}^*\mathbb{R}^n$, est composé de tous les *points hyperréels*, c'est-à-dire de tous les n -uples $({}^*x_1, {}^*x_2, \dots, {}^*x_n)$ dont les n *coordonnées* sont des nombres hyperréels quelconques.

Bien entendu, l'espace réel \mathbb{R}^n est inclus dans l'espace hyperréel ${}^*\mathbb{R}^n$.

Pour deux points arbitraires

$${}^*P = ({}^*x_1, {}^*x_2, \dots, {}^*x_n) \text{ et } {}^*Q = ({}^*y_1, {}^*y_2, \dots, {}^*y_n)$$

de ${}^*\mathbb{R}^n$, on définit une distance, notée $d({}^*P, {}^*Q)$, à l'aide de l'égalité suivante :

$$d({}^*P, {}^*Q) = \sqrt{\sum_{i=1}^n ({}^*x_i - {}^*y_i)^2} ;$$

cette distance généralise bien entendu la distance euclidienne classique utilisée dans \mathbb{R}^n . Les deux points *P et *Q sont dits *infinitement proches* lorsque leur distance $d({}^*P, {}^*Q)$ est un hyperréel infinitement petit, ce qui se note comme suit : ${}^*P \approx {}^*Q$.

La distance de ${}^*P = ({}^*x_1, {}^*x_2, \dots, {}^*x_n)$ à l'origine $O = (0, 0, \dots, 0)$ est la *norme* du point *P : elle est encore notée $\|{}^*P\|$ et vaut donc :

$$\|{}^*P\| = d(O, {}^*P) = \sqrt{\sum_{i=1}^n ({}^*x_i)^2}.$$

Un point *P de ${}^*\mathbb{R}^n$ est soit *limité*, soit *infinitement éloigné* selon que sa norme est respectivement un hyperréel limité ou un infinitement grand positif. En vertu des règles de LEIBNIZ, la première éventualité ne se produit que si et seulement si chaque coordonnée *x_i de *P est un nombre hyperréel limité.

Un point limité ${}^*P = ({}^*x_1, {}^*x_2, \dots, {}^*x_n)$ qui n'appartient pas à \mathbb{R}^n est infinitement proche d'un et d'un seul point $P = (r_1, r_2, \dots, r_n)$ de \mathbb{R}^n : ce point P est appelé la *partie standard* de *P et se note encore $P = \text{st}({}^*P)$ ou bien $P \approx {}^*P$; dans ce cas, pour tout indice i de 1 à n , on a $r_i = \text{st}({}^*x_i)$. Plus généralement, la partie standard d'un ensemble *E de points limités *X de ${}^*\mathbb{R}^n$ est l'ensemble E composé des points $X = \text{st}({}^*X)$ de \mathbb{R}^n lorsque *X parcourt *E .

Tous les points hyperréels infinitement proches du point réel P forment le *halo* $H(P)$ de P . Ils peuvent être "visualisés" au moyen d'un *microscope virtuel*

dirigé sur le point P : un tel microscope, dont la *puissance* est un hyperréel ω infiniment grand positif, agit comme une transformation de ${}^*\mathbb{R}^n$ vers ${}^*\mathbb{R}^n$ définie par

$$({}^*x_1, {}^*x_2, \dots, {}^*x_n) \mapsto ({}^*X_1, {}^*X_2, \dots, {}^*X_n) : {}^*X_i = \omega \times ({}^*x_i - r_i).$$

Les coordonnées initiales *x_i sont encore reliées aux coordonnées correspondantes *X_i (du nouveau repère auquel on se réfère avec le microscope) par cette égalité :

$${}^*x_i = r_i + \frac{{}^*X_i}{\omega} \quad \forall i = 1, \dots, n.$$

Dans la suite, \mathcal{M}_ω^P désignera indistinctement le microscope virtuel de puissance ω dirigé vers le point P et la transformation définie ci-dessus.

Signalons encore que les règles d'extension et de transfert données pour les fonctions à une seule variable s'étendent aux fonctions comptant plusieurs variables ; nous noterons d'ailleurs pareillement une fonction et son extension naturelle.

10.2 Continuité

Soit f une fonction définie sur un sous-ensemble E de \mathbb{R}^n :

$$f : X = (x_1, x_2, \dots, x_n) \in E \mapsto f(x_1, x_2, \dots, x_n) = f(X).$$

On dit que la fonction f est *continue sur* un ensemble A lorsque A est inclus dans E et que, pour tout point X de A , on a

$${}^*X \in H(X) \cap {}^*A \implies f(X) \cong f({}^*X),$$

où *A désigne l'extension naturelle de A ; dans le cas particulier où A se réduit en un point X , on dit alors que la fonction f est continue en X . Notons que la conséquence de l'implication peut encore se mettre sous la forme équivalente

$$f(X) = \text{st}(f({}^*X)).$$

Exemples

– La fonction

$$f : (x, y) \mapsto x$$

est continue sur \mathbb{R}^2 .

– La fonction

$$f : (x, y) \mapsto y$$

est continue sur \mathbb{R}^2 .

– Si a et b désignent deux réels, la fonction linéaire

$$f : (x, y) \mapsto f(x, y) = ax + by$$

est continue sur \mathbb{R}^2 .

– La fonction

$$f : (x, y) \mapsto \frac{xy}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0) \text{ et } f(0, 0) = 0$$

n'est pas continue en l'origine. De fait, si ε désigne un réel infiniment petit, le point $*X = (\varepsilon, 2\varepsilon)$ est infiniment proche de l'origine $(0, 0)$ mais $f(*X) \not\approx 0 = f(0, 0)$.

Proposition 10.2.1 *Si f et g sont deux fonctions continues en un point P de \mathbb{R}^n , alors $|f|$, $f + g$, $f - g$ et $f \times g$ sont continues en P ; il en va de même du quotient $\frac{f}{g}$ lorsque $g(P) \neq 0$.*

Preuve. Il suffit d'appliquer les règles de LEIBNIZ.

Proposition 10.2.2 *Soient h une fonction définie sur un intervalle I de \mathbb{R} et g une fonction définie sur un sous-ensemble A de \mathbb{R}^n et telle que $g(X) \in I$ pour tout $X \in A$. Si h est continue sur I et g est continue sur A , alors la fonction composée*

$$f = h \circ g : X \in A \mapsto h(g(X))$$

est continue sur A .

Preuve. Considérons un point arbitraire X de A et un point $*X$ de $H(X)$. Par la continuité de g , on sait que

$$g(X) \approx g(*X).$$

La continuité de h entraîne alors

$$h(g(X)) \approx h(g(*X)).$$

Remarque. La plupart des fonctions à plusieurs variables courantes sont continues sur leur ensemble de définition, puisqu'elles sont composées de fonctions continues.

10.3 Dérivabilité

Dans cette section, on considèrera une fonction f à n variables qui est définie sur un ouvert D de \mathbb{R}^n ; $P = (r_1, r_2, \dots, r_n)$ sera un point de D , tandis que U_i désignera le i -ème point unitaire, c'est-à-dire le point de \mathbb{R}^n dont toutes les coordonnées sont nulles à l'exclusion de la i -ème égale à 1. Pour un hyperréel ε infiniment petit, on fera appel au quotient

$$\Delta_i f(P; \varepsilon) = \frac{f(P + \varepsilon U_i) - f(P)}{\varepsilon},$$

dans lequel intervient l'extension naturelle de f (également notée f selon notre convention).

Définitions 10.3.1 – Si, quel que soit l'infiniment petit ε , le quotient $\Delta_i f(P; \varepsilon)$ est limité et possède une partie standard qui ne dépend pas de ε , alors le nombre réel

$$\text{st}(\Delta_i f(P; \varepsilon))$$

est appelé la dérivée partielle, par rapport à la i -ème variable, de f au point P et est noté indifféremment

$$\frac{\partial f}{\partial x_i}(P) \text{ ou } f'_i(P).$$

– La fonction f est dite dérivable au point P si les dérivées partielles $f'_i(P)$ existent pour tout i compris entre 1 et n .

Exemples. – La fonction $f : (x, y) \mapsto x^2 y^3$ est dérivable en tout point $P = (r, s)$. En effet, on a

$$\frac{(r + \varepsilon)^2 s^3 - r^2 s^3}{\varepsilon} = 2rs^3 + \varepsilon s^3,$$

d'où l'on tire

$$\frac{\partial f}{\partial x}(P) = 2rs^3;$$

de même,

$$\frac{r^2(s + \varepsilon)^3 - r^2 s^3}{\varepsilon} = r^2(3s^2 + 3s\varepsilon + \varepsilon^2),$$

d'où l'on tire

$$\frac{\partial f}{\partial y}(P) = 3r^2 s^2.$$

– Soit f la fonction définie sur \mathbb{R}^2 par $f(x, y) = 0$ si $x = 0$ ou $y = 0$, et par $f(x, y) = 1$ sinon; f est dérivable en l'origine puisque l'on a

$$\frac{f(\varepsilon, 0) - f(0, 0)}{\varepsilon} = \frac{f(0, \varepsilon) - f(0, 0)}{\varepsilon} = \frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0.$$

Remarquons toutefois que f n'est pas continue en l'origine.

– Soit f la fonction à quatre variables définie par

$$f(x_1, x_2, x_3, x_4) = 2x_1 + x_1^2 + (\sin x_1)|x_2| + |x_3| + (x_4)^{2/3}.$$

On a visiblement

$$\frac{\partial f}{\partial x_1}(0, 0, 0, 0) = 2, \quad \frac{\partial f}{\partial x_2}(0, 0, 0, 0) = 0.$$

Par contre, les dérivées partielles de f par rapport à x_3 et x_4 n'existent pas en l'origine, puisque

$$\frac{f(0, 0, \varepsilon, 0) - f(0, 0, 0, 0)}{\varepsilon} = \frac{|\varepsilon|}{\varepsilon}$$

vaut $+1$ ou -1 selon que ε est positif ou négatif, tandis que

$$\frac{f(0, 0, 0, \varepsilon, 0) - f(0, 0, 0, 0)}{\varepsilon} = \frac{1}{(\varepsilon)^{1/3}}$$

n'est pas limité. La fonction f n'est donc pas dérivable en l'origine de l'espace \mathbb{R}^4 bien qu'elle y soit continue.

Théorème 10.3.2 (Théorème de dérivation totale). *Il s'agit de démontrer la formule suivante :*

$$\frac{dF}{dt} = \frac{dx}{dt} \frac{\partial f}{\partial x} + \frac{dy}{dt} \frac{\partial f}{\partial y},$$

lorsque f est une fonction continûment dérivable en les deux variables x et y qui, elles, sont dérivables en la variable t , avec $F(t) = f(x(t), y(t))$.

Preuve. Nous adoptons les notations suivantes :

$$F(t) = f(x(t), y(t)), \quad x = x(t), \quad y = y(t), \quad x(t+\Delta t) = x+\Delta x, \quad y(t+\Delta t) = y+\Delta y.$$

Dans ces conditions, la différence

$$\Delta F = F(t + \Delta t) - F(t)$$

peut se mettre sous la forme :

$$\begin{aligned} \Delta F &= f(x + \Delta x, y + \Delta y) - f(x, y) \\ &= f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) + f(x, y + \Delta y) - f(x, y) \end{aligned}$$

En posant

$$g(x) = f(x, y + \Delta y) \text{ et } h(y) = f(x, y),$$

on peut encore écrire

$$\Delta F = g(x + \Delta x) - g(x) + h(y + \Delta y) - h(y).$$

En appliquant la formule des accroissement finis aux fonctions univariées g et h , on trouve deux réels Θ_1 et Θ_2 compris entre 0 et 1 tels que

$$\Delta F = \Delta x g'(x + \Theta_1 \Delta x) + \Delta y h'(y + \Theta_2 \Delta y),$$

d'où l'on tire

$$\frac{\Delta F}{\Delta t} = \frac{\Delta x}{\Delta t} g'(x + \Theta_1 \Delta x) + \frac{\Delta y}{\Delta t} h'(y + \Theta_2 \Delta y).$$

En prenant Δt infiniment petit, on peut écrire :

$$\Delta x \cong 0, \quad \Delta y \cong 0, \quad \frac{\Delta x}{\Delta t} \cong x'(t), \quad \frac{\Delta y}{\Delta t} \cong y'(t).$$

De plus, comme les dérivées partielles de f sont supposées continues au point (x, y) , on a encore

$$g'(x + \Theta_1 \Delta x) \cong \frac{\partial f}{\partial x}(x, y) \text{ et } h'(y + \Theta_2 \Delta y) \cong \frac{\partial f}{\partial y}(x, y).$$

En prenant la partie standard du quotient $\frac{\Delta F}{\Delta t}$, on trouve bien la thèse.

Théorème 10.3.3 (Théorème de SCHWARZ - YOUNG). *Lorsque la fonction f est deux fois continûment dérivable en P , on a*

$$\frac{\partial^2 f}{\partial x \partial y}(P) = \frac{\partial^2 f}{\partial y \partial x}(P).$$

Preuve. Considérons l'expression suivante

$$E = f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y) - f(x, y + \Delta y) + f(x, y).$$

En posant

$$g(x) = f(x, y + \Delta y) - f(x, y) \text{ et } h(y) = f(x + \Delta x, y) - f(x, y),$$

on obtient

$$E = g(x + \Delta x) - g(x) = h(y + \Delta y) - h(y).$$

En appliquant la formule des accroissements finis aux fonctions univariées g et h , on trouve deux réels Θ_1 et Θ_2 compris entre 0 et 1 tels que

$$\Delta x g'(x + \Theta_1 \Delta x) = \Delta y h'(y + \Theta_2 \Delta y),$$

ce qui peut encore s'écrire sous la forme équivalente

$$\begin{aligned} & \Delta x \left[\frac{\partial f}{\partial x}(x + \Theta_1 \Delta x, y + \Delta y) - \frac{\partial f}{\partial x}(x + \Theta_1 \Delta x, y) \right] \\ &= \Delta y \left[\frac{\partial f}{\partial y}(x + \Delta x, y + \Theta_2 \Delta y) - \frac{\partial f}{\partial y}(x, y + \Theta_2 \Delta y) \right]. \end{aligned}$$

Une nouvelle application de la formule des accroissements finis livre des réels Θ_3 et Θ_4 encore compris entre 0 et 1 tels que

$$\Delta x \Delta y (f'_x)'_y(x + \Theta_1 \Delta x, y + \Theta_3 \Delta y) = \Delta y \Delta x (f'_y)'_x(x + \Theta_4 \Delta x, y + \Theta_2 \Delta y).$$

On peut simplifier les deux membres de cette égalité par le produit $\Delta x \Delta y$, prendre ces deux nombres Δx et Δy infiniment petits, puis égaliser les parties standards de ces expressions : comme les dérivées partielles secondes f''_{xy} et f''_{yx} sont supposées continues, on en déduit la thèse.

10.4 Courbes algébriques

Considérons une courbe \mathcal{C} d'équation

$$F(x, y) = 0, \text{ avec } F(x, y) = \sum_{i=0}^m \sum_{j=0}^p a_{ij} x^i y^j :$$

il s'agit d'une courbe algébrique de degré égal à

$$d = \sup_{a_{ij} \neq 0} (i + j).$$

La règle de transfert permet d'écrire, au sein de l'espace ${}^*\mathbb{R}^2$:

$$F({}^*x, {}^*y) = 0.$$

Sur un point $P = (r, s)$ quelconque de \mathcal{C} , pointons le microscope virtuel \mathcal{M}_ω^P de puissance ω (avec ω infiniment grand positif). L'image, notée \mathcal{C}_ω , de \mathcal{C} par \mathcal{M}_ω^P est obtenue, en remplaçant *x par $r + \frac{{}^*X}{\omega}$ et *y par $s + \frac{{}^*Y}{\omega}$ dans l'égalité $F({}^*x, {}^*y) = 0$, ce qui donne la courbe algébrique d'équation exprimée dans les nouvelles coordonnées ${}^*X, {}^*Y$:

$$\sum_{i=0}^m \sum_{j=0}^p a_{ij} \left(r + \frac{{}^*X}{\omega}\right)^i \left(s + \frac{{}^*Y}{\omega}\right)^j = 0.$$

Développons les diverses puissances intervenant dans cette égalité en exploitant la formule du binôme de NEWTON. Toutefois, contentons-nous d'écrire les deux premiers termes de chaque développement en regroupant tous les autres dans un reste, matérialisé par le symbole "...", dont tous les termes contiennent un facteur $\frac{1}{\omega^k}$ avec $k \geq 2$; on peut ainsi écrire l'équation de \mathcal{C}_ω :

$$\sum_{i=0}^m \sum_{j=0}^p a_{ij} \left(r^i + ir^{i-1} \frac{{}^*X}{\omega} + \dots\right) \left(s^j + js^{j-1} \frac{{}^*Y}{\omega} + \dots\right) = 0 \quad (1)$$

On constate aisément que les termes indépendants de *X et de *Y se simplifient puisque $P \in \mathcal{C}$ implique que $F(r, s) = 0$. Par ailleurs, dans le produit des deux parenthèses, le terme en ${}^*X{}^*Y$ est rejeté dans le reste indiqué par le signe "...", car il contient le facteur $\frac{1}{\omega^2}$; on aboutit alors à cette équation de \mathcal{C}_ω :

$$m_1 \frac{{}^*X}{\omega} + m_2 \frac{{}^*Y}{\omega} + \dots = 0 \quad (2)$$

où les coefficients m_1 et m_2 ne sont rien d'autre que les dérivées partielles de la fonction polynomiale F au point P , soit

$$m_1 = F'_1(r, s) \text{ et } m_2 = F'_2(r, s).$$

A ce stade, il nous faut distinguer deux cas selon que les deux coefficients m_1 et m_2 sont simultanément nuls ou pas, c'est-à-dire selon que le gradient de F au point P est nul ou ne l'est pas. Dans le premier cas, le point P est qualifié de *singulier* pour la courbe \mathcal{C} et dans le second cas, il est dit *régulier*.

Si le point P est régulier, l'équation de l'image \mathcal{C}_ω (de \mathcal{C} par \mathcal{M}_ω^P) peut s'écrire, après multiplication par le facteur ω des deux membres de l'égalité (2) ci-dessus :

$$m_1 *X + m_2 *Y + \dots = 0,$$

où le reste "... " est composé de termes comprenant tous en dénominateur une puissance naturelle de ω . La partie "observable" de cette image est la courbe dont l'équation est obtenue en prenant la partie standard des deux membres de l'équation ci-dessus, c'est-à-dire, en notant $X = \text{st}(*X)$ et $Y = \text{st}(*Y)$:

$$m_1 X + m_2 Y = 0.$$

En revenant aux variables initiales x et y , on obtient l'équation de la tangente à la courbe \mathcal{C} au point P , soit la droite d'équation

$$m_1(x - r) + m_2(y - s) = 0 :$$

il s'agit d'une tangente horizontale, oblique ou verticale selon que $m_1 = 0$, $m_1 m_2 \neq 0$ ou $m_2 = 0$ respectivement. Remarquons que dans les deux premiers cas, on retrouve bien le coefficient angulaire de la tangente donné par le théorème de dérivation des fonctions implicites, à savoir le nombre

$$-\frac{m_1}{m_2} = -\frac{F'_1(r, s)}{F'_2(r, s)},$$

dans le dernier cas, on obtient l'équation de la tangente alors que la relation implicite $F(x, y) = 0$ ne définit pas une fonction (en la variable x) dérivable au point d'abscisse r .

Lorsque P est un point singulier de \mathcal{C} , c'est-à-dire lorsque $m_1 = m_2 = 0$, il importe de développer davantage les calculs en précisant ce que vaut le reste dans l'égalité (1). Nous nous contenterons ici de traiter le cas où les trois dérivées partielles secondes de F au point P ne sont pas toutes nulles¹. Dans le développement du premier membre de (1), tous les termes non nuls contiennent un facteur en $\frac{1}{\omega^k}$ avec $k \geq 2$. En multipliant les deux membres de (1) par ω^2 , on obtient comme équation de \mathcal{C}_ω :

$$m_{11} *X^2 + 2m_{12} *X *Y + m_{22} *Y^2 + \dots = 0,$$

où les coefficients m_{11} , m_{12} , m_{22} sont égaux à des multiples positifs des dérivées partielles secondes $F''_{11}(r, s)$, $F''_{12}(r, s)$, $F''_{22}(r, s)$ respectivement, et

1. Dans le cas contraire, il suffit de poursuivre les calculs encore plus loin.

où tous les termes rassemblés dans le reste “...” contiennent un facteur en $\frac{1}{\omega^k}$ avec $k \geq 1$. En passant à la partie standard comme ci-dessus, on obtient :

$$m_{11}X^2 + 2m_{12}XY + m_{22}Y^2 = 0,$$

ou encore en les variables x, y originelles :

$$m_{11}(x - r)^2 + 2m_{12}(x - r) \times (y - s) + m_{22}(y - s)^2 = 0.$$

Il s’agit d’une conique dégénérée en un point, en une droite ou en deux droites, selon que le nombre $\Delta = m_{11} \times m_{22} - m_{12}^2$ est respectivement positif, nul ou négatif ; le point P est alors respectivement un point isolé pour \mathcal{C} , un point de rebroussement, ou encore un noeud avec deux tangentes.

Nous allons dégager de ce qui précède une méthode générale permettant de traiter systématiquement le cas d’un point quelconque $P = (r, s)$ de n’importe quelle courbe algébrique \mathcal{C} ; ensuite, nous l’illustrerons par quelques exemples typiques.

Algorithme relatif à l’application d’un microscope virtuel en un point d’une courbe algébrique arbitraire

1. Dans l’équation de \mathcal{C} , remplacer la variable x par $r + \frac{*X}{\omega}$ et la variable y par $s + \frac{*Y}{\omega}$.
2. Développer les calculs par distributivité, en exploitant la formule du binôme de NEWTON et en regroupant entre eux les termes contenant une même puissance de $\frac{1}{\omega}$. Toutefois, ne calculer effectivement que les termes en la plus petite puissance $\frac{1}{\omega^k}$ de $\frac{1}{\omega}$, les autres étant regroupés dans un reste matérialisé par le symbole “...” ; on obtient ainsi une équation de la courbe \mathcal{C}_ω .
3. Multiplier les deux membres de l’équation obtenue par le facteur ω^k , puis en prendre la partie standard. En remplaçant st $(*X)$ par $x - r$ et st $(*Y)$ par $y - s$ dans l’équation obtenue, on a une équation d’une courbe \mathcal{C}_0 dans le repère initial.
4. Interpréter géométriquement les résultats obtenus. Notamment, si $k = 1$, \mathcal{C}_0 est la droite tangente au point P qui est régulier pour \mathcal{C} ; si $k = 2$, le point P est singulier et peut être un noeud, un point de rebroussement ou un point isolé.

Exemple. Appliquons cette méthode à la lemniscate de GERONO \mathcal{C} définie par l’égalité

$$x^4 - 4x^2 + 4y^2 = 0.$$

Pointons tout d’abord un microscope de puissance ω sur le point $P_1 = (1, \frac{\sqrt{3}}{2})$ qui appartient à cette courbe. L’image \mathcal{C}_ω de \mathcal{C} par $\mathcal{M}_\omega^{P_1}$ est donnée par

$$\left(1 + \frac{*X}{\omega}\right)^4 - 4\left(1 + \frac{*X}{\omega}\right)^2 + 4\left(\frac{\sqrt{3}}{2} + \frac{*Y}{\omega}\right)^2 = 0.$$

Développons le premier membre sachant que les termes indépendants vont s'éliminer et que les termes contenant en dénominateur une puissance de ω d'exposant au moins égal à 2 sont regroupés dans le reste "...": il suffit donc de n'écrire que les termes en $\frac{1}{\omega}$, ce qui donne

$$\frac{4 *X}{\omega} - \frac{8 *X}{\omega} + \frac{4\sqrt{3} *Y}{\omega} + \dots = 0.$$

Une simplification par ω et un passage par la partie standard livrent la droite d'équation

$$-X + \sqrt{3}Y = 0;$$

le point P_1 est régulier pour \mathcal{C} : la tangente en ce point est la droite d'équation

$$-(x-1) + \sqrt{3} \left(y - \frac{\sqrt{3}}{2} \right) = 0.$$

Dirigeons à présent un microscope de puissance ω sur le point $P_2 = (0, 0)$. L'image \mathcal{C}_ω de la courbe par le microscope $\mathcal{M}_\omega^{P_2}$ est fournie par l'égalité

$$\frac{*X^4}{\omega^4} - 4 \frac{*X^2}{\omega^2} + 4 \frac{*Y^2}{\omega^2} = 0.$$

Ici, le plus petit exposant k de ω est 2. Partant, le passage par la partie standard fournit, après une simplification par ω^2 , la courbe d'équation

$$X^2 = Y^2.$$

Le point P_2 est singulier pour \mathcal{C} : il s'agit d'un noeud où la courbe possède deux tangentes, à savoir les deux droites définies, dans le repère initial, par

$$y = x \text{ et } y = -x.$$

10.5 Fonctions implicites

Si $F : (x, y) \mapsto F(x, y)$ est une fonction à deux variables, l'équation

$$F(x, y) = 0$$

définit, dans le plan numérique \mathbb{R}^2 , l'ensemble composé de tous les points $P = (x, y)$ dont les coordonnées vérifient l'équation en question (c'est-à-dire tels que $F(x, y) = 0$): cet ensemble est généralement une courbe \mathcal{C} .

Si l'on dirige vers un point $P = (r, s)$ appartenant à cette courbe un microscope \mathcal{M}_ω^P , dont la puissance est un hyperréel infiniment grand positif ω , l'image observée est, comme nous l'avons vu avec les courbes algébriques, le plus souvent une droite passant par le centre de l'oculaire et dès lors d'équation

$$aX + bY = 0.$$

Si cette droite n'est pas verticale, elle est la représentation graphique de la fonction linéaire

$$l : X \mapsto -\frac{a}{b}X.$$

Ainsi, au voisinage du point P , la courbe \mathcal{C} définit une fonction univariée f telle que la partie observable du graphe à l'aide du microscope virtuel \mathcal{M}_ω^P coïncide avec le graphe de la fonction linéaire l . Cette fonction f , qui est dite *définie implicitement* par l'égalité $F(x, y) = 0$, est certes inconnue, mais on sait que son graphe passe par le point P et l'on en connaît le nombre dérivé, puisque la tangente à la courbe \mathcal{C} au point P est la droite obtenue par le microscope, c'est-à-dire la droite d'équation

$$y - s = f'(r)(x - r) \text{ avec } f'(r) = -\frac{a}{b}.$$

De plus, ce nombre dérivé $f'(r)$ peut être calculé à partir des dérivées partielles de la fonction bivariable F calculées au point $P = (r, f(r))$.

Proposition 10.5.1 *Si $\frac{\partial F}{\partial y}(P) \neq 0$, alors*

$$f'(r) = -\frac{\frac{\partial F}{\partial x}(P)}{\frac{\partial F}{\partial y}(P)}.$$

Preuve. Il suffit de dériver en l'abscisse r la fonction univariée définie par

$$g(x) = F(x, f(x)).$$

En tenant compte des égalités $s = f(r)$ et $F(x, f(x)) = 0$, on peut écrire, grâce au théorème de dérivation totale :

$$g'(r) = \frac{\partial F}{\partial x}(P) + \frac{\partial F}{\partial y}(P)f'(r) = 0,$$

d'où la thèse.

Bien plus, on peut souvent calculer le nombre $f''(r)$ qui va caractériser la "concavité" de la courbe \mathcal{C} au voisinage du point P . On dispose en effet de ce résultat exprimé avec les notations ci-dessus.

Proposition 10.5.2 *Si F est une fonction deux fois dérivable en P et telle que $\frac{\partial F}{\partial y}(P) \neq 0$, alors*

$$f''(r) = \frac{|\mathcal{H}\bar{F}(P)|}{\left(\frac{\partial F}{\partial y}(P)\right)^3},$$

où $|\mathcal{H}\bar{F}(P)|$ désigne le déterminant hessien bordé calculé au point P .

Preuve. En ayant recours à un microscope virtuel \mathcal{M}_ω^P , de puissance infiniment grande ω , l'égalité

$$F(*x, *y) = 0,$$

déduite des règles d'extension et de transfert, permet d'écrire

$$F\left(r + \frac{*X}{\omega}, s + \frac{*Y}{\omega}\right) = 0.$$

Le théorème de TAYLOR appliqué à la fonction bivariable F livre, après simplification par ω et en supprimant les symboles $*$:

$$F'_1 X + F'_2 Y + \frac{1}{2\omega} [F''_{11} X^2 + 2F''_{12} XY + F''_{22} Y^2] + \dots = 0,$$

où toutes les dérivées partielles sont calculées au point P , et où le signe "... " indique la présence en dénominateur d'une puissance de ω supérieure à 1. On peut en tirer

$$Y = -\frac{F'_1}{F'_2} X - \frac{1}{2\omega F'_2} [F''_{11} X^2 + 2F''_{12} XY + F''_{22} Y^2] + \dots$$

En remplaçant dans le second membre Y par $-\frac{F'_1}{F'_2} X + \dots$, il vient :

$$Y = -\frac{F'_1}{F'_2} X - \frac{X^2}{2\omega F'_2} \left[F''_{11} + 2F''_{12} \frac{-F'_1}{F'_2} + F''_{22} \left(\frac{-F'_1}{F'_2} \right)^2 \right] + \dots;$$

Le nombre cherché $f''(r)$ est le coefficient de $\frac{X^2}{2\omega}$ dans ce second membre, c'est-à-dire le nombre

$$\frac{1}{F'_2} \left[-F''_{11} + 2F''_{12} \frac{F'_1}{F'_2} - F''_{22} \left(\frac{F'_1}{F'_2} \right)^2 \right],$$

d'où la thèse après quelques calculs algébriques élémentaires.

10.6 Différentiabilité

Nous allons formaliser, dans un cadre général, ce que nous avons trouvé sur des cas particuliers.

Soit f une fonction à n variables définie sur un ouvert de \mathbb{R}^n contenant le point $P = (r_1, r_2, \dots, r_n)$.

Le graphe \mathcal{G} de f est d'équation dans \mathbb{R}^{n+1}

$$x_{n+1} = f(x_1, x_2, \dots, x_n).$$

Par les règles d'extension et de transfert, le graphe ${}^*\mathcal{G}$ de l'extension naturelle de f est d'équation dans ${}^*\mathbb{R}^{n+1}$

$${}^*x_{n+1} = f({}^*x_1, {}^*x_2, \dots, {}^*x_n).$$

L'image de ${}^*\mathcal{G}$ par $\mathcal{M}_\omega^{\bar{P}}$, où \bar{P} désigne le point $(r_1, r_2, \dots, r_n, f(P))$ de ${}^*\mathbb{R}^{n+1}$ tandis que ω est un infiniment grand positif, est d'équation

$$f(P) + \frac{{}^*X_{n+1}}{\omega} = f\left(r_1 + \frac{{}^*X_1}{\omega}, r_2 + \frac{{}^*X_2}{\omega}, \dots, r_n + \frac{{}^*X_n}{\omega}\right).$$

La *partie standard* de cette image (restreinte évidemment aux points limités) est un ensemble de \mathbb{R}^n qui est noté $G(f; \bar{P}; \omega)$.

Définition 10.6.1 *La fonction f est dite différentiable au point P lorsque, pour tout hyperréel ω infiniment grand positif, l'ensemble $G(f; \bar{P}; \omega)$ est un hyperplan de \mathbb{R}^{n+1} d'équation*

$$X_{n+1} = \sum_{j=1}^n m_j X_j,$$

m_j désignant un réel pour tout indice j de 1 à n .

Dans ce cas, la fonction linéaire

$$l : (h_1, h_2, \dots, h_n) \mapsto \sum_{j=1}^n m_j h_j$$

est appelée la *différentielle totale*, ou simplement la *différentielle de f en P* .

En remarquant que chaque *X_j dans l'oculaire du microscope correspond, dans le repère initial, à un *x_j infiniment proche de r_j , on remplace fréquemment $X_j = h_j$ dans la fonction linéaire l par dx_j : en effet, dans la pratique, dx_j représente souvent une variation infinitésimale de la variable x_j ; en conséquence, on note encore la différentielle totale de f en P sous la forme

$$df_P(dx_1, dx_2, \dots, dx_n) = \sum_{j=1}^n m_j dx_j$$

ou encore plus simplement quand aucune confusion n'est à craindre

$$df = \sum_{j=1}^n m_j dx_j.$$

Ainsi, lorsque f est différentiable en P , si $\Delta P = (dx_1, dx_2, \dots, dx_n)$ appartient au halo de l'origine de \mathbb{R}^n , on constate que la variation

$$\Delta f = f(P + \Delta P) - f(P)$$

est infiniment proche de la valeur correspondante de la différentielle totale :

$$\Delta f \cong df.$$

Exemples. – Soient la fonction $f : (x, y) \mapsto x^2 + y^2$ et $P = (1, 1)$; comme $f(P) = 2$, le point $\bar{P} = (1, 1, 2)$ est situé sur le graphe de f . Pointons sur \bar{P} un microscope virtuel de puissance ω infiniment grande : l'image observable du graphe de f par $\mathcal{M}_\omega^{\bar{P}}$ est d'équation

$$2 + \frac{{}^*Z}{\omega} = \left(1 + \frac{{}^*X}{\omega}\right)^2 + \left(1 + \frac{{}^*Y}{\omega}\right)^2.$$

En prenant la partie standard des deux membres de cette égalité transformée par quelques calculs algébriques élémentaires, on trouve, en posant $X = \text{st}({}^*X)$, $Y = \text{st}({}^*Y)$ et $Z = \text{st}({}^*Z)$, l'image observée d'équation :

$$Z = 2X + 2Y.$$

La fonction f est donc différentiable en P et sa différentielle totale y est donnée par

$$df = 2dx + 2dy.$$

– La fonction f à quatre variables définie par

$$f(x_1, x_2, x_3, x_4) = 2x_1 - 3x_2 + x_3 - 2x_4 + x_1^2 \left(\sqrt{|\sin x_2|} - |x_3| \times \sqrt{x_4^2} \right)$$

est différentiable en l'origine de l'espace \mathbb{R}^4 à 4 dimensions.

En effet, un microscope $\mathcal{M}_\omega^{\bar{P}}$, où \bar{P} désigne l'origine de l'espace \mathbb{R}^5 et ω est un infiniment grand positif, permet d'écrire, avec des notations habituelles

$$\frac{{}^*X_5}{\omega} = 2\frac{{}^*X_1}{\omega} - 3\frac{{}^*X_2}{\omega} + \frac{{}^*X_3}{\omega} - 2\frac{{}^*X_4}{\omega} + \frac{{}^*X_1^2}{\omega^2} \left(\sqrt{\left| \sin \frac{{}^*X_2}{\omega} \right|} - \left| \frac{{}^*X_3}{\omega} \right| \times \sqrt{\frac{{}^*X_4^2}{\omega^2}} \right),$$

d'où on tire en faisant appel à la partie standard :

$$X_5 = 2X_1 - 3X_2 + X_3 - 2X_4.$$

La différentielle de f en l'origine de \mathbb{R}^4 est donc donnée par

$$df = 2dx_1 - 3dx_2 + dx_3 - 2dx_4.$$

Proposition 10.6.2 *Si f est différentiable en P , alors f est continue en P .*

Preuve. Avec les notations exposées ci-dessus, on peut écrire

$$\Delta f = f(P + \Delta P) - f(P) \approx \sum_{j=1}^n m_j dx_j,$$

où m_j et dx_j désignent respectivement des réels et des hyperréels infiniment petits pour tout indice j .

En vertu des règles de LEIBNIZ, l'expression $\sum_{j=1}^n m_j dx_j$ est également un infiniment petit. Il en résulte que

$$f(P + \Delta P) \cong f(P).$$

De ce dernier résultat, on peut en déduire que, pour une même fonction à plusieurs variables et en un même point, les notions de dérivabilité et de différentiabilité ne coïncident pas : en effet, il est possible de construire des fonctions multivariées² dérivables mais non continues, donc forcément non différentiables. Néanmoins, les deux énoncés suivants vont préciser les liens existant entre ces deux notions.

Proposition 10.6.3 *Si f est différentiable en P , alors f est dérivable en P .*

Preuve. Fixons un indice i entre 1 et n ; soit $\Delta P = (dx_1, dx_2, \dots, dx_n)$, avec $dx_i = \varepsilon$ infiniment petit et $dx_j = 0$ pour tout indice j différent de i . L'image observable du graphe de f donnée par le microscope virtuel $\mathcal{M}_\omega^{\bar{P}}$, avec $\bar{P} = (r_1, r_2, \dots, r_n, f(P))$ et $\omega = \frac{1}{|\varepsilon|}$ est donnée par l'équation

$$X_{n+1} = \sum_{j=1}^n m_j X_j,$$

pour des réels m_j . Il existe donc un α infiniment petit tel que

$$\omega \times [f(P + \Delta P) - f(P)] = \sum_{j=1}^n m_j \times (x_j - r_j) \times \omega + \alpha = m_i \times \omega \times \varepsilon + \alpha.$$

On en déduit

$$\frac{f(P + \Delta P) - f(P)}{\varepsilon} = m_i + \frac{\alpha}{\omega \times \varepsilon}.$$

Comme $\omega \times \varepsilon$ est appréciable, un passage à la partie standard conduit à cette égalité définissant la partie observée du graphe de f :

$$\text{st} \left(\frac{f(P + \Delta P) - f(P)}{\varepsilon} \right) = m_i,$$

d'où la thèse

$$\frac{\partial f}{\partial x_i}(P) = m_i.$$

2. Rappelons que les deux notions de dérivabilité et de différentiabilité coïncident dans le cas d'une fonction à une seule variable.

Remarque. La preuve de ce résultat permet de réécrire la différentielle totale de f sous la forme suivante qui est la plus souvent trouvée dans la littérature³

$$df = \sum_{j=1}^n \frac{\partial f}{\partial x_j} dx_j,$$

la différentielle et toutes les dérivées partielles étant prises au point P .

Proposition 10.6.4 *Si f est dérivable en P et si ses dérivées partielles premières sont également continues en P , alors f est différentiable en P .*

Preuve. Au moyen du microscope $\mathcal{M}_\omega^{\bar{P}}$, où $\bar{P} = (r_1, r_2, \dots, r_n, f(P))$ et ω est un infiniment grand positif, le graphe de f conduit à cette égalité

$$f(P) + \frac{{}^*X_{n+1}}{\omega} = f\left(r_1 + \frac{{}^*X_1}{\omega}, r_2 + \frac{{}^*X_2}{\omega}, \dots, r_n + \frac{{}^*X_n}{\omega}\right).$$

En appliquant la formule des accroissements finis et en effectuant quelques calculs algébriques élémentaires, on trouve un hyperréel α infiniment petit tel que

$${}^*X_{n+1} = \sum_{j=1}^n \left({}^*X_j \times \frac{\partial f}{\partial x_j}(P) \right) + \alpha.$$

Un passage à la partie standard permet de conclure.

Bibliographie du chapitre 10

- [1] ARCHINARD G. - GUERRIEN B., *Analyse mathématique pour économistes*, 3^{ème} édition, Economica, 1988, Paris.
- [2] BAIR J., *Analyse mathématique*, Editions de l'Université de Liège, 2003 - 2004.
- [3] BAIR J. - HENRY V., *Analyse infinitésimale : le calculus redécouvert*, édition Academia Bruylant, 2008, Louvain-la-Neuve.
- [4] BAIR J. - HENRY V., Implicit Differentiation with microscopes, *The Mathematical Intelligencer*, 31 (1), 2010, pp. 53 - 55.
- [5] HAESBROECK G. - HENRY V., *Exercices résolus d'analyse mathématique multivariée*, Editions de l'Université de Liège, 2003 - 2004.
- [6] HENRY V., Une proposition d'ingénierie pour définir la différentielle, CD-Rom du Colloque international du CREM, Mons, 2005.
- [7] HENRY V., An Introduction to Differentials with Hyperreal Numbers and Infinite Microscopes, *PRIMUS (Problems Resources, and Issues in Mathematics Undergraduate Studies)*, 20(1), 2010, pp. 39-49.

3. Cette présentation de la différentielle est toutefois dangereuse dans la mesure où elle n'indique pas le point P considéré, et où elle ne met pas en évidence le fait que la différentielle est en réalité une fonction linéaire.

- [8] KEISLER H. J., *Elementary Calculus*, Prindle, Weber & Schmidt, 1966, Boston.
- [9] STEWART J., *Analyse : concepts et contextes. Volume 2. Fonctions de plusieurs variables*, De Boeck Université, 2001, Bruxelles.

Chapitre 11

Diverses présentations du concept de différentielle

Résumé. Dans un cours universitaire de mathématiques destiné à de futurs économistes ou gestionnaires, le concept de différentielle paraît incontournable aussi bien d'un point de vue théorique que pour les applications pratiques.

Après une courte introduction historique sur le sujet, nous montrons, au départ d'une assez large analyse de la littérature spécialisée, que ce concept peut être présenté de multiples manières et qu'il semble exister un "décalage interdisciplinaire" entre certains enseignants selon qu'ils sont mathématiciens ou économistes. Cette étude est reprise du chapitre 12 de la thèse [17].

Nous proposons une présentation nouvelle, basée sur l'emploi de microscopes infiniment puissants utilisés dans notre approche de l'analyse non standard. Elle a été testée auprès de nombreuses cohortes d'étudiants inscrits dans des études d'ingénieur de gestion à l'Université de Liège.

11.1 Introduction

La notion de *différentielle* est due à LEIBNIZ. Selon DIDEROT et D'ALEMBERT dans l'*Encyclopédie*

« On appelle dans la haute Géométrie "quantité différentielle" ou simplement "différentielle" une quantité infiniment petite, ou moindre que toute grandeur assignable. On l'appelle différentielle ou quantité différentielle parce qu'on la considère ordinairement comme la différence infiniment petite de deux quantités finies, dont l'une surpasse l'autre infiniment peu. (...) Cette méthode est une des plus belles et des plus fécondes de toutes les Mathématiques; M. Leibniz qui l'a publiée le premier, l'appelle "calcul différentiel", en considérant les grandeurs infiniment petites comme les différences

des quantités finies ; c'est pourquoi il les exprime par la lettre d qu'il met au-devant¹ de la quantité différenciée ; ainsi la différentielle de x est exprimée par dx , celle de y par dy , etc. » (cité dans [13], p. 129)

Dès le début de son histoire, elle est apparue assez particulière. En effet, une lecture des textes originaux *Nova Methodus pro Maximis et Minimis* révèle que son créateur s'est contenté d'en énoncer les principales propriétés, pratiquement sans justification théorique, mais en a proposé diverses applications géométriques. En fait, comme l'a écrit M. PARMENTIER,

« LEIBNIZ ne s'embarrasse guère de discussions métaphysiques sur les fondements de son calcul. (...) Il recourt à la très approximative analogie entre les différents ordres d'infiniment petits et les "incomparables" que sont un grain de poussière au regard du diamètre terrestre. (...) L'analogie avec les incomparables se présente sous sa plume comme un commode artifice pédagogique, à seule fin de mieux faire comprendre, non les fondements métaphysiques du calcul, mais son usage, plus exactement son maniement. » ([20], pp. 34-35)

De nos jours encore, cette notion de différentielle n'a pas toujours une signification claire ; toujours selon M. PARMENTIER,

« Les différentielles sont des êtres amphibies à deux faces, dont l'avvers est quantitatif, mais le revers est la face aveugle qu'elles présentent aux règles et aux opérations du calcul. » ([20], p. 39)

Paradoxalement, cette polysémie ne l'empêche pas d'être très efficace dans des applications nombreuses et variées, tant d'un point de vue local (petites variations de diverses grandeurs) que d'un point de vue global (mise en équations de problèmes non linéaires menant à des intégrales ou à des équations différentielles) (voir [1], p. 57).

Ce double statut de la différentielle, à savoir la faiblesse de sens et la puissance de l'outil, se retrouve constamment dans la formation mathématique des futurs économistes ou gestionnaires. La notion est souvent présentée fort succinctement par les mathématiciens, qui semblent attacher beaucoup plus d'importance à la notion de dérivée, la différentielle apparaissant alors, d'un point de vue théorique, comme un "sous-produit" de la dérivée. Par contre, les économistes sont des grands utilisateurs de différentielles et, quelquefois, ne font guère de distinction entre les concepts de dérivabilité et de différentiabilité, et dès lors entre ceux de dérivée et de différentielle. Ce décalage, que nous qua-

1. La lettre d , pour désigner une différentielle, est apparue la première fois dans un manuscrit daté du 11 octobre 1675 ; notons que, dans un premier temps, LEIBNIZ écrivait cette lettre d sous la variable et notait par exemple $\frac{x}{a}$ ce que l'on note aujourd'hui dx ([20], p. 100).

lifions d'« interdisciplinaire », entre l'utilisation du même concept dans deux contextes différents peut se révéler être la source d'obstacles pour l'apprenant.

11.2 Présentation de la différentielle dans les mathématiques pour économistes

Une analyse de la façon dont la notion de différentielle est introduite dans différents livres, en langue française, destinés à des étudiants entreprenant des études universitaires en économie ou en gestion révèle une grande variété de présentations.

Néanmoins, que ce soit pour des fonctions à une variable ou pour des fonctions à plusieurs variables réelles, les auteurs consultés mentionnent un certain nombre de points communs, les textes différant alors essentiellement par le choix des points privilégiés ainsi que par l'ordre dans lequel ceux-ci sont détaillés.

Nous dressons ci-dessous la liste de ces points communs. Chacun de ceux-ci est présenté, de manière succincte, souvent en référence avec un ou plusieurs des ouvrages sélectionnés. Ensuite, nous relevons, pour chacun des points, le nombre d'ouvrages où il apparaît et nous terminons en détaillant, pour chaque référence, les points traités ainsi que leur ordre d'apparition.

Remarquons d'emblée qu'il s'avérera souvent opportun de distinguer les cas de fonctions à une seule variable réelle de ceux des fonctions à plusieurs variables réelles ; pour clarifier l'exposé, nous noterons systématiquement dans cette section f une fonction d'une seule variable et F une fonction à au moins deux variables

11.2.1 Dix points rencontrés dans la littérature

– **(1) : Hypothèse de dérivabilité**

Dans le cas d'une fonction f d'une seule variable, elle est fréquemment supposée dérivable au point x étudié ; on peut alors écrire

$$f(x + \Delta x) - f(x) = f'(x)\Delta x + \varepsilon(x)\Delta x$$

pour une fonction $\varepsilon(\Delta x)$ tendant vers 0 avec Δx ([14], p. 126). Une fonction F à plusieurs variables est quelquefois supposée dérivable au point considéré, avec de plus des dérivées partielles continues ([24], p. 62) ; parfois F est supposée dérivable au point considéré mais dans toutes les directions ([15]).

– **(2) : Définition de la différentiabilité**

– Une fonction f est dite *différentiable* en un point x_0 s'il existe un réel $\delta > 0$ tel que

$$|\Delta x| < \delta \implies \Delta f = K\Delta x + o(x_0, \Delta x),$$

où $\Delta f = f(x_0 + \Delta x) - f(x_0)$ et

(i) K est une constante dont la valeur dépend de x_0 (mais évidemment pas de Δx)

(ii) $o(\dots)$ est une fonction de x_0 et de Δx satisfaisant la condition suivante :

$$\lim_{\Delta x \rightarrow 0} \frac{o(x_0, \Delta x)}{\Delta x} = 0.$$

([18], p. 64).

- Une fonction F à deux variables est différentiable en (x_0, y_0) s'il existe un voisinage V de (x_0, y_0) tel que, pour tous $\Delta x, \Delta y$ pour lesquels $(x_0 + \Delta x, y_0 + \Delta y) \in V \cap D$ on ait

$$\Delta F = A_1 \Delta x + A_2 \Delta y + \omega_1 \Delta x + \omega_2 \Delta y$$

où

(i) A_1 et A_2 sont indépendants de Δx et Δy mais peuvent dépendre de x_0 et y_0 ;

(ii) ω_1 et ω_2 peuvent dépendre de $x_0, y_0, \Delta x$ et Δy et de plus

$$\lim_{(\Delta x, \Delta y) \rightarrow (0,0)} \omega_1(x_0, y_0, \Delta x, \Delta y) = 0$$

$$\lim_{(\Delta x, \Delta y) \rightarrow (0,0)} \omega_2(x_0, y_0, \Delta x, \Delta y) = 0$$

([18], p. 382). Cette version peut évidemment être adaptée au cas d'un nombre quelconque de variables.

– **(3) : La dérivée comme quotient de différentielles**

La dérivée d'une fonction f d'une variable en un point x vaut

$$f'(x) = \frac{df(x)}{dx}.$$

On en déduit que

$$df = f'(x)dx.$$

[26], p. 106).

– **(4) : Définition en tant que fonction linéaire à coefficients connus**

Si la fonction f , à une seule variable, est dérivable en x_0 , on appelle différentielle de f en x_0 la fonction linéaire en la variable réelle h définie par

$$df(x_0) : \mathbb{R} \mapsto \mathbb{R} : h \mapsto f'(x_0)h$$

([11], p. 198). De même pour une fonction F à n variables, la différentielle de F en un point P_0 est la fonction linéaire

$$dF(P_0) : \mathbb{R}^n \mapsto \mathbb{R} : (h_1, h_2, \dots, h_n) \mapsto \sum_{i=1}^n \frac{\partial F(P_0)}{\partial x_i} h_i.$$

- **(5) : Définition en tant que fonction linéaire à coefficients inconnus**

On appelle différentielle d'une fonction f (resp. F) en un point x_0 (resp. M_0) toute fonction linéaire tangente à f (resp. F) en x_0 (resp. M_0). ([7], p. 401), sachant que deux fonctions ϕ et φ sont dites tangentes en M_0 si le rapport

$$\frac{\phi(M) - \varphi(M)}{\|M - M_0\|}$$

tend vers 0 en même temps que $\|MM_0\|$ ([7], p. 152).

- **(6) : Définition en tant que fonction comprenant un nombre double de variables**

Partant de l'observation que la différentielle df dépend à la fois du point auquel on se place et des accroissements considérés ([18], p. 382), on se place sur un ensemble A de points en lesquels f est dérivable; on peut alors définir la différentielle $df(x)$ comme suit :

$$df(x) : A \times \mathbb{R} : (x, h) \mapsto f'(x)h$$

([11], p. 199). On pourrait procéder de même pour une fonction F à plusieurs variables.

- **(7) : Fonctions particulières**

À une variable, plusieurs auteurs traitent le cas particulier où la fonction f est la fonction identité; pour des fonctions à plusieurs variables, ils envisagent pour F les fonctions projections sur les axes de coordonnées ([2]; [3], pp. 122 et 271; [11], p. 199; [21], p. 262 et [27]).

- **(8) : Formule**

La différentielle est très fréquemment donnée par la formule utilisant la (les) dérivée(s) (partielles).

Pour une variable x , on a

$$df(x_0) = f'(x_0)dx \quad ([26], p.106).$$

Pour deux variables x et y , on peut écrire

$$dF = F'_x dx + F'_y dy \quad ([26], p.109).$$

La généralisation à n variables livre

$$dF = \sum_{i=1}^n \frac{\partial F}{\partial x_i} dx_i.$$

- **(9) : Approximation**

La différentielle est aussi présentée comme un moyen d'approximation des valeurs de la fonction au voisinage d'un point.

On peut dire que df est la variation de la fonction entraînée par une variation infiniment petite de la variable ou des variables ([26], pp. 106-109).

On en déduit une approximation affine pour les valeurs de la fonction elle-même : ainsi, si F est une fonction de \mathbb{R}^2 dans \mathbb{R} , différentiable en (a, b) , alors l'image par F d'un point $(a + h, b + k)$ voisin de (a, b) est approximée par ([9], p. 189) :

$$F(a + h, b + k) \approx F(a, b) + \frac{\partial F}{\partial x}(a, b)h + \frac{\partial F}{\partial y}(a, b)k.$$

– **(10) : Interprétation graphique**

A une variable, une interprétation graphique de la différentielle est régulièrement présentée à partir de la tangente au graphe de la fonction au point considéré.

De même, on peut faire référence au plan tangent en un point du graphe pour une fonction à deux variables.

11.2.2 Quelques réflexions didactiques

Diverses constatations peuvent être émises à la lecture des points qui précèdent.

- Au point (1), on suppose les fonctions dérivables à une variable, tandis que, dans le cas multivarié, les fonctions sont supposées non seulement dérivables mais avec des dérivées partielles continues. Cette différence du cadre référentiel peut provoquer des erreurs par “entraînement analogique” ([25], p. 107) en oubliant, dans le cas multivarié, l’hypothèse additionnelle relative à la dérivabilité.
- La présentation du point (2) est inutilement compliquée. Le quantificateur existentiel ($\exists \delta$ ou $\exists V$) ainsi que l’implication résultante ne sont d’aucune utilité en raison de la présence de la limite (ou des deux limites) ultérieure(s). La présentation peut être simplifiée : par exemple, une fonction F à deux variables est différentiable en (a, b) lorsqu’il existe deux réels c et d tels que

$$\lim_{(x,y) \rightarrow (a,b)} \frac{F(x, y) - [F(a, b) + c(x - a) + d(y - b)]}{\sqrt{(x - a)^2 + (y - b)^2}} = 0$$

([9], p. 182).

- La différentielle de f en x se note quelquefois df_x : on a donc

$$df_x(\Delta x) = f'(x)\Delta x$$

([3], p. 122). Dans cette notation, le symbole d se réfère à la différentielle, la lettre f indique la fonction dont on prend cette différentielle, tandis que l’indice x rappelle que la différentielle de f est prise au point x donné.

« Il s’agit donc d’une application linéaire de Δx et non de x .
Ce dernier point est extrêmement important : le lecteur doit

toujours le garder en tête, car c'est probablement là que se situe la difficulté à propos de la définition de différentielle. » ([3], p. 122)

- Il est peut-être surprenant de constater qu'aucun livre ne signale qu'une dérivée partielle d'une fonction peut s'exprimer comme un quotient de deux "différentielles" : par exemple, pour la fonction $(x, y) \mapsto F(x, y)$, on peut écrire, sous réserve d'existence :

$$\frac{\partial F}{\partial x}(x_0, y_0) = \frac{dF(x_0, y_0)|_{dy=0}}{dx}$$

où le numérateur du second membre désigne la valeur de la différentielle de F en (x_0, y_0) pour un accroissement dx de l'abscisse avec une variation nulle de l'ordonnée. Pourtant, les économistes raisonnent souvent de la sorte, en ne faisant varier qu'une seule des variables indépendantes, rappelant cette démarche par la phrase "et toutes choses égales par ailleurs" (ou "ceteris paribus" en latin, ou encore l'acronyme TACRE signifiant « toute autre chose restant égale »).

- Les différents points trouvés dans la littérature donnent trois représentations différentes pour la notion de différentielle.
 - a) Selon la première conception (point (4) ou (5)), une différentielle est une fonction linéaire introduite pour approximer au mieux la fonction étudiée par une autre plus simple ([7], p. 401).
 - b) La deuxième version (point (6)) peut difficilement être interprétée intuitivement puisqu'on associe, à une fonction à n variables, une nouvelle fonction à $2n$ variables. Son introduction permet néanmoins de justifier la formule (8) grâce au point (7) ; de fait, à une variable, en appliquant (6) à la fonction x elle-même, on trouve

$$dx : \mathbb{R}^2 \mapsto \mathbb{R} : (x, h) \mapsto x'h = h;$$

ainsi, on peut remplacer h par dx dans (4).

- c) Enfin, la troisième conception (point (8)) ne se situe pas dans le cadre fonctionnel, mais bien dans le cadre numérique puisqu'une différentielle apparaît alors comme une quantité. A ce propos, relevons une remarque émise dans un cours d'analyse mathématique pour économistes :

Il existe une, mauvaise, habitude chez les économistes consistant à se servir de la formule $dy = \sum_i f'_{x_i}(X)dx_i$ (appelée souvent différentielle totale alors que c'est la différentielle "tout court") en considérant dy et les dx_i comme des "infinitement petits", ce qui n'a pas de sens et ce qui ne permet pas de comprendre le fait que la différentielle est une "application linéaire tangente" et qu'elle sert à "linéariser" une fonction au voisinage d'un point ([3], p. 271).

Nous présenterons ultérieurement une nouvelle approche du concept de différentielle qui semble répondre (au moins partiellement) à ces objections.

- A propos du point (10), relevons une nouvelle différence entre les cas univarié et multivarié. Alors que la notion de différentielle d'une fonction à une seule variable est presque toujours illustrée par une figure, il est rare de retrouver une illustration graphique pour les fonctions à deux variables, le cas de plus de deux variables ne pouvant évidemment pas être représenté par un dessin élémentaire. La visualisation dans l'espace est certes plus délicate, mais peut souvent être aisément réalisée de façon suggestive grâce à la technologie moderne.

11.2.3 Occurrence des différents points

Nous avons consulté 22 ouvrages présentant l'analyse mathématique à de futurs économistes ou gestionnaires; 2 de ces livres ne traitent pas les différentielles de fonctions à une seule variable, bien qu'ils développent le calcul différentiel dans ce cadre, mais présentent cette notion pour des fonctions de plusieurs variables. Par contre, 7 auteurs font le contraire : ils introduisent donc la notion de différentielle uniquement pour les fonctions d'une seule variable et pas pour celles à plusieurs variables pour lesquelles ils développent toutefois le calcul différentiel. Dans le premier cas, on peut imaginer que les rédacteurs des manuels n'ont pas estimé utile d'introduire la notion de différentiabilité parce qu'elle équivaut alors à celle de dérivabilité; ils privent toutefois leurs lecteurs d'un outil fréquemment exploité par les économistes. Dans le second cas, nous pensons que les auteurs ont voulu, par souci de simplification, éviter de présenter la notion délicate de différentiabilité qui diffère ici de celle de dérivabilité, cette dernière apparaissant incontournable en vue des applications économiques.

Voici un tableau reprenant les pourcentages (arrondis à l'unité) d'occurrence des 10 points introduits ci-dessus. La première ligne du tableau comprend les numéros des différents points; la deuxième (resp. la troisième) ligne se réfère aux fonctions à une seule variable (resp. à n variables, avec $n \geq 2$) et concerne donc 20 (resp. 15) des 22 ouvrages consultés.

Points	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$n = 1$	70	15	35	70	0	10	35	80	60	65
$n > 1$	33	53	0	47	27	7	33	87	73	47

Relevons quelques caractéristiques des données précédentes.

- Pour la plupart des points, les résultats diffèrent sensiblement selon que l'on travaille avec des fonctions à une seule variable ou des fonctions à plusieurs variables.

- La définition de la différentiabilité (2) n'est guère donnée, surtout pour les fonctions univariées.
- Même lorsque la notion de différentiabilité est introduite, la définition (4) de la différentielle est préférée à (5).
- Deux tiers des auteurs consultés n'envisagent les cas particuliers (7) de la fonction identité à une variable et des projections à plusieurs variables. Ils ne justifient dès lors pas la formule (8) qu'ils donnent pourtant majoritairement.
- La différentielle n'est pas toujours définie comme étant une fonction : même en cumulant les pourcentages relatifs aux points (4), (5) ou (6), le total de 100 % n'est pas atteint ; plus précisément, le caractère fonctionnel de la différentielle est présenté dans 80 % (resp. 73 %) des cas pour les fonctions univariées (resp. multivariées).
- L'interprétation géométrique de la différentielle n'est donnée que par 65 % (resp. 46 %) des auteurs dans le cas d'une variable (resp. de plusieurs variables) ; et pourtant, les graphiques représentent généralement ce qui subsiste le mieux et le plus longtemps dans la mémoire des apprenants.

11.2.4 Ordre d'apparition des différents points

Il nous a semblé intéressant d'analyser succinctement les points abordés par un même auteur, ainsi que leur ordre de présentation. Le tableau suivant synthétise ces données ; les numéros se réfèrent à ceux des points introduits ci-dessus, tandis que les ouvrages recensés sont indiqués par leur référence se rapportant à la bibliographie située en fin de ce texte ; les symboles *** signalent que cette matière n'est pas développée dans la référence.

Auteurs	1 variable	plusieurs variables
[2]	(1)-(4)-(10)-(7)-(8)	(2)-(1)-(4)-(9)-(10)-(7)-(8)
[3]	(1)-(4)-(9)-(10)	***
[5]	(1)-(10)-(4)-(8)-(9)	(2)-(5)-(4)-(8)-(9)-(10)
[6]	(3)-(8)	***
[7]	(1)-(4)-(8)-(3)	(2)-(5)-(9)-(8)
[8]	(1)-(4)-(10)-(9)	***
[9]	***	(2)-(10)-(4)-(8)
[10]	(3)-(8)	(9)-(8)
[11]	(1)-(4)-(10)-(6)-(7)-(8)	***
[12]	***	(9)-(8)
[14]	(1)-(4)-(9)-(10)	***
[15]	(2)-(4)-(9)	(2)-(4)-(1)-(9)
[16]	(3)-(8)-(9)-(10)	(8)-(9)
[18]	(2)-(1)-(4)-(7)-(8)-(6)-(10)-(9)	(2)-(1)-(4)-(6)-(7)-(8)-(10)-(9)

Auteurs	1 variable	plusieurs variables
[19]	(1)-(4)-(7)-(8)-(10)-(9)	***
[21]	(1)-(4)-(7)-(3)-(8)	(2)-(5)-(4)-(8)-(10)
[22]	(4)-(7)-(8)-(9)-(10)	(4)-(7)-(8)
[23]	(2)-(1)-(10)-(4)-(8)-(9)	(1)-(2)-(5)-(4)-(7)-(8)-(9)
[24]	(1)-(8)-(10)-(9)	(1)-(8)
[26]	(3)-(8)-(9)	(9)-(8)
[27]	(1)-(9)-(8)-(10)-(3)	(9)-(10)-(8)-(7)

Epinglons quelques constatations issues de ces données.

- Les mêmes points peuvent apparaître dans un ordre différent d’une référence à l’autre.
- Tous les auteurs n’adoptent pas la même présentation pour les cas univarié ou multivarié.
- Les livres écrits par des économistes ([10], [16], [26]) présentent le concept de différentielle essentiellement à l’aide des formules (3) et (8), c’est-à-dire en ne considérant que le cadre numérique. Par contre, les ouvrages rédigés par des mathématiciens se placent principalement dans le cadre fonctionnel. Cette différence de présentation met en évidence le “décalage” pouvant exister entre les utilisateurs des mathématiques et les mathématiciens.
- Plusieurs auteurs ayant introduit la notion de différentiabilité pour des fonctions à plusieurs variables en déduisent la dérivabilité, ce qui leur permet de donner la définition (4) au lieu de (5).
- Les auteurs qui considèrent le point (1) ne mentionnent pas nécessairement le point (5), ce qui est naturel. Par contre, ceux qui reprennent le point (2) considèrent souvent (5) avant (4).
- Les points (9) et (10) surviennent fréquemment après les définitions (4) ou (5), alors que ces interprétations numérique et graphique pourraient servir de situation favorisant l’introduction d’une définition.
- Les auteurs qui, au sein d’un chapitre consacré à l’analyse multivariée, ne traitent pas de la notion de différentielle pour des fonctions à plusieurs variables ont tous introduit la différentielle pour une fonction univariée par la séquence (1)-(4). On peut émettre la supposition qu’ils n’ont pas traité le cas multivarié parce que le schéma unidimensionnel ne peut plus alors être appliqué.

11.3 Nouvelle approche de la différentielle

Le point de départ consiste à remarquer que l’image du graphe d’une fonction dans l’oculaire d’un microscope infiniment puissant dirigé vers un point de ce graphe apparaît généralement comme étant une droite dans le cas univarié, et un plan dans le cas bivarié; de plus, cette image contient chaque fois

l'origine du nouveau repère auquel se réfère le microscope. Analytiquement, cette visualisation provient donc d'une fonction linéaire ; celle-ci possède dès lors pratiquement les mêmes valeurs que la fonction de départ aux points vus dans le microscope, ces derniers étant infiniment proches du point de référence.

11.3.1 Définitions

On considère ici une fonction f définie sur une partie D de l'espace \mathbb{R}^n , un point $P = (r_1, r_2, \dots, r_n)$ intérieur à D et $\bar{P} = (r_1, r_2, \dots, r_n, r_{n+1})$ un point de \mathbb{R}^{n+1} avec $r_{n+1} = f(P)$.

Définition 11.3.1 *f est différentiable en P lorsque, quel que soit l'hypperréel ω infiniment grand positif, la partie standard du graphe de f dans l'oculaire du microscope \mathcal{M}_ω^P est le graphe d'une fonction linéaire, c'est-à-dire d'une fonction définie par une égalité du type suivant :*

$$X_{n+1} = \sum_{j=1}^n m_j X_j.$$

Pour une fonction différentiable, on introduit la notion correspondante de différentielle.

Définition 11.3.2 *Si f est différentiable en P , l'ensemble de \mathbb{R}^{n+1} défini par l'équation*

$$X_{n+1} = \sum_{j=1}^n m_j X_j$$

et vu dans l'oculaire du microscope \mathcal{M}_ω^P est le graphe d'une fonction linéaire appelée la différentielle de f en P .

Ces définitions doivent bien entendu être illustrées par des exemples variés, puis commentées ; ensuite, les propriétés classiques des différentielles peuvent en être déduites.

11.3.2 Réflexions didactiques

Relevons divers points qui, au vu de nos réflexions et de nos expériences, nous semblent représenter, par rapport à une présentation plus classique, des avantages pour les étudiants.

- Il est possible de “voir concrètement” la différentielle en un point grâce à un microscope infiniment puissant. La *vision géométrique* ([4], p. 19) est ainsi bien présente ; elle peut être rendue plus concrète pour des fonctions à une ou à deux variables en recourant à un logiciel adéquat et en effectuant alors des zooms successifs.

- Ces définitions insistent sur l’aspect fonctionnel ([4], p. 49) du concept de différentielle.
- Ce que l’on voit de la fonction différentielle concerne des points infiniment proches du point de référence. Cette remarque est à relier à l’idée que *l’approximation différentielle ne présente d’intérêt que localement* [4], p. 2) ; elle est également conforme avec son usage par des praticiens pour qui une différentielle désigne essentiellement une quantité infiniment petite.
- Les calculs sont aisés, tout du moins pour des fonctions algébriques ; ils peuvent alors être menés sans recourir à des calculs de dérivées (éventuellement partielles).
- Une même définition vaut pour toutes les fonctions, qu’elles soient à une ou à plusieurs variables. Au contraire des définitions classiques, qui prennent en charge des fonctions dérivables à une variable mais des fonctions continûment dérivables à plusieurs variables, cette nouvelle présentation peut être formulée sans hypothèse préalable sur les fonctions.

Bibliographie du chapitre 11

- [1] ALIBERT D. - ARTIGUE M., *Mise au point théorique sur les différentielles*, Annexe du dossier *Questionnaire de travail sur les différentielles*, Irem et Université de Paris VII, 76, 1989.
- [2] ARCHINARD G. - GUERRIEN B., *Analyse mathématique pour économistes*, 3^{ème} édition, Economica, Paris, 1988.
- [3] ARCHINARD G. - GUERRIEN B., *Principes mathématiques pour économistes*, Economica, Paris, 1992.
- [4] ARTIGUE M. - MENIGAUX J. - VIENNOT L., Some aspects of students conceptions and difficulties about differentials, *European Journal of Physics*, n° 11, 1990, pp. 262 - 267.
- [5] BAIR J., *Mathématiques générales à l’usage des sciences économiques, de gestion et A.E.S.*, De Boeck Université, Bruxelles, 1993 ; *Analyse multivariée*, Editions Point de Vue, Liège, 1995.
- [6] BARA M.F. - GERMAT L. - LABOURE M.J. - LOUAT B. - MOYNOT S. - MULLER H., *Mathématiques et statistiques appliquées à l’économie*, Collection “Grand Amphi Economie”, Bréal, Rosny cedex, 2001.
- [7] BOURSIN J.L., *Éléments de mathématiques - Deug économie*, 2ème édition, Montchrestien, Paris, 1990.
- [8] COURTADE-COULOMB G., *Bases mathématiques pour l’économie et la gestion*, tome 2, Les Editions d’Organisation, Paris, 1991.
- [9] DEBIÈVE C. - FÉLIX Y., *Analyse infinitésimale pour HEC et ingénieurs commerciaux*, De Boeck Université, Bruxelles, 1995.

- [10] DOWLING E.T., *Mathématiques pour l'économiste*, Série Schaum, McGraw-Hill Inc., Paris, 1990.
- [11] ESCH L., *Mathématiques pour économistes et gestionnaires*, De Boeck Universtité, Bruxelles, 1992.
- [12] FOURASTIÉ J., *Mathématiques appliquées à l'économie*, 4ème édition, Dunod - module Economie, Paris, 1994.
- [13] GAUD D. - GUICHARD J. - SICRE J.P. - CHRÉTIEN C., *Des tangentes aux infiniment petits : réflexions et travaux pour la classe*, Irem de Poitiers, 1998.
- [14] GUERRIEN B., *Initiation aux mathématiques : sciences économiques et sociales*, 2ème édition, Economica, Paris, 1991.
- [15] HAYEK N. - LECA J.P., *Mathématiques pour l'économie : analyse - algèbre*, Dunod, Paris, 2001.
- [16] HENDERSON J.M. - QUANDT R.E., *Microéconomie : formulation mathématique élémentaire*, 2ème édition, Dunod, Paris, 1972.
- [17] HENRY V., *Questions de didactique soulevées par un enseignement de l'analyse non standard à de futurs économistes*, thèse doctorale soutenue à l'Université Paul Sabatier (Toulouse III), 2004.
- [18] JANSSEN J., *Traité de Mathématiques Appliquées pour l'assurance, l'économie et la finance*, tome 2 *Fonctions*, Office international de Librairie, Bruxelles, 1993.
- [19] KLOTZ G., *Mathématiques pour les sciences économiques et sociales*, Presses Universitaires de France, Paris, 1998.
- [1] LECOUTRE J.P. - PILIBOSSIAN P., *Analyse I*, 2ème édition, Dunod, Paris, 2002.
- [20] LEIBNIZ G.W., *La naissance du calcul différentiel*, 26 articles des *Acta eruditorum*; introduction, traduction et notes par M. PARMENTIER, Librairie Philosophique J. Vrin, Paris, 1995.
- [21] MICHEL P., *Cours de Mathématiques pour Economistes*, Economica, Paris, 1984.
- [22] PASSET R., *Introduction aux mathématiques de l'analyse économique*, tome 3, *la mesure des variations corrélatives*, Editions Cujas, Paris, 1971.
- [23] PLANCHE A., *Mathématiques pour économistes : Analyse*, 2ème édition, Dunod, Paris, 1999.
- [24] POUPALION G. - PUPION, G., *Les mathématiques de l'économiste*, Vuibert, 1999.
- [25] ROUCHE N., Rôle de l'erreur dans l'apprentissage et l'enseignement de la mathématique, *Compte rendu de la 39e rencontre internationale de la CIEAEM*, Les Editions de l'Université de Sherbrooke, 1988, pp. 97 - 121.

- [26] SCHLACTHER D., *Comprendre la formulation mathématique en économie*, Hachette supérieur, Paris, 1994.
- [27] SIMON C.P. - BLUME L., *Mathématiques pour économistes*, De Boeck Université, Bruxelles, 1998.

Table des figures

1.1	Axiome d'Archimède	9
1.2	Deux sortes de continuum	13
1.3	Ellipse	14
1.4	Parabole	15
3.1	Ordre pour deux réels	37
3.2	Fonction supérieure à une autre	37
3.3	Ordre local entre deux fonctions	38
3.4	Ordre local entre deux fonctions (bis)	38
5.1	Angle corniculaire	52
6.1	Courbe cinétique	62
6.2	Espace parcouru en fonction du temps (cas d'un MU)	63
6.3	Espace parcouru en fonction du temps (cas d'un MUA)	64
6.4	Vitesse en fonction du temps (cas d'un MUA)	64
6.5	Angle de contact	65
6.6	Somme de deux angles	68
6.7	Ordre sur les angles	69
8.1	Sous-tangente	81
9.1	Parabole et cercles tangents à l'oeil nu	94
9.2	Vision avec un microscope	95
9.3	Microscope de microscope	98
9.4	Approximation circulaire d'une ellipse	101

Table des matières

Avant-propos	3
1 Deux approches	7
1.1 Infini actuel <i>vs</i> infini potentiel	7
1.2 Axiome d'Archimède	9
1.3 Lois heuristiques de Leibniz	11
1.4 Analyse classique	15
1.5 Analyse non standard	18
1.6 L'analyse dans l'enseignement	22
1.7 Conclusion	25
2 Existence des infinitésimaux	29
2.1 Historique succinct	29
2.2 Relativité de la notion de nombre	30
2.3 Arguments apologétiques	31
3 Cadre fonctionnel	35
3.1 Existence d'infinitésimaux	36
3.2 Introduction des superréels	39
4 Vision positiviste	43
4.1 Introduction	43
4.2 Angles rectilignes	44
4.3 Angles corniculaires	45
4.4 Angles de demi-cercles	46
4.5 Angles recti-circulaires	47
4.6 Conclusion	48
5 Nombres superréels	51
5.1 Introduction	51
5.2 Angles mixtilignes	52
5.3 Introduction	54
5.4 Mesure d'un angle corniculaire	57
5.5 Rapport de mesures	59

6	Vitesses et infinitésimaux	61
6.1	Vitesse géométrique	62
6.1.1	MU	62
6.1.2	MUA	63
6.1.3	Quelques compléments	66
6.2	Outils mathématiques	67
6.2.1	Un concept général d'angle planaire	67
6.2.2	Addition d'angles	68
6.2.3	Un ordre pour les angles	68
6.2.4	Mesures d'angles	69
6.2.5	Nombres hyperréels	70
6.2.6	Le plus grand angle rectiligne sous une parabole	70
7	Construction de Robinson	73
7.1	Une \mathbb{N} -mesure	74
7.2	Définition et propriétés de ${}^*\mathbb{R}$	74
7.3	Infiniment petits ou grands	75
7.4	Une expérimentation	76
8	Algorithme de Sluse	79
8.1	Un savant liégeois	79
8.1.1	Une courte biographie	79
8.1.2	L'œuvre mathématique	80
8.2	Algorithme	81
8.2.1	Un exemple	81
8.2.2	Justification de l'algorithme	82
8.3	Au moyen de microscopes	83
8.3.1	Le cadre de la méthode	83
8.3.2	La méthode proprement dite	84
8.3.3	Comparaison avec l'algorithme de Sluse	85
9	Cercle osculateur	89
9.1	Introduction	89
9.2	Présentations classiques	91
9.2.1	Courbure	91
9.2.2	Une première définition	91
9.2.3	Une deuxième définition	92
9.2.4	Une troisième définition	93
9.3	Une autre approche	94
9.3.1	Un premier exemple	94
9.3.2	Nombres hyperréels	95
9.3.3	Microscope	96
9.3.4	Microscope de microscope	97
9.3.5	Cercles tangents	98

<i>TABLE DES MATIÈRES</i>	139
9.4 Nouvelle présentation	99
9.4.1 Une quatrième définition	99
9.4.2 Un deuxième exemple	100
10 Analyse multivariée	103
10.1 Espace numérique	103
10.2 Continuité	105
10.3 Dérivabilité	106
10.4 Courbes algébriques	110
10.5 Fonctions implicites	113
10.6 Différentiabilité	115
11 Différentielle	121
11.1 Introduction	121
11.2 Pour économistes	123
11.2.1 Littérature	123
11.2.2 Réflexions didactiques	126
11.2.3 Occurrence	128
11.2.4 Ordre d'apparition	129
11.3 Nouvelle approche	130
11.3.1 Définitions	131
11.3.2 Réflexions didactiques	131