Quelques aspects de l'histoire des équations fonctionnelles liés à l'évolution du concept de fonction

J. DHOMBRES

Présenté par A. P. YOUSCHKEVITCH

Fixer une origine précise à l'histoire des équations fonctionnelles entraîne une discussion *a priori* oiseuse puisque c'est le concept de fonction qui fait problème. H. LEBESGUE soulignait à ce propos la variation du sens des mots de la langue mathématique: "*le mot fonction qui servit autrefois à désigner les puissances de la variable a acquis lui aussi un sens extrêmement large*"[1]. Et les équations fonctionnelles font précisément partie des problèmes qui conduisirent les mathématiciens, à partir du XVIII^ème^ siècle, à élargir puis à vraiment définir le concept de fonction. Notre propos est d'éclairer ces différentes définitions qui contribuèrent à forger le concept de fonction au cours du temps, *en focalisant sur la mise en application pratique de ces définitions dans le cadre des équations fonctionnelles*. En ce sens notre travail se veut une illustration particulière de l'article de A. P. YOUSCHKEVITCH sur le concept de fonction[2].

1. La théorie grecque des proportions

Certes les tables babyloniennes associent des valeurs à d'autres dès le deuxième millénaire avant Jésus-Christ. Tout aussi bien, PTOLÉMÉE dans *l'Almageste* vers le 2ème siècle après Jésus-Christ détaille avec précision des tables trigonométriques. Cependant, même dans le monde alexandrin d'EUCLIDE ou d'APOLLONIUS, l'idée de fonctionnalité n'est certainement pas mise en évidence. Aucun mot particulier ne la désigne et on peut à peine parler avec E. T. BELL d'"*instinct for functionali-*

[1] H. LEBESGUE: A propos de quelques travaux mathématiques récents. Cet article de LEBESGUE fut écrit en 1905, mais publié seulement dans l'Enseignement mathématique, tome 17, Fasc 1, 1971, avec des notes de G. CHOQUET.

[2] Sur le concept de fonction à proprement parler, c'est le texte très documenté suivant qui nous a servi de fil d'Ariane dans ce travail. A. P. YOUSCHKEVITCH, The concept of function up to the middle of the 19th century, Arch. Hist. Exact Sc. **16** n° 1 (1976/77), p. 37–85. (Traduction française dans Fragments d'histoire des mathématiques, Brochure APMEP, 1981, p. 7–68.)

ty"[3]. Un exemple assez parlant nous est fourni par EUCLIDE, vers 300 avant Jésus-Christ, au Livre XII des *Eléments*, Proposition 2. Il y établit, par ce qu'il est convenu d'appeler depuis GRÉGOIRE DE SAINT-VINCENT au 17ème siècle *la méthode d'exhaustion*[4], que "*deux cercles sont entre eux en raison double de leurs diamètres*". Le plus remarquable est sans doute le besoin, satisfait brillamment, d'une démonstration que l'intuition de similitude rend presque évidente, par exemple à partir du théorème élémentaire prouvé pour les triangles semblables. Toutefois, il est sensible qu'EUCLIDE ne donne pas au résultat énoncé une quelconque forme de relation fonctionnelle, laquelle d'ailleurs l'obligerait à préciser le *λόγος* que nous notons Π depuis EULER. Non qu'EUCLIDE ignore comment calculer, approximativement, l'aire d'un cercle connaissant son diamètre. Mais, dans les *Eléments*, ce n'est certainement pas la technique numérique qui prédomine. En fait, quant à l'étude de ce qui relève du continu, comme par exemple la mesure des aires, la mathématique grecque pratique avec une virtuosité aisée une manipulation des proportions dont la définition et les opérations permises sont codifiées au livre V des *Eléments*, livre que la tradition fait remonter à EUDOXE[5]. Ce calcul sur les proportions peut certes conduire à des aspects numériques fort distingués, comme ARCHIMÈDE (287–212 avant J. C.) le démontre avec aisance avec sa *Mesure du Cercle*, dans laquelle il encadre Π entre $3 + \frac{10}{71}$ et $3 + \frac{1}{7}$ et pourrait affiner au besoin. Cependant le but de la théorie des proportions est autre. Son langage vise à l'énoncé de propriétés de type universel. Pour ce qu'en langage fonctionnel nous considérons comme une liaison de l'aire d'un cercle au carré du diamètre de ce cercle (par la formule $4S = \Pi D^2$), la théorie

---

[3] E. T. BELL, *The Development of Mathematics*, New York, McGraw-Hill, 1940. (La citation est page 32, et est appliquée seulement aux mathématiques babyloniennes.)

[4] Cette méthode se traduit par l'épuisement au terme d'un raisonnement par l'absurde (exhaustion logique) de deux hypothèses parmi les trois seules possibles. Le style de cette méthode est caractéristique et, malgré sa particularité, il servit de modèle admiré et repris jusqu'au 17ème siècle (*cf.* J. DHOMBRES, L'écriture mathématique, dans *L'accession à l'écriture*, Textes et Langages, I, Université de Nantes, 1978, p. 67–103.

[5] Et rien ne vient controuver cette tradition. En langage moderne, une raison est un élément du corps des fractions associé à un semi-groupe commutatif, divisible, totalement ordonné et archimédien. Il s'agit donc d'un rapport agréablement noté A/B. Cette notation est *inexistante* jusque vers la fin du XVIIème siécle. Aussi des régles presque évidentes sous cette notation, comme $A/B = C/D$ entraîne $A/C = B/D$, ou $A/B = C/D$ entraîne $A/B = (A + C)/(B + D)$, font l'objet d'une mémorisation dans la pratique "euclidienne" des proportions. Et les manuels pédagogiques du 17ème siècle attestent encore une telle pratique dont il reste à faire l'histoire. Une étude du livre V et de ses avatars dans l'histoire fut entreprise dans J. DHOMBRES, *Nombre, mesure et continu. Epistémologie et histoire*, Nathan, Paris, 1978. Beaucoup plus fouillée et érudite est l'étude du texte euclidien par W. KNORR, *The evolution of the Euclidean elements. A study of the theory of incommensurable magnitudes and its significance for early Greek geometry*, D. Reidel, 1975. En ce qui concerne les proportions au Moyen-Age, citons deux commentaires particulièrement intéressants: E. GRANT, Part I of Nicole Oresme's *Algorismus proportionum*, ISIS, **56**, n° 185, 1965, p. 327–342, J. E. MURDOCH, The medieval language of Proportions: Elements of the interaction with Greek foundations and the development of new mathematical techniques, in *Scientific Change*, Ed. A. C. CROMBIE, Heinemann, 1962.

des proportions compare deux éléments de même espèce (les aires de deux cercles quelconques) à deux autres éléments éventuellement d'une autre espèce (les carrés des deux diamètres). Elle énonce: l'aire d'un cercle est à l'aire d'un autre cercle comme le carré du diamètre du premier cercle est au carré du diamètre du deuxième cercle. On raccourcit: deux cercles sont entre eux en raison double de leurs diamètres (double est mis au sens de carré). Ce vocabulaire, cette comparaison faisant intervenir 4 éléments au lieu de 2, malgré son élégance, bloque l'idée de fonctionnalité et *a fortiori* celle d'équation fonctionnelle. Disons, pour amoindrir ce propos, que le seul *style* mathématique disponible alors pour évoquer une relation fonctionnelle est celui des proportions et qu'il est mal adapté à l'idée fonctionnelle. Et c'est pourtant avec cette seule technique des proportions que le concept va être dégagé, mais avec difficulté. Un parcours passionnant!

2. Fonctions linéaires et affines: une caractérisation au Moyen-Age?

Pour nous, lecteurs du XX$^{\text{ème}}$ siècle oublieux des techniques rituelles de la théorie des proportions, c'est d'ailleurs ce langage même qui alourdit l'exposé des idées nouvelles développées par les Ecoles d'Oxford et de Paris au XIV$^{\text{ème}}$ siècle. N. ORESME (1323–1382) en est le mathématicien le plus accompli. Même si le langage utilisé est quelquefois cinématique, ces Ecoles conçoivent une représentation "*abstraite*" ("*ymaginationes*" dit ORESME ou "*secundum imaginationem*" à Oxford) du changement, par une théorie de la latitude des formes comme dans le *Tractatus de latitudinibus formarum*. Il s'agit dans cette théorie médiévale de prendre en compte les variations ou les intensités de quantités continues, des *qualités* selon le vocabulaire utilisé, telles que la distance d'un point mobile par rapport à un point fixe, à partir d'autres quantités telles que le temps, lequel devient l'exemple type du continu. L'étude de ces variations n'est pas seulement théorique. La représentation graphique se fait en latitude (*latitudo*), notre ordonnée, et en longitude (*longitudo*), notre abscisse. ORESME dénombre plusieurs genres de qualités. D'une part les qualités uniformes (*qualitas uniformis*) que nous appelerions des fonctions constantes, et dont la représentation graphique fournit des rectangles. Ensuite les qualités uniformément difformes (*uniformiter difformis*) dont la représentation graphique fournit des trapèzes. Pour ces dernières, une définition équivalente chez ORESME utilise le langage des proportions et énonce verbalement ce que nous écrivons aujourd'hui sous la forme d'une égalité de quotients:

$$(1) \qquad \frac{f(x_1) - f(x_2)}{f(x_2) - f(x_3)} = \frac{x_1 - x_2}{x_2 - x_3}, \quad \text{pour tous } x_1, x_2 \text{ et } x_3 \text{ (distincts).}$$

"*Qualitas vero uniformiter difformis est cuius omnium trium punctorum proportio distantie inter primum et secundum ad distantiam inter secundum et tertium est sicut proportio excessus primi supra secundum ad excessum secundi supra tertium in intensione ...*"

"*La qualité uniformément difforme est celle pour qui, trois points quelconques* *étant donnés, le rapport de la distance du premier au second à la distance du second au troisième est égal au rapport de l'excès de l'intensité du premier sur le second à l'excès de l'intensité du second sur le troisième ...*".

L'équivalence entre la représentation graphique du trapèze (figure 1) et la relation (1) est *explicitement* soulignée chez ORESME (*De configurationibus*, Partie I, chap. XI, [6]). ORESME toutefois ne *démontre* qu'un sens de l'équivalence, à savoir que la représentation graphique par une droite implique la relation fonctionnelle (1). Il le fait en utilisant la géométrie des triangles semblables, distinguant le cas du trapèze de celui du triangle rectangle. Cependant *l'équivalence entre la définition graphique et la relation* (1) *lui paraît acquise* puisque le troisième genre de qualité, les qualités difformément difformes (*difformiter difformis*), n'est défini que par la *négation* des deux autres qualités (négation effectuée sur la relation (1) dans le cas le plus notable).

"*Omnis autem qualitas se habens alio modo a predictis dicitur difformiter difformis et potest describi negative, scilicet qualitas que non est in omnibus partibus subiecti equaliter intensa nec omnium trium punctorum ipsius proportio excessus primi supra secundum ad excessum secundi supra tertium est sicut proportio distantiarum eorum*"

"*Cependant toute qualité ayant un mode autre que ceux prescrits est dite difformément difforme. Elle peut être décrite négativement comme une qualité qui n'est ni égale d'intensité dans toutes les parties du sujet ou qui n'est pas celle pour qui, trois points quelconques étant donnés, le rapport de l'excès de l'intensité du premier au second sur l'excès de l'intensité du second au troisième est égal au rapport des distances*"[6].

Puisque la relation fonctionnelle (1) fait intervenir des proportions, un changement d'unité dans la mesure de l'intensité $f(x)$ ne modifie en rien le caractère uniformément difforme de la qualité en jeu. En bon manipulateur des proportions,

[6] Ce passage est tiré du *De Configurationibus qualitatum et motuum* d'ORESME (Pars I, Cap. XI). Le titre du chapitre est: Des qualités uniformes et difformes. Une traduction anglaise et un commentaire copieux se trouvent dans M. CLAGETT, *Nicole Oresme and the medieval geometry of qualities and motions. A treatise on the uniformity and difformity of intensities known as Tractatus de configurationibus qualitatum et motuum.* Madison, The University of Wisconsin Press 1968.

Voir aussi E. GRANT, *A source book in medieval science*, Cambridge, Massachusetts, Harvard Univ. Press, 1974. Précisons qu'ORESME appelle premier point celui dont l'intensité est la plus grande, puisqu'il ne manipule pas les quantités négatives. Aussi termine-t-il la phrase citée par "*ita quod punctum intensiorem illorum trium voco primum*". Il écarte de même, à nos yeux, les quantités négatives dans le second membre de l'équation (1) en ne parlant que de distances. On remarquera que ces restrictions sont pénibles à porter sur les notations modernes utilisées selon l'équation (1), ce qui donne une mesure de l'écart commis en réduisant le texte d'ORESME à l'équation (1). En effet, à l'équation fonctionnelle (1), il faudrait rajouter les conditions: $f(x_1) > f(x_2) > f(x_3)$; $x_1 > x_2 > x_3$, ou éventuellement mettre des valeurs absolues au second membre. La deuxième citation est le dernier paragraphe du *De configurationibus*, Pars I, Cap. XI. (Nous soulignons la description négative.)

ORESME sait cela. Semblablement donc, une affinité verticale sur la représentation graphique transforme un trapèze en un autre trapèze, ou un triangle rectangle en un autre triangle rectangle si l'intensité terminale est nulle. ("*Qualitas uniformiter difformis terminata in intensione ad non gradum*"). ORESME n'est pas abusé par une équivalence grossière entre (1) et la représentation graphique. Il prend bien soin de ne conclure sur la relation (1) d'une propriété de la seule représentation graphique que dans la mesure où cette propriété lue géométriquement est invariante par affinité verticale[7]. Du coup, dans sa classification des quantités difforméement difformes, ORESME introduit la difformité difforme simple ("*simplex difformis difformitas*") qui revient à notre convexité géométrique, laquelle est bien invariante par affinité verticale (de rapport positif bien entendu). Mathématiquement, ORESME n'ira guère plus loin mais l'ouverture fonctionnelle est effectuée en dépassant les limitations de la théorie des proportions. La *caractérisation fonctionnelle* énoncée par (1) est le pivot central de la classification en trois genres: les qualités uniformes (fonctions constantes), les qualités uniformément difformes (fonctions affines) et les qualités difformément difformes (les autres). Et avec (1), cette idée explicitement fonctionnelle pouvait déboucher sur une nouvelle équation fonctionnelle en prenant pour point x_2 le milieu des points x_1 et x_3. On déduit de (1), sous cette condition, l'équation fonctionnelle dite de JENSEN aujourd'hui et sur l'histoire de laquelle nous aurons l'occasion de revenir,

$$f\left(\frac{x_1 + x_3}{2}\right) = \frac{f(x_1) + f(x_3)}{2}. \tag{2}$$

De fait, ce calcul fut effectivement fait. Non seulement fait, mais souligné par ce que l'on a coutume aujourd'hui d'appeler le théorème de Merton. Eponymie, car certains auteurs auxquels on attribue communément le résultat, comme R. SWINESHEAD, W. HEYTESBURY et J. DUMBLETON, travaillaient au Merton College d'Oxford vers les années 1330. Ce résultat figure aussi chez ORESME vers 1350 et fut souvent répété dans des manuels ultérieurs, ce qui en souligne l'importance pédagogique.

Le théorème de Merton sous sa forme cinématique énonce que la distance parcourue par un mobile en déplacement rectiligne, dont la vitesse suit une variation

---

[7] Nous utilisons à preuve le texte *Tractatus de figuratione potentiarum et mensurarum difformitatum* (De latitudinibus formarum ab Oresme, porte un manuscrit et s'il n'est pas d'Oresme, son influence s'en ressent). Des extraits sont traduits et commentés avec finesse dans P. DUHEM, *Le Système du Monde*, Hermann, Paris, Tome VII, chap. VI et VII, 1956. Fournissons seulement les deux extraits suivants, très éclairants.

"*Toute qualité représentable par un triangle rectangle dont l'angle droit a la longitude pour côté, peut être représentée par tout autre triangle rectangle qui aurait un angle droit placé de même, et ne peut être représentée par aucune autre figure*".

A l'inverse, ORESME constate qu'une représentation graphique par un demi-cercle n'est pas stable par affinité verticale (*cf.* P. DUHEM, *opus cité*, p. 545):

"*Mais je dis qu'elle ne peut être représentée par aucune figure plus haute que le demi-cercle et qui soit une portion de cercle*".

"*uniformément difforme*" (mouvement uniformément accéléré), est exactement celle parcourue par un mobile dont la vitesse est uniforme et égale à la moitié de la somme de la vitesse initiale et de la vitesse finale du premier mobile. En notations modernes, en appelant $x(t)$ l'abscisse au temps t du premier mobile et en notant $v(t)$ sa vitesse, on dispose pour des temps t_1 et t_2, de l'égalité,

$$x(t_2) - x(t_1) = \frac{v(t_1) + v(t_2)}{2}(t_2 - t_1). \tag{3}$$

Mais l'énoncé du théorème est plutôt donné en utilisant la propriété fonctionnelle (2) quant à la vitesse, c'est-à-dire:

$$v\left(\frac{t_1 + t_2}{2}\right) = \frac{v(t_1) + v(t_2)}{2}. \tag{4}$$

Soit la forme générale du théorème de MERTON, *l'equivalence* (en termes modernes) entre un mouvement uniformément accéléré et un mouvement uniforme dont la vitesse est celle du mouvement accéléré à mi-temps

$$x(t_2) - x(t_1) = v\left(\frac{t_1 + t_2}{2}\right)(t_2 - t_1). \tag{5}$$

Cette dernière formulation est en effet particulièrement adaptée à un énoncé visualisé par la représentation graphique. Ainsi chez ORESME, on lit une *équivalence* entre deux qualités: "*Omnis qualitas, si fuerit uniformiter difformis, ipsa est tanta quanta foret qualitas eiusdem subiecti vel equalis uniformis secundum gradum puncti medii eiusdem subiecti*". (*Toute qualité uniformément difforme a même quantité que si elle informait uniformément le même sujet ou un sujet égal selon le degré du point milieu de ce sujet*[8].)

Chez ORESME, mais déjà chez HEYTESBURY, le théorème de MERTON prend une tournure générale, sans référence à une cinématique, mais pour des qualités abstraites. Et chez ORESME ces qualités abstraites sont celles pour lesquelles la mesure des intensités peut s'effectuer grâce à la théorie des proportions, c'est-à-dire la théorie du Livre V d'EUCLIDE que nous résumons à l'excès en parlant de théorie des qualités continues. Comme ORESME considère que le paradigme de cette théorie est la mesure des longueurs, on comprend son insistance pour une représentation graphique.

"*Omnis igitur intensio successive acquisita ymaginanda est per lineam rectam perpendiculariter erectam super aliquod punctum aut aliquot puncta intensibilis spacii vel subiecti ...*"

"*Donc toute intensité susceptible d'être acquise d'une manière successive doit être*

---

[8] De figuratione potentiarum, Pars III, Cap. VII (De mensura qualitatum et velocitatum difformarum). Nous avons pris la traduction française un peu libre fournie par P. DUHEM, dans *Le Système du Monde*, Hermann, Paris (Tome VII, page 556): en y ajoutant l'expression "*ou un sujet égal*".

imaginée au moyen d'une ligne droite élevée verticalement à partir de chaque point de l'espace ou du sujet qu'affecte cette intensité"[9].

En tout cas, (5) est exprimée, géométriquement, pour des "*qualités*" que nous notions $x(t)$ soumises à des variations mesurées par une vitesse que nous notons $v(t)$. Nous ne soulevons pas ici la signification médiévale attribuée à la vitesse de variation d'une qualité (sa dérivée!). Il nous suffit de savoir que la démonstration choisie par ORESME pour ce théorème de Merton est de nature géométrique à partir de la représentation graphique des seules vitesses de variation. Il y a égalité (figure 1) entre l'aire du trapèze $ABCD$ et celle du rectangle $ABEF$. Puisque

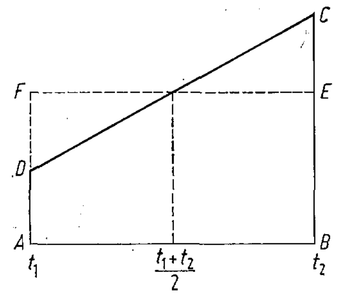

Fig. 1

l'intensité de la vitesse de variation de la qualité en jeu est portée en ordonnée, l'aire est interprétée comme la mesure de la variation totale de la qualité. Il faut donc bien noter que le théorème de Merton sous la forme (3) fédère deux résultats: l'un, résultat le plus novateur, est de calcul intégral engrangé sous forme géométrique de l'aire décrite, et l'autre résultat est une relation fonctionnelle caractéristique d'une fonction affine. Pour le résultat de calcul intégral, la représentation graphique si originale d'ORESME est bien adaptée: il appelle quantité de la qualité l'expression $x(t_2) - x(t_1)$:

"*Quod enim quantitas talis qualitatis per huiusmodi superficiem possit ymaginari patet, quia contingit dare superficiem illi quantitati equalem in longitudine sue extensioni, et similem in altitudine eidem quantitati in intensione*".

"*Que la quantité d'une telle qualité se puisse imaginer à l'aide d'une telle surface, cela est évident, puisqu'on peut donner une surface égale en longitude à l'extension de cette quantité, et semblable, en altitude, à l'intensité de cette quantité*"[9bis].

On imagine bien que des confusions aient pu naître dans le langage du XIV$^{\text{ème}}$ siècle entre la quantité de la qualité ($x(t)$) et la seule intensité de la qualité ($v(t)$),

---

[9] N. ORESME, *Tractatus de figuratione potentiarum et mensurarum difformitatum*, Pars I, Cap. I (De continuitate intensionis, traduit par DUHEM, p. 537, op. cité).

[9bis] Citation tirée du texte indiqué en [9], Pars I, Cap. IV, De quantitate qualitatis, trad. DUHEM, op. cité p. 541.

confusions que la représentation graphique d'ORESME abolit, confusions qui expliquent sans doute les tergiversations au sujet du théorème de Merton[10]. La seule forme (5) du théorème de Merton est privilégiée dans le texte cité, sans doute grâce à son élégante interprétation géométrique. Elle l'est d'autant plus volontiers qu'ORESME n'hésite pas à considérer ce que nous appelons des fonctions de deux ou de trois variables. En effet, nous l'avons vu jusqu'ici porter en longitude des grandeurs semblables à des longueurs, ce qu'ORESME appelle des quantités linéaires. Mais, en connaisseur raffiné de la théorie des proportions, ORESME admet bien que l'on puisse porter en longitude des surfaces comme des rectangles (qualités superficielles) voire des volumes (qualités corporelles). Non pas la mesure de ces surfaces ou volumes, mais bien une grandeur à deux ou trois dimensions. Dans le cas des surfaces planes, il porte alors perpendiculairement l'intensité correspondante. La représentation graphique est ainsi à trois dimensions et ORESME indique à nouveau la distinction entre qualités uniformes et qualités uniformément difformes. Pour celles-ci la représentation est effectuée à partir d'un plan non parallèle à la base. Le cas des volumes est abstrait. Le théorème de Merton est précisé dans ces deux cas-là.

"*Et hoc intelligo; secundum gradum puncti, si qualitas fuerit linealis; et si fuerit superficialis, secundum gradum linee medie; et si fuerit corporalis, secundum gradum medie superficiei, semper conformiter intelligendo*".

"*En disant: selon le degré du point milieu, je sous-entends: si la qualité est linéaire; si elle superficielle, il faudra dire: selon le degré de la ligne moyenne; et si elle est corporelle, il faudra toujours comprendre selon le degré de la surface moyenne*"[9ter].

Avec cette généralisation en vue, on comprend bien l'emploi de la forme (5) du

---

[9ter] Citation tirée du texte indiqué en [9], Pars III, Cap VII, De mensura qualitatum et velocitatum difformarum, trad. DUHEM, op. cité p. 556. Signalons que lorsque le sujet est un temps, non une longueur, c'est-à-dire pour la version cinématique du théorème de Merton, ORESME précise temps moyen au lieu de temps milieu.

[10] Des indications de nature arithmétique, quant à une démonstration du théorème dit de Merton, plus exactement des indications basées sur la théorie eudoxienne des proportions sans référence à une interprétation graphique par les surfaces, furent antérieurement données, notamment par R. SWINESHEAD, dit le Calculateur dans son *Liber Calculationum*, ou HEYTESBURY dans *Regule solvendi sophismata*. Voir par exemple C. WILSON, *W. Heytesbury*, Univ. of Wisconsin Press, Madison, 1956. P. DUHEM a contesté la priorité d'Oxford au sujet de la pleine compréhension du théorème de Merton et cité plusieurs passages contradictoires de THOMAS BRADWARDINE dans ses *Proportiones*, des *Regule solvendi sophismata* de WILLIAM HEYSTESBURY, de JOHN OF DUMBLETON dans *Summa logicae et naturalis philosophiae* et de ROGER SWINESHEAD enfin. DUHEM souligne en tout cas le rôle du traité anonyme et antérieur *De proportionalitate motuum et magnitudinum*, cité explicitement par BRADWARDINE. Quoiqu'il en soit, la représentation graphique d'ORESME constitue un pas fondamental dans la compréhension du résultat que l'on qualifie de théorème de Merton. Une référence très nourrie sur ce théorème est M. CLAGETT, *The Science of Mechanics in the Middle Ages*. Univ. of Wisconsin Press, Madison, 1959, corr. reprint, 1961.

théorème, car la forme (3) dans ces cas plus généraux exigerait des précisions incommodes comme le dessin ci-contre permet de s'en rendre compte (figure 2).

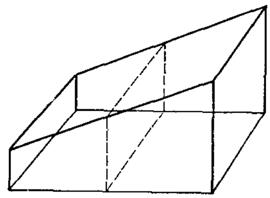

Fig. 2

Chez ORESME donc, l'équivalence de la forme (3) du théorème de Merton et de la forme (5), grâce à (4), n'est qu'implicite, ne serait-ce que dans la démonstration du résultat par utilisation du théorème de THALÈS. Par contre la forme (3) du théorème de Merton et l'équivalence (4) se trouvent bien dans d'autres textes médiévaux, ce qui nous permet de mieux séparer les deux résultats réunis par le seul théorème de Merton. Nous prendrons comme exemple un texte de HEYTESBURY, écrit sans doute vers 1335, dont le titre est: *Regule solvendi sophismata.* On y trouve explicitement des calculs du genre $v\left(\frac{t_2}{2}\right) = \frac{v(t_2)}{2}$ c'est-à-dire une spécialisation de (4) lorsque le mouvement est uniformément accéléré avec $v(t_1) = 0$ en $t_1 = 0$:

"*D'ailleurs, pour toute telle latitude qui commence au repos et se termine à un certain degré, le degré moyen est la moitié du degré qui termine cette même latitude*"[10bis].

La démonstration de (4) lorsque $v(t_1) = 0$, que fournit HEYTESBURY, n'est pas facile à interpréter, d'autant que le titre même de l'ouvrage où se trouve la preuve en indique déjà la portée philosophique générale. Une lecture possible nous paraît être la suivante. Mais rappelons, ce qui est essentiel pour la compréhension du passage, qu'HEYTESBURY raisonne *sans* l'interprétation graphique d'ORESME.

D'abord la définition d'un mouvement uniformément accéléré n'a pas la simplicité théorique générale de la relation fonctionnelle (1). Elle restreint (voir (2)) en fait la relation (1) au cas $x_1 - x_2 = x_2 - x_3$. C'est-à-dire qu'il y a mouvement uniformément accéléré ("*uniformiter intenditur*") lorsque pendant deux parties égales quelconques de temps il y a égal accroissement de vitesse. Et d'ailleurs HEYTESBURY doit distinguer le cas de l'uniforme accélération de celui d'uniforme décélération, ce qu'ORESME évite par son choix astucieux d'ordre des points. Reprenons les choses avec les temps. Ce qui semble pensé chez HEYTESBURY en place de la relation fonctionnelle (1) pour la vitesse, c'est une propriété de

[10bis] W. HEYTESBURY, *Regule solvendi sophismata*, Pars VI, De tribus praedicamentis, (1ère division, de motu locali, du mouvement local). Nous fournissons la traduction de P. DUHEM, op. cité, p. 641. La définition de la vitesse instantanée que fournit HEYTESBURY est très remarquable par sa modernité: c'est celle de GALILÉE.

permanence sur les vitesses, soit la permanence de *certaines* proportions seulement, quand ces proportions sont réalisées sur les temps. Même si cette permanence est le noeud de la définition du mouvement uniformément accéléré, elle ne s'applique pas pour le temps où il y a repos, donc absence de mouvement. Ainsi si en $t = 0$ nous avons une vitesse nulle, il s'agit effectivement de prouver $v\left(\frac{t}{2}\right) = \frac{v(t)}{2}$ pour les autres temps. C'est un approfondissement notable de la théorie des proportions qui permet d'avancer. HEYTESBURY exploite un résultat général sur les proportions continues et que nous pouvons, en notations modernes, expliquer ainsi. Si $\frac{a_1}{a_2} = \frac{a_2}{a_3} = \frac{a_3}{a_4} = etc.$, alors $\frac{a_1}{a_2} = \frac{a_1 - a_2}{a_2 - a_3}$. Donc si a_2 vaut la moitié de a_1, $\frac{a_1}{a_2} = 2$ et $2 = \frac{a_1 - a_2}{a_2 - a_3} = \frac{a_2 - a_3}{a_3 - a_4} = etc.$ HEYTESBURY énonce:

$$a_1 - a_2 = (a_2 - a_3) + (a_3 - a_4) + (a_4 - a_5) + \ldots$$

On peut imaginer comme suit une preuve de cette égalité, laquelle preuve doit peut-être exister dans d'autres traités contemporains tant elle paraît acquise chez HEYTESBURY.

$$\begin{aligned}
a_1 - a_2 &= 2(a_2 - a_3)\\
&= (a_2 - a_3) + (a_2 - a_3)\\
&= (a_2 - a_3) + 2(a_3 - a_4)\\
&= (a_2 - a_3) + (a_3 - a_4) + (a_3 - a_4)\\
&= (a_2 - a_3) + (a_3 - a_4) + 2(a_4 - a_5)\\
&= (a_2 - a_3) + (a_3 - a_4) + (a_4 - a_5) + \ldots
\end{aligned}$$

Il s'agit d'un argument de dichotomies successives justifiable par la méthode d'exhaustion et qui ne fait intervenir que les proportions du style $a_1 - a_2 = 2(a_2 - a_3)$ et ainsi de suite. Si l'on applique cette proposition aux temps, avec $a_1 = t$, $a_2 = \frac{t}{2}$, le temps moitié, on a la relation

$$t - \frac{t}{2} = \left(\frac{t}{2} - \frac{t}{4}\right) + \left(\frac{t}{4} - \frac{t}{8}\right) + \cdots$$

Le résultat lu est trivial $t - \frac{t}{2} = \frac{t}{2}$. Mais la conséquence ne l'est pas. Comme le mouvement est uniformément accéléré, certaines proportions pour les vitesses son égales aux proportions correspondantes pour les temps. Cette permanence sur ces proportions permet donc d'écrire pour les vitesses, en ne faisant jouer de la même façon que les proportions du style $v(t) - v\left(\frac{t}{2}\right) = 2\left(v\left(\frac{t}{2}\right) - v\left(\frac{t}{4}\right)\right)$,

$$\begin{aligned}
v(t) - v\left(\frac{t}{2}\right) &= \left(v\left(\frac{t}{2}\right) - v\left(\frac{t}{4}\right)\right) + \left(v\left(\frac{t}{4}\right) - v\left(\frac{t}{8}\right)\right) + \cdots\\
&= v\left(\frac{t}{2}\right).
\end{aligned}$$

La deuxième égalité "*provenant*" de l'élimination successive des termes dans la somme infinie et du fait que $v(0) = 0$, le repos au temps initial. D'où le résultat qu'il fallait démontrer,

$$v\left(\frac{t}{2}\right) = \frac{v(t)}{2}.$$

Comme HEYTESBURY admet, sans preuve formelle, le théorème de Merton sous sa forme (5), il peut déduire du résultat précédent l'égalité,

$$x(t) = \frac{v(t)}{2}\,t.$$

Naturellement ce résultat est donné sous forme d'un rapport, c'est-à-dire en comparant $x(t)$ à $X(t) = v(t)\,t$ la distance parcourue par un mouvement uniforme à la vitesse finale $v(t)$. Ce qui fournit le rapport,

$$\frac{x(t)}{X(t)} = \tfrac{1}{2}.$$

"*Il résulte de la proposition précédente que si un mobile part du repos et si l'intensité de son mouvement croit uniformément jusqu'à un certain degré, il parcourra deux fois moins de chemin en un certain temps que s'il se mouvait uniformément, pendant le même temps, avec le degré qui termine la latitude*"[10bis].

HEYTESBURY indique aussi que pour un tel mouvement, si la vitesse au départ n'est pas nulle $(v(t_1) \neq 0)$, la vitesse au temps milieu dépasse strictement la moitié de la vitesse finale,

$$v\left(\frac{t_1 + t_2}{2}\right) > \frac{v(t_2)}{2}.$$

Donc il déduit dans ce cas, et en utilisant les notations précédentes, que

$$x(t) > \frac{X(t)}{2}.$$

C'est-à-dire que la distance parcourue dépasse la moitié de la distance qui serait parcourue dans le même temps si le mouvement était uniforme, de vitesse égale à la vitesse finale. On constate ainsi que la relation (2) n'est pas perçue en toute généralité. HEYTESBURY ne parvient donc pas à la formulation générale (4). Il essaie certes lorsque $v(t_1) \neq 0$ un calcul analogue à ce qu'il a fait par dichotomies successives pour le cas $v(t_1) = 0$ et conclut à une impasse.

"*Mais un calcul de ce genre offre plus de difficultés que d'avantages*"[10bis].

Il serait intéressant comme test de comparer dans différents textes du XIV$^{\text{ème}}$ siècle, et sans l'interprétation graphique d'ORESME qui annule la difficulté et marque donc un grand progrès, le traitement de (4). Nous ne le ferons pas ici. HEYTESBURY, toujours dans le cas d'une vitesse nulle en t_1, prouve aisément que la distance parcourue de t_1 à $(t_1 + t_2)/2$ est le tiers de celle parcourue de $(t_1 + t_2)/2$ à t_2. En effet, de t_1 à $(t_1 + t_2)/2$ la distance est celle réalisée par un mobile qui serait animé d'une vitesse uniforme, celle en $t_1 + t_2/4$, laquelle vaut $v(t_2)/4$. Il calcule aussitôt $v(t_2)/4 \cdot (t_2 - t_1)/2$ pour la distance. Donc de $(t_1 + t_2)/2$ à t_2, par différence, la distance parcourue vaut

$$\frac{v(t_2)}{2}(t_2 - t_1) - \frac{v(t_2)}{4}\left(\frac{t_2 - t_1}{2}\right) = 3\frac{v(t_2)}{4}\,\frac{t_2 - t_1}{2}.$$

Un tel résultat devient géométriquement évident et est généralisé chez ORESME. Celui-ci, lui aussi, dépasse le seul théorème de Merton en résolvant (4) pour un temps terminal à vitesse nulle. Il utilise géométriquement cette résolution dans l'équation (5) dont il propose une solution explicite. Ceci est fait dans les *Questions sur les éléments d'Euclide*[11], texte que l'on date généralement des années précédant 1350 tandis que le *De Configurationibus* viendrait juste après cette date. ORESME part donc de la représentation triangulaire en introduisant un temps terminal t_0 pour lequel la vitesse $v(t_0)$ est nulle (*cf.* figure 3). Et il considère des intervalles égaux de temps, c'est-à-dire $t_3 - t_2 = t_2 - t_1$, ou encore $t_2 = (t_1 + t_3)/2$ pour faire référence à (4). Il mesure des temps positifs, c'est-à-dire de la forme $t_0 - t_1$, $t_0 - t_2$ ou $t_0 - t_3$. En langage moderne, (4) fournit $v(t) = a(t - t_0)$ et fidèle naturellement au langage de la théorie des proportions, à partir de (3), ORESME énonce ce que l'écriture contemporaine réduit à

$$(6) \qquad \frac{x(t_3) - x(t_2)}{x(t_2) - x(t_1)} = \frac{(t_0 - t_2) + (t_0 - t_3)}{(t_0 - t_1) + (t_0 - t_2)}, \quad \text{avec} \quad t_2 = \frac{t_1 + t_3}{2}.$$

L'évêque de Lisieux conclut en donnant à l'équation fonctionnelle (6) une forme numérique très avantageuse pour une éventuelle vérification expérimentale, à laquelle ni lui, ni ses contemporains immédiats, ne semblent songer. Cette forme est la suivante. A partir d'une unité quelconque de temps — on ne mesure en fait que des rapports, c'est l'avantage des proportions — pour des valeurs entières successives de cette unité, ORESME calcule le rapport des accroissements successifs de la qualité de la quantité, accroissements représentés par des aires. Au second membre de (6) ce rapport apparaît comme rapport de deux nombres impairs successifs. Ainsi avec $t_0 - t_1 = 1$, $t_0 - t_2 = 2$ et $t_0 - t_3 = 3$, le rapport de $x(t_3) - x(t_2)$ à $x(t_2) - x(t_1)$ vaut 5/3 et ainsi de suite (*cf.* figure 3):

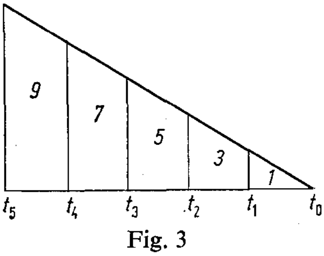

Fig. 3

[11] N. ORESME, *Questiones super geometriam Euclidis*, Ed. H. L. L. BUSARD, 2 vol., Leiden, 1961. (Voir aussi la référence [6] pour une traduction anglaise. Une traduction française est en cours de réalisation.)

"*Secundo conclusio est quod subiecto taliter diviso, (per aliquas partes equales) et vocetur semper pars remissior prima, proportio partialium qualitatum et habitudo earum ad invicem est sicut series imparium numerorum ubi prima est 1, secunda 3, tertia 5, etc., ut patet in figura*"

"*La seconde conclusion est qu'en divisant ainsi le sujet (en parties égales), et en désignant la plus petite partie comme la première, le rapport des qualités partielles, c'est-à-dire leur relation mutuelle, est comme la série des nombres entiers impairs où le premier est 1, le second 3, le troisième 5, etc, ce qui est évident sur la figure*"[11bis].

Il me paraît que cette conclusion va nettement plus loin que le calcul de HEYTESBURY en ce sens qu'une nouvelle équation fonctionnelle est introduite où interviennent cette fois des entiers, équation évidemment énonçable sous forme de rapports,

$$\frac{x[(n+1)\,t] - x(nt)}{x(nt) - x[(n-1)\,t]} = \frac{n + (n+1)}{n + (n-1)}.$$

Nous reviendrons sur cette équation. En tout cas, l'équation (6), sous la condition $t_2 = \dfrac{t_1 + t_3}{2}$, est bien une *équation fonctionnelle* déduite du théorème de MERTON, et donc en particulier de la première équation fonctionnelle (1) via la relation (4) spécialisée. ORESME connaît une solution de cette équation (6), solution qu'il indique sous la forme suivante:

"*Quod omnis subiecti uniformiter difformis ad non gradum proportio qualitatis totius ad qualitatem partis terminate ad non gradum est sicut proportio totius subiecti ad istam partem, proportio duplicata*".

– *Dans le cas de tout sujet uniformément difforme à partir d'un degré nul, le rapport de la qualité entière à la qualité d'une partie terminée à un degré nul est le carré du rapport du sujet entier à cette partie*[11ter].

Ce qui revient à écrire en notations modernes pour une solution de l'équation (6),

$$\frac{x(t)}{x(t')} = \left(\frac{t_0 - t}{t_0 - t'}\right)^2.$$

Et, rappelons-le, t_0 est le temps terminal, celui où $v(t_0) = 0$ (*cf.* figure 3). On peut donc conclure cette partie sur le monde médiéval en assurant qu'ORESME ne se contente pas d'introduire explicitement, et pour des problèmes abstraits, la représentation graphique d'une variation. Il en induit, pour un type de fonction, l'équivalence d'une définition géométrique, lue sur le graphique, et d'une défini-

---

[11 bis] Question 14 de l'ouvrage cité en [11].

[11ter] Question 13 de l'ouvrage cité en [11].

Signalons qu'une étude minutieuse des idées des écoles d'Oxford et de Paris se trouve dans A. P. JUSCHKEWITCH, *Geschichte der Mathematik im Mittelalter*, Leipzig, Teubner, 1964, S. 402–413.

tion analytique telle que formulée par une équation fonctionnelle[12]. D'autres équations fonctionnelles interviennent chez ORESME telles que celle d'une fonction quadratique. Autrement dit, *l'idée fonctionnelle est bien lancée.* Il est pourtant périlleux, faute de documents, de préciser l'influence pourtant vraisemblable d'ORESME et de l'école de Paris sur un GALILÉE ou sur un DESCARTES (par l'intermédiaire de I. BEECKMAN) au sujet de cette présentation fonctionnelle.

3. Utilisation essentielle de la résolution d'une équation fonctionnelle chez Galilée

Par une autre forme de calcul intégral, après les proportions continues d'HEYTESBURY, l'interprétation graphique d'ORESME ou les indivisibles de CAVALIERI, GALILÉE (1564–1642) expose le théorème de Merton et en déduit l'aspect que lui donna ORESME par utilisation de rapports de nombres impairs successifs. Son génie est d'appliquer ce résultat à l'étude de la chute des corps. Dans une lettre de 1604 à PAOLO SARPI, GALILÉE énonce:

"*Les espaces traversés par un mouvement naturel sont en proportion double* (le carré) *des temps et en conséquence les espaces traversés dans des temps égaux sont comme les entiers impairs à partir de l'unité*"[13].

Lors de la publication, en 1638, du *Discours et démonstrations mathématiques sur deux nouvelles sciences,* GALILÉE fournit une explication mathématique. Le théorème de Merton sous la forme (3) est d'abord démontré (Th. 1, Prop. 1), puis le résultat sur le carré des temps (Th. 2, Prop. 2). Enfin,

"*Si l'on prend à partir du premier instant, ou du début du mouvement, un nombre quelconque de temps successifs égaux, les espaces sont alors l'un à l'autre comme les entiers impairs à l'unité, c'est-à-dire comme 1, 3, 5, 7*"[14] (*cf.* figure 4, laquelle

---

[12] On mesure bien, par cette équivalence entre une propriété analytique et une propriété géométrique, tout l'écart entre le travail d'ORESME et celui de ses prédécesseurs d'Oxford et *a fortiori* avec les représentations graphiques de phénomènes physiques bien déterminés comme la variation des latitudes des planètes (à partir de la longitude). Ces dernières sont déjà décrites dans un manuscrit du XI$^{\text{ème}}$ siècle. (*Cf.* page 392, Tome 1 du livre de A. C. CROMBIE, *Histoire des Sciences de Saint-Augustin à Galilée*, trad. française, Paris, P.U.F., 1958.)

[13] G. GALILEI, Opere, FAVARO e LONGO, Firenze, Vol. 10, p. 115.

[14] G. GALILEI, *Discorsi e dimostrazioni matematiche intorno à due nuove scienze*, Leyde, 1638. (Opere, Vol. 8, cité en [13].) Voir aussi la nouvelle édition, A. CARUGO et L. GEYMONAT, Torino, Boringhieri, 1958. Le texte cité vient en corollaire 1 de la proposition II, Théorème 2 à la Troisième Journée intitulée, du mouvement local. La proposition énonce que les espaces parcourus sont comme les carrés des temps:

"*Si aliquod Mobile motu uniformiter accelerato descendat ex quiete*; *spatia, quibuscunque temporibus ab ipso peracta, sunt inter se in duplicata ratione eorundem temporum*: *nempe ut eorundem temporum quadrata*".

L'expérience quant à la mesure du mouvement d'une boule sur un plan incliné fut très se trouve chez GALILÉE et qui parle d'elle-même). De fait ce résultat peut apparaître comme une conséquence numérique, d'alleurs envisagée déjà par ORESME. Mais GALILÉE donne à ce résultat une forme fonctionnelle que nous fournissons en notations modernes. Si l'on garde $x(t)$ comme en (6) pour désigner les espaces parcourus, mais avec $x(0) = 0$

$$\frac{x[(n+1)t] - x(nt)}{x(nt) - x[(n-1)t]} = \frac{2n+1}{2n-1}, \tag{7}$$

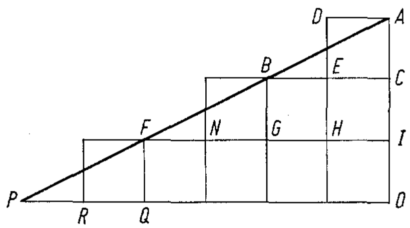

Fig. 4

équation valable pour t ($t > 0$) et tout entier n, positif ou nul. Ici le temps t apparaît comme une unité, d'ailleurs arbitraire comme il l'est bien souligné. Par ce choix arbitraire, une grandeur continue comme $x(t)$ est ramenée à satisfaire une condition discrète, exprimée en entiers. Habile manipulation des proportions! L'étonnante application que fait GALILÉE de la résolution de l'équation fonctionnelle (7) est plus importante, pour notre propos relatif aux équations fonctionnelles et au concept de fonction, que la méthode utilisée pour parvenir à cette résolution, réminiscence ou non d'ORESME et d'une tradition mécaniste parisienne.

GALILÉE considère en effet que l'équation fonctionnelle (7) *caractérise* les mouvements uniformément accélérés (dont la vitesse au départ est nulle). Or cette propriété (7) est expérimentalement constatable, non peut-être directement dans la chute libre des corps mais au moins par l'astucieux truchement physique de la descente d'une boule de bronze sur un plan incliné, descente suffisamment ralentie pour se prêter à la mesure effective. Et GALILÉE réalise l'expérience en question, peut-être dès 1604, en tout cas bien avant 1638, date à laquelle l'expérience est explicitement décrite[14]. Cette expérience prouve la validité de la relation fonctionnelle (7) pour la chute des corps. La vérification expérimentale est d'autant plus sûre, malgré les mesures imprécises que les premiers rapports d'entiers sont bien séparés.

Donc la résolution de l'équation fonctionnelle (7) a une conséquence dynamique considérable. Mais les conséquences quant à la méthode scientifique, et donc quant à la philosophie du $17^{\text{ème}}$ siècle, sont tout aussi considérables. Pour nous en convaincre, citons le propos fort que tient SALVIATI, l'un des personnages inventés

imprécise, faute d'une horloge convenable. La mesure du temps était effectuée au moyen de la quantité d'eau écoulée à partir de grosses bonbonnes. On notera que le choix d'une certaine quantité d'eau écoulée est libre et correspond à l'unité de temps t.

par GALILÉE afin de donner de la vie à son exposé de 1638. SALVIATI vient de mentionner différentes appréciations suggérées par divers auteurs pour expliquer la cause de l'accélération de la chute des corps; brutalement il les ridiculise:

"*Maintenant, toutes ces fantaisies, et d'autres encore, doivent être examinées, mais cela n'en vaut vraiment pas la peine. Désormais le but de notre auteur est seulement de rechercher et de prouver certaines des propriétés du mouvement accéléré (quelle que puisse être la cause de cette accélération) en appelant ainsi un mouvement dont, après avoir quitté le repos, les moments de la vitesse s'accroissent proportionnellement avec le temps. Ce qui revient à dire que, dans des intervalles égaux de temps, le corps reçoit des accroissements égaux de vitesse. Et si nous trouvons que les propriétés, lesquelles seront démontrées plus tard, sont réalisées pour des corps tombant en chute libre et accélérée, nous serons en droit de conclure que la définition supposée inclut un mouvement tel que celui de la chute des corps pesants et que leur vitesse s'accroît avec le temps et la durée du mouvement*"[15].

Déroulons le cheminement précis de GALILÉE dans sa démonstration (en notations modernes), car il s'agit d'un paradigme de l'induction en physique à partir des mathématiques. Il *définit* d'abord un mouvement uniformément accéléré comme celui satisfaisant

$$\frac{dx}{dt} = at + b,$$

et en outre il prend $b = 0$, c'est-à-dire une vitesse initiale nulle, puisqu'il s'agit de l'appliquer à la chute d'un corps lâché sans impulsion au départ. Les choses sont donc bien claires à condition d'accepter sa définition controversable de la vitesse instantanée. D'une part, GALILÉE établit un résultat d'intégration, c'est-à-dire une autre écriture pour un mouvement accéléré, dans le cas $b = 0$, c'est-à-dire la forme

$$x(t) = \frac{at^2}{2} + c. \tag{8}$$

Puis GALILÉE déduit *mathématiquement* une propriété particulière de ce mouvement uniformément accéléré, à savoir l'équation fonctionnelle (7) qu'il énonce verbalement, ce qui lui est d'autant plus familier que l'énoncé est une belle utilisation

---

[15] Les dialogues entre SALVIATI, SAGREDO et SIMPLICIO, les trois héros inventés par GALILÉE, sont en italien. Le passage de SALVIATI est situé à la Troisième Journée de l'ouvrage cité en [14]. Le texte italien original est le suivant:

"*Per ora basta al nostro Autore, che noi intendiamo, che egli ci vuole investigare, e dimostrare alcune passioni di un moto accelerato (qualunque si sia la causa della sua accelerazione) talmente che i momenti della sua velocità vadano accrescendosi dopo la sua partita dalla quiete con quella semplicissima proporzione, con la quale cresce la continuazion del tempo, che è quanto dire, che in tempi equali si facciano eguali additamenti di velocità. E se s'incontrerà, che gli accidenti, che poi saranno dimostrati si verifichino nel moto de i gravi naturalmente descendenti, & accelerati, potremo reputare, che l'assunta definizione comprenda cotal moto de i gravi, e che vero sia che l'accelerazione loro vadia crescendo secondo, che cresce il tempo, e la durazione del moto*".

de la théorie des proportions. D'autre part, GALILÉE considère explicitement que cette équation (7) est *caractéristique* du mouvement uniformément accéléré. Ce qui réalise donc *pratiquement* une résolution de l'équation fonctionnelle (7). GALILÉE ne démontre pas mathématiquement l'équivalence entre (8) et (7) mais seulement que (8) implique (7). Il induit pourtant le mouvement accéléré de la vérification de (7). En conséquence, la vérification *expérimentale* de l'équation fonctionnelle (7), pour une unité de temps t quelconque et quelques entiers successifs n, *équivaut* à la dépendance quadratique de la distance parcourue en fonction du temps, formulée par (8). La loi de la chute des corps est établie[16]. L'analyse est passée du "*secundum imaginationem*" au "*secundum cursum naturae*", de l'imagination créative à l'ordre naturel.

De même que pour le théorème de Merton, on aura noté la combinaison d'un résultat de calcul intégral et de résultats sur les équations fonctionnelles.

Avec ORESME, l'idée fonctionnelle permettait une lecture analytique des propriétés des fonctions affines sous la forme d'une équation fonctionnelle. Avec GALILÉE, *l'équivalence* est poussée à d'autres fonctions et elle sert en outre de *pivot* entre la démarche expérimentale et la conclusion théorique. C'est indéniablement mettre les équations fonctionnelles à un très haut niveau d'intérêt.

4. La relation fonctionnelle du logarithme

De fait, à partir de GALILÉE, lequel proclamait que "*la nature est écrite en langue mathématique*"[17], l'idée de relation fonctionnelle prend une réelle extension

[16] GALILÉE ne démontre pas mathématiquement l'équivalence de l'équation fonctionnelle (7) et de la forme (8). C'est-à-dire que GALILÉE ne résout pas explicitement l'équation fonctionnelle (7) Explicitement, il considère cependant que (7) équivaut à (8). GALILÉE établit certes que toute fonction donnée par (8) satisfait (7). La réciproque peut s'obtenir moyennant une hypothèse de régularité de la fonction $x(t)$, régularité d'autant plus supposée dans le contexte galiléen que $\dfrac{dx}{dt}$ existe. Mais cette remarque nous fait évidemment sortir de l'histoire. A titre de curiosité, on peut indiquer:

$$\frac{x[(n+1)t]-x(nt)}{x(t)-x(0)}=\frac{x[(n+1)t]-x(nt)}{x(nt)-x[(n-1)t]}\cdot\frac{x(nt)-x[(n-1)t]}{x[(n-1)t]-x[(n-2)t]}\cdots\frac{x(2t)-x(t)}{x(t)-x(0)}.$$

On utilise maintenant l'équation fonctionnelle (7), d'où

$$\begin{aligned}\frac{x[(n+1)t]-x(nt)}{x(t)-x(0)}&=\frac{2n+1}{2n-1}\cdot\frac{2n-1}{2n-3}\cdots\frac{3}{1}\\&=2n+1.\end{aligned}$$

Avec $y(t)=x(t)-x(0)$, pour tout $t>0$, on a pour tout entier $n\geqq 0$ la relation

$$\begin{aligned}y[(n+1)t)]&=(2n+1)\,y(t)+y(nt)\\&=[(2n+1)+(2n-1)+\ldots+3+1]\,y(t)=(n+1)^2\,y(t).\end{aligned}$$

On déduit $y\left(\dfrac{n}{m}t\right)=\left(\dfrac{n}{m}\right)^2 y(t)$ et si y est continue, par exemple, $y(t)=\dfrac{at^2}{2}$, $a=2y(1)$. Soit $x(t)=\frac{1}{2}at^2+C$, $C=x(0)$, ce qui fournit la relation (8).

[17] La citation en question se trouve dans *Il Saggiatore*, question 6. Un ouvrage

poursuivie tout au long du XVII$^{\text{ème}}$ siècle. D'une part sur le plan conceptuel grâce à la ferme liaison établie par DESCARTES (1596–1650) et FERMAT (1601–1665) entre algèbre et géométrie, ce qui permet une description systématique des courbes algébriques considérées comme une relation polynomiale entre une abscisse et une ordonnée. D'autre part, par la création de nouveaux types non algébriques de relations fonctionnelles, relations souvent obtenues par des séries infinies de puissances dans la dernière partie du 17$^{\text{ème}}$ siècle. Mais avant même la découverte de ces développements infinis, on avait souvent remarqué l'analogie entre les propriétés multiplicatives de la suite géométrique (ou d'une partie de cette suite) pour une valeur a positive (différente de 0 ou de 1)

$$\ldots, a^{-1}, 1, a, a^2, a^3, a^4, \ldots, a^n, \ldots$$

et les propriétés additives de la suite arithmétique

$$\ldots, -1, 0, 1, 2, 3, 4, \ldots, n, \ldots$$

Analogie qui repose sur la loi opératoire

$$a^{n+m} = a^n a^m. \tag{9}$$

N. BOURBAKI[18] fait même remonter cette loi à la proposition 11 du livre IX d'EUCLIDE, au moins pour les coefficients n et m entiers. Enonçons cette proposition.

"*Dans toute suite de nombres en proportion continue à partir de l'unité, le terme suivant l'unité divise le plus grand terme de la suite et le quotient appartient à la suite*".

La démonstration euclidienne revient à $a^n/a = a^{n-1}$, c'est-à-dire à la stabilité

---

publié en 1623. Voit *Opere*, vol. 6, cité en [13]. Cette citation, nous semble particulièrement appropriée pour illustrer l'intérêt des équations fonctionnelles, et repérer la démarche inductive de la physique mathématisée.

"*La Filosofia è Scritta in questo grandissimo libro, che continuamente ci sta aperto innanzi a gli occhi (io dico l'universo) ma non si può intendere se prima non s'impara à intender la lingua, e conoscer i caratteri, ne'quali è scritto. Egli è scritto in lingua matematica. Li caratteri son triangoli, cerchi et altre figure geometriche, senza i quali mezziè impossibile a intenderne umanamente parola, senza questi è un aggirarsi vanamente per un'oscuro labirinto*".

On peut traduire

"*La philosophie est écrite dans ce très vaste livre qui est éternellement ouvert devant nos yeux – je veux dire l'Univers – mais on ne peut le lire avant d'avoir appris la langue et s'être familiarisé avec les caractères dans lesquels elle est écrite. Elle est écrite en langue mathématique, et ses lettres sont des triangles, des cercles et d'autres figures géométriques, moyens sans lesquels il est humainement impossible de comprendre un seul mot, sans lesquels l'on erre en vain dans un obscur labyrinthe*".

[18] N. BOURBAKI, *Eléments d'histoire des mathématiques*, N$^{\text{elle}}$ édition augmentée, Hermann, 1969 (page 196).

de la suite par division par la raison. Un corollaire (πόρισμα) précise que le rang du quotient d'un terme quelconque par un diviseur est le même dans l'ordre inverse que celui du diviseur dans le dénombrement naturel. Soit

$$\frac{a^n}{a^m} = a^{n-m}; \quad n \geqq m > 0.$$

En fait, dans cette proposition, l'accent est mis sur l'aspect du dénombrement. Mais je ne suis pas sûr qu'il faille y trouver l'aspect opératoire de (9). En tout cas, la pleine compréhension de (9) fut longue. Dix-sept siècles plus tard, N. ORESME est conscient de la relation (9). D'ailleurs, dans la plupart des manuels ultérieurs, la théorie des proportions fait un parallèle qui se veut pédagogique entre une proportion arithmétique, c'est-à-dire quatre grandeurs A, B; C et D telles que $A - B = C - D$ et une proportion géométrique, c'est-à-dire quatre grandeurs $A, B; C, D$ telles que $A/B = C/D$. D'un côté interviennent les différences, de l'autre les quotients. Et avec les proportions continues des deux genres, c'est-à-dire les suites arithmétique et géométrique, c'est bien la relation (9) qui est à l'oeuvre dans l'analogie. Pourtant, dans ces mêmes manuels, ce n'est pas tellement l'analogie par (9) sur laquelle on insiste, mais bien plutôt sur la similitude des règles de calcul codifiées et souvent complexes à nos yeux et qui constituent l'algèbre des proportions[19]. Le langage analogique risque donc d'être trompeur qui nous ferait croire que (9) résume au Moyen-Age le lien entre progression arithmétique et progression géométrique. Ainsi, ORESME ayant introduit les exposants fractionnaires positifs, n'étend que partiellement la relation (9), par exemple sous la forme

$$a^{\frac{2}{n}} = (a^2)^{\frac{1}{n}} = a^{\frac{1}{n}} a^{\frac{1}{n}}.$$

Une pleine conscience de l'analogie fournie par (9) devrait le conduire plus loin, jusqu'à étendre (9) à tous les n, m rationnels. Plus tard N. CHUQUET (1445–1500) dans un texte, *le triparty en la science des nombres*, rédigé en 1484 mais fort peu diffusé[20], fait usage d'exposants négatifs et prend $a^0 = 1$, un choix remarquable, et qui revient à s'appuyer fortement sur la relation (9) prolongée pour toutes les valeurs entières de n et m, sans toutefois que cela soit explicitement indiqué. Au cours de l'histoire, les équations fonctionnelles seront souvent et judicieusement utilisées pour justifier des écritures: c'est un rôle lié aux définitions que nous n'aurons malheureusement pas le temps d'explorer malgré son intérêt.

---

[19] C'est là le danger de toute théorie, comme celle des proportions, dont on perd de vue le but (la manipulation algébrique des grandeurs continues) pour ne plus y voir qu'une technique traditionnelle de calcul dont les tours font figure de théorèmes. Tard dans le 17ème siécle, on trouvera de nombreux manuels mathématiques où l'analogie entre suite arithmétique et géométrique n'est que l'occasion de justification de ces tours. On peut citer par exemple trois ouvrages qui eurent à la fois du succès et de l'influence: A. ARNAULD, *Nouveaux élémens de géométrie*, Paris, 1667; J. PRESTET, *Elémens des Mathématiques*, Paris, 1675; M. OZANAM, *Cours de Mathématiques*, Paris, 1693. Ces manuels ratent l'idée fonctionnelle par une vision sclérosée, scolaire, des proportions.

[20] N. CHUQUET, *Le Triparty en la science des nombres*, 1ère édition imprimée, Paris, 1880.

STIFEL (1487–1567) est explicite dans son *Arithmetica Integra* de 1544 puisque l'on y trouve le tableau suivant[21]:

-3	-2	-1	0	1	2	3	4	5	6
$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	1	2	4	8	16	32	64

D'où la règle: "*Qualiscunque facit progressio Geometricae multiplicando et dividendo talis facit progressio Arithmeticae addendo et subtrahendo*".

"*Ce qui se fait par multiplication et division dans les progressions géométriques se fait par addition et soustraction dans les progressions arithmétiques*".

Mais là encore le langage est en avance sur la perception réelle de (9) en dehors des entiers relatifs. Cette perception va se faire progressivement au cours du XV$^{\text{ème}}$ et du XVI$^{\text{ème}}$ siècle et il serait intéressant de relever systématiquement les utilisations de (9), sous tous ses avatars, dans les nombreux manuels de cette période.

L'horloger suisse JOOST BÜRGI (1552–1632), inspiré par l'oeuvre de STIFEL, construit en tout cas le logarithme à partir de la relation (9) et par une intuition de continuité fonctionnelle. Son œuvre ne parut que fort tardivement à Prague en 1620 (*Progress Tabulen*). Entre temps, l'écossais JOHN NAPIER (1550–1617) l'avait devancé. En fait, NAPIER, pour construire le logarithme, se sert d'une image mécaniste de nature théorique. Il compare deux mouvements rectilignes. Ce qui lui permet, dans la construction, de tirer parti de la continuité liée au mouvement, c'est-à-dire au déroulement du temps. L'un des mouvements utilisé est en effet uniforme, l'autre est à vitesse proportionnelle à la distance à un point fixe[22]. On peut reconstruire en termes modernes l'argument de NAPIER afin de manifester le rôle des dépendances fonctionnelles. NAPIER envisage deux mouvements rectilignes gouvernés par les équations différentielles suivantes, avec les conditions initiales indiquées:

$$\frac{dY}{dt} = V_0, \quad Y(0) = 0; \quad \frac{dX}{dt} = -V_0\, 10^{-7} X, \quad X(0) = 10^7.$$

La dépendance est à établir entre Y et X c'est-à-dire en éliminant le temps. Soit

---

[21] M. STIFEL, *Arithmetica integra*, Nuremberg, 1544, (page 249).

[22] La première moûture du travail de NAPIER est descriptive sur l'utilisation de tables numériques. Elle parut en 1614, *Mirifici Logarithmorum Canonis Descriptio*, Edinburgh. Une explication de la méthode d'obtention des tables, le procédé numérique, et une justification cinématique qui nécessite d'ailleurs une reconstruction pour être compréhensible, parut en 1619 à titre posthume, *Mirifici Logarithmorum Canonis Constructio*, Edinburgh.

Pour plus de précision, on pourra consulter: E. W. HOBSON, *John Napier and the invention of logarithms*. Cambridge Univ. Press, 1914; C. G. KNOTT, *Napier tercentenary memorial volume*, Longmans Green, 1915; et trouver les textes de NAPIER dans F. MASÈRES, *Scriptores Logarithmici*, 6 vol., London, 1791–1807.

en termes modernes,

$$Y = 10^7 \operatorname{Log} \frac{10^7}{X}.$$

Si l'on manifeste cette liaison par $Y = M(X)$, on déduit les relations fonctionnelles

$$M(xy) = M(x) + M(y) - 10^7 \operatorname{Log} 10^7,$$

$$M\left(\frac{x}{y}\right) = M(x) - M(y) + 10^7 \operatorname{Log} 10^7.$$

De telles relations ont l'inconvénient de termes constants supplémentaires. Toutefois dès que $x_1, x_2, \ldots, x_n$ sont des nombres en progression géométrique, la deuxième formule établit que les nombres $M(x_1), M(x_2), \ldots, M(x_n)$ sont en progression arithmétique. C'est bien la permanence cherchée par NAPIER: non par une exploitation directe de (9), mais la mise en oeuvre du parallélisme entre les deux sortes de progression. Le vocabulaire élaboré par NAPIER est fidèle à la construction réalisée: le logarithme, utilisation du λόγος, rapport, et de ἀριτμὸς, nombre entier. Par de judicieuses approximations que l'image mécaniste soutient en évoquant des mouvements uniformes par morceaux, NAPIER constate que l'on peut remplacer 10^7 par $-1/\operatorname{Log}(1 - 10^{-7})$, pour utiliser les notations actuelles, ce qui ne provoque qu'une erreur inférieure à 10^{-7} pour les valeurs choisies couvrant la table. Ce changement revient à écrire une nouvelle dépendance fonctionnelle puisque l'approximation faite fournit

$$\operatorname{Log} \frac{10^7}{X} = \frac{Y}{10^7} \sim -Y \operatorname{Log}(1 - 10^{-7}) = \operatorname{Log} \frac{1}{(1 - 10^{-7})^Y}.$$

Soit la forme,

$$X = 10^7(1 - 10^{-7})^Y.$$

Posant cette fois $Y = N(X)$ pour manifester la dépendance fonctionnelle, on constate l'équation fonctionnelle,

$$N(xy) = N(x) + N(y) + \frac{\operatorname{Log} 10^7}{\operatorname{Log}(1 - 10^{-7})}.$$

Il y a encore un terme constant, mais la relation possède aussi la même permanence, échangeant des progressions géométriques en progressions arithmétiques. Aussitôt NAPIER se sert fondamentalement de cette permanence pour construire sa table par interpolation[22].

Toujours en termes modernes, la liaison qu'utilise BÜRGI entre X et Y est assez voisine de celle de NAPIER,

$$X = (1 + 10^{-4})^{10^4 Y}.$$

Posant ici $Y = B(X)$, l'équation fonctionnelle satisfaite par B, pour le produit, ne possède pas de termes constants, puisque l'on a,

$$B(xy) = B(x) + B(y).$$

En particulier $B(1) = 0$ et $B((1 + \frac{1}{10^4})^{10^4}) = 1$. Numériquement, on n'est pas loin du logarithme naturel de base e. Par opposition, chez NAPIER, on a $N(10^7) = 0$ et $N(10^7 - 1) = 1$.

Ce qui est sûr, c'est que la relation calculatoire (9) a pris une certaine tournure fonctionnelle et permis la *création* d'un être nouveau, le logarithme, c'est-à-dire une correspondance nouvelle. Et ainsi H. BRIGGS (1561–1631), dans sa construction d'une table de logarithmes de base 10, part fondamentalement de la relation fonctionnelle

$$\log xy = \log x + \log y, \tag{10}$$

qui n'est que la lecture inverse de la relation (9) convenablement étendue[23]. En effet, BRIGGS semble poser *a priori* l'existence d'une correspondance entre x et ce que l'on nomme aujourd'hui $\log x$ de sorte que (10) soit satisfaite, donc en particulier que $\log 1 = 0$, et que $\log 10 = 1$. Itérant (10), BRIGGS calcule par des extractions successives de racines carrées, et avec la précision considétable (10^{-30}) qu'il s'est fixée, une puissance en $1/2^n$ de 10 suffisamment voisine de 1. Il peut alors constituer une table de ses logarithmes par utilisation systématique de l'équation fonctionnelle (10). L'image mécaniste de NAPIER prédomine toujours pour assurer l'existence d'une correspondance satisfaisant (10), à tel point que pendant longtemps les logarithmes ne sont pas définis par exponentiation. *Il n'y a donc pas réflexion théorique poussée sur l'équation* (9) *et l'équation fonctionnelle* (10) *apparaît avant tout comme utilitaire.* Il faut dire que les idées fonctionnelles de BÜRGI et de BRIGGS furent vite bloquées par le flou épistémologique concernant les nombres réels, sans parler de la continuité.

L'évolution des logarithmes en dehors de leur utilisation numérique, vite maîtrisée, est d'être rattachée sur le plan théorique à ce qui va devenir le calcul intégral. Et ce rattachement essentiel change le point de vue et réduit la portée fonctionnelle créatrice des logarithmes en ce sens que ceux-ci apparaissent comme des aires sous l'hyperbole équilatère, une courbe familière acceptée et décrite par les Anciens. GRÉGOIRE DE SAINT-VINCENT (1584–1667) dont *l'Opus Geometricum* est écrit dès 1625 mais tardivement publié en 1647, dispose de la propriété de ces aires. A la proposition 109 (Page 986) de son gros ouvrage[24], il énonce: "*Quand*

---

[23] Voir *Mirifici Logarithmorum Canonis Descriptio* de J. NAPIER (Annotationes aliquot, D. HENRICI BRIGGII), Londres 1620 et *Arithmétique logarithmique* de BRIGGS parue en 1628, traduction de *Arithmetica logarithmica* publiée à Londres en 1624.

[24] GREGORIO A S[te] VINCENTO, *Opus geometricum quadraturae circuli et sectionem coni*, 1647, Antverpiae. (Titre exact: *Problema austriacum. Plus ultra quadratura circuli.* Un ouvrage posthume de 1668 s'intitule *Opus geometricum posthumum ad Mesolabium*). Le résultat de la proposition 109 figure déjà à la proposition 673 du volume 13 des manuscrits de GRÉGOIRE DE SAINT-VINCENT, ce qui le date de la période 1617–1625. (Voir H. VAN LOOY, Chronologie et analyse des manuscrits mathématiques de Grégoire de Saint-Vincent, Archivum Hist. Societatis Jesu, Vol. **49**, 1980, p. 279–303). Après une attaque d'apoplexie en 1627, GRÉGOIRE DE SAINT VINCENT perdit beaucoup de son acuité, mais il publia tardivement faute d'autorisation de ses supérieurs jésuites. Voir pour l'analyse de l'Opus Geometricum: J. E. HOFMANN, Das Opus Geometricum des Gregorius a S. Vincento und seine Einwirkung auf Leibniz, Abh. Preussischen Akad. der Wissenschaften 1941, n° 13, 79 pages.

les abscisses croissent en progression géométrique, les aires (obtenues sous l'hyperbole) *croissent en progression arithmétique*". Et la proposition 130 établit la réciproque. Si des points $x_1, x_2, \ldots, x_n, \ldots$ sont tels que les aires découpées sous l'hyperbole sont égales, les x_i et donc les y_i sont en progression géométrique

$$\frac{y_1}{y_2} = \frac{y_2}{y_3} = \frac{y_3}{y_4} = \ldots$$

Autrement dit, les abscisses suivent une progression géométrique tandis que les aires correspondantes additionées a, $2a$, $3a, \ldots$ suivent une progression arithmétique (figure 5). Les démonstrations se font par la méthode d'exhaustion[25]. C'est A. DE SARASA (1618–1667), élève de GRÉGOIRE DE SAINT-VINCENT, qui interprète ce double aspect d'une progression géométrique des abscisses et arithmétique des aires. Il remarque que le résultat de GRÉGOIRE DE SAINT-VINCENT conduit explicitement à l'équation fonctionnelle (10) quant aux aires. En notations modernes, appelons $A(x)$ l'aire sous l'hyperbole entre l'abscisse 1 et l'abscisse x. Soient x et y deux abscisses ($1 < x < y$, pour fixer les idées). Soit z la moyenne géométrique de x et y, c'est-pà-dire $\dfrac{x}{z} = \dfrac{z}{y}$ (ou $z = \sqrt{xy}$). D'après le résultat de GRÉGOIRE DE SAINT-VINCENT,

$$A(z) - A(x) = A(y) - A(z),$$

donc $A(z) = \frac{1}{2}(A(x) + A(y))$. En particulier, avec $x = 1$ et $y = x$, $A(\sqrt{x}) =$

[25] Indiquons succinctement le principe de la démonstration très archimédienne, mais dont nous ne détaillons pas les différentes étapes (raisonnement par l'absurde par exemple).

Supposons x_1, x_2 et x_3 en progression géométrique. Un calcul simple établit l'égalité des aires des rectangles x_1, x_2, y_2, Y_1 et x_2, x_3, y_3, Y_2 (voir figure A). Rappelons que la courbe représentée est une branche de l'hyperbole $xy = 1$. On prend x_1' comme moyenne géométrique de x_1 et x_2, puis x_1'' comme moyenne géométrique de x_2 et x_3. L'égalité analogue des quatre rectangles a lieu puisque $x_1, x_1', x_2, x_1'', x_3$ sont en progression géométrique. On insère ainsi successivement des moyennes géométriques pour exhauster à la limite les aires curvilignes x_1, x_2, y_2, y_1 et x_2, x_3, y_3, y_2, lesquelles sont donc égales.

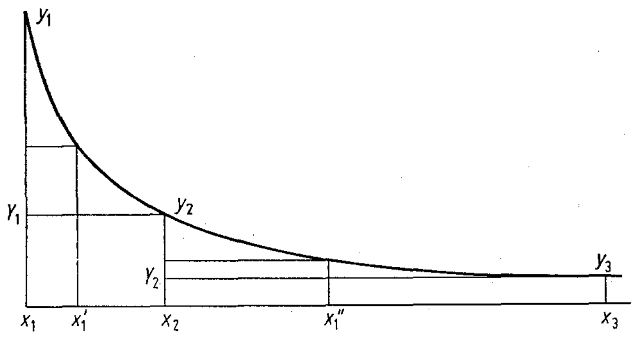

Fig. A

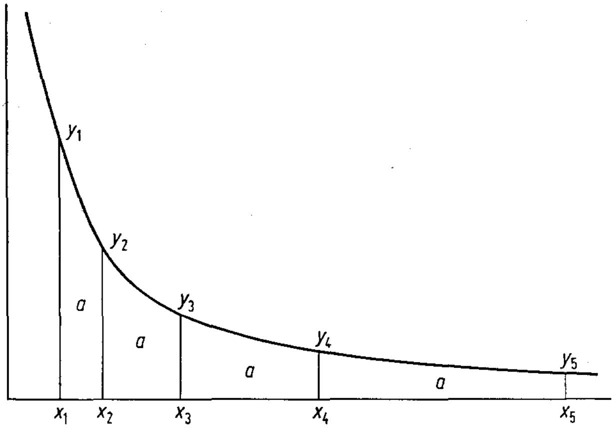

Fig. 5

$A(x)/2$. D'où

(11) $$A(xy) = A(x) + A(y).$$

Et A. DE SARASA conclut (Proposition 10 de [26]) "*Unde hae superficies supplere possent locum Logarithmorum*", ("*donc ces aires peuvent servir de logarithmes*"). Il s'agit bien là d'une *résolution de l'équation fonctionnelle* (10). D'une part, on en construit explicitement une solution grâce aux aires sous l'hyperbole. D'autre part, de SARASA indique que l'équation fonctionnelle (10) est caractéristique des logarithmes avec liberté de la base. Car DE SARASA ne spécifie pas la base en jeu qui est ici celle des logarithmes ultérieurement qualifiés de naturels par N. MERCATOR en 1668. Et de fait, l'opuscule de DE SARASA fit un certain bruit, bien indépendamment de cette caractérisation du logarithme, car se situant dans le cadre du problème de la quadrature du cercle, problème qui entâchait la réputation de GRÉGOIRE DE SAINT-VINCENT[27]. Faut-il mettre sur le compte de cette

[26] A. DE SARASA, *Solutio problematis a R. P. Marino Mersenno minimo propositi*, 1649, Antverpiae.

[27] Le texte de A. DE SARASA répond à un problème du Père MERSENNE, ce correspondant infatigable entre savants. L'ouvrage de GRÉGOIRE DE SAINT-VINCENT prétendait avoir résolu le problème de la quadrature du cercle (construction à la règle et au compas). Une erreur sérieuse n'apparaît qu'à la page 1120 du texte et ne fut d'ailleurs pas remarquée des contemporains (sauf de HUYGENS). Mais l'on doutait de l'exactitude du résultat final et du coup l'on méconnaissait les qualités remarquables de l'ouvrage, bien que tardif par rapport à sa conception, donc nécessairement dépassé. MERSENNE proposa imprudemment dans ses réflexions physico-mathématiques (*Novarum observationum physico-mathematicarum*, tomus III, Paris, 1647) de seulement résoudre, pour cette quadrature du cercle, le problème suivant:

"*Connaissant trois grandeurs quelconques, rationnelles ou irrationnelles et le logarithme de deux d'entre elles, en déduire le logarithme de la troisième*".

Un simple problème de détermination d'une base de logarithme. A. DE SARASA n'a pas de peine à résoudre ce problème au moyen des résultats de GRÉGOIRE DE SAINT-VINCENT et donc de justifier ce dernier quant à sa prétention d'avoir résolu le problème de la quadrature du cercle — à condition d'admettre l'implication de MERSENNE!

réputation ébranlée le peu d'intérêt apparent pour la caractérisation fonctionnelle du logarithme et la perte d'allant de l'idée fonctionnelle? Sans doute, mais il faut pousser plus loin l'analyse épistémologique. Nous prendrons d'abord l'exemple de FERMAT qui, dans un texte[28] vraisemblablement écrit vers 1660, parvient à réaliser les quadratures de courbes $y^{\alpha}x^{\beta} = a$. Examinons sa façon de traiter le cas $\alpha = 1$, $\beta = +2$. Dès le départ, FERMAT nous indique que sa méthode repose sur l'utilisation de la progression géométrique "*très féconde en quadratures*". Il prend des points *G, H, O, M, R etc.* en progression géométrique (*cf.* figure 6),

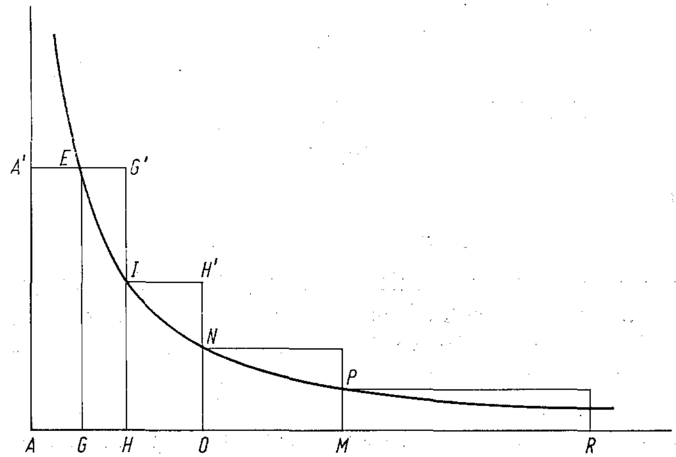

Fig. 6

disons selon une raison $1/r$ ($r < 1$). FERMAT indique ensuite qu'en prenant r suffisamment voisin de 1, on peut "adégaler" un rectangle tel que $GHG'E$ et le trapèze curviligne $GHIE$. Il explique en effet que l'on peut "*facilement employer la méthode d'Archimède de réduction à l'impossible, par circonscriptions et inscriptions*". Et il ajoute: "*Il suffit de faire cette remarque une fois pour ne pas s'obliger à revenir et à insister constamment sur un artifice bien connu de tous les géomètres*"[28]. Par manipulation des proportions, il constate que les rectangles successifs $GHG'E$, $HOH'I$ *etc.* forment une progression géométrique de raison r. Il suffit de sommer cette progression. En écrivant $r = AG/AH$, il vient facilement pour la valeur des sommes des aires des rectangles depuis le rectangle $HOH'I$, l'égalité à l'aire du rectangle $AGEA'$. Faisant alors tendre la raison vers 1, c'est-à-dire H vers G, FERMAT déduit l'égalité du rectangle $AGEA'$ à l'aire curviligne déterminée par l'axe des abscisses, la droite GE et la courbe. De même, si $\alpha = 1$, $\beta = +n$, pour un entier $n > 1$, l'aire du rectangle analogue $AGEA'$ vaut $(n - 1)$ fois l'aire curviligne. FERMAT envisage naturellement le cas $n = 1$, c'est-à-dire le cas de l'hyperbole. Constatant l'égalité des rectangles successifs

[28] P. FERMAT, *Oeuvres*, ed. TANNERY et HENRY, Paris, Gauthier-Villars, (1891–1912), Tome I, p. 255–259. FERMAT ne travaille pas nécessairement avec des axes orthogonaux, mais pour simplifier nous avons parlé de rectangles au lieu de parallélogrammes.

$GHG'E$, il déduit que sa méthode ne convient pas puisqu'elle ne fournit pas une aire déterminée égale à l'aire curviligne (qui est infinie).

Si la méthode est habile, il me paraît essentiel de faire sentir la différence, du point de vue fonctionnel, par rapport à GRÉGOIRE DE SAINT-VINCENT complété par DE SARASA. FERMAT trouve un équivalent géométrique pour l'aire curviligne que nous notons $f(x_0) = \int_{x_0}^{+\infty} \frac{dx}{x^2}$ pour $x_0 > 0$ (abscisse de G), mais n'étudie pas du tout la correspondance entre x_0 et cette aire. Or la propriété fonctionnelle de f est notable et facilement vérifiable géométriquement, à condition toutefois de vouloir chercher une telle relation fonctionnelle,

$$f(x_0 x_0') = f(x_0) f(x_0').$$

Plus généralement avec $f_n(x) = \int_{x_0}^{+\infty} \frac{dx}{x^n}$, il vient $f_n(x_0 x_0') = (n-1) f_n(x_0) f_n(x_0')$ pour $n > 1$. Rien de tel chez FERMAT au contraire de la relation fonctionnelle avec $A(x_0) = \int_1^{x_0} \frac{dx}{x}$ établie par de SARASA. Et ce n'est pas parce que FERMAT travaille avec une aire qui s'étend à l'infini en abscisse. Lorsqu'il applique sa méthode à $\int_0^{x_0} \frac{dx}{x^{1/n}}$ pour $n > 1$, il n'indique pas plus de propriété fonctionnelle. Mais allons plus loin dans l'analyse de la preuve de FERMAT. C'est bien la manipulation des progressions géométriques qui est à l'oeuvre. Dès le départ en effet, la définition des courbes en jeu est fournie dans le seul langage des proportions. Si l'on prend deux points quelconques en abscisses G et H, et E et I les points correspondants de la courbe alors $\left(\frac{HI}{GE}\right)^\alpha = \left(\frac{AG}{AH}\right)^\beta$, α et β étant entiers ou inverses d'entiers. Cette définition par proportions me paraît alors essentielle pour la justification de l'adégalisation, "*cette vérité qu'il serait facile de confirmer par une démonstration plus prolixe menée à la façon d'Archimède*"[28]. Comme dit FERMAT lui-même, en effet, ce qui entre bien dans le cadre de la méthode d'exhaustion, c'est d'établir pour deux rectangles successifs quelconques $GHG'E$ et $HOH'I$ l'égalité de deux rapports,

$$\text{(A)} \qquad \frac{\text{aire curviligne } GHIE}{\text{aire curviligne } HONI} = \frac{\text{aire } GHG'E}{\text{aire } HOH'I},$$

lorsque AG, AH et AO forment une progression géométrique de raison $1/r$ avec $r < 1$ (*cf.* figure 6). Il suffit pour le voir d'adapter simplement la preuve de GRÉGOIRE DE SAINT-VINCENT (*cf.* [25]) en insérant entre G et H, puis entre H et O des moyennes géométriques J et K. La définition de la courbe, au moyen des proportions, établit à l'évidence une permanence, c'est-à-dire la transformation d'une progression géométrique sur les abscisses en une autre progression géométrique de raison généralement différente quant aux aires des rectangles correspondants. Et FERMAT le sait bien qui précise que lorsque $\beta/\alpha = 1$, la progression géométrique est triviale, de raison 1 "*et c'est précisément la différence qui fait tout le mystère de cette affaire*"[28]. Dès lors (Fig. 7), on constate aisément que ce rapport de l'aire du rectangle $GJJ'E$ à l'aire $HKK'I$ est égal au rapport des aires des rectangles $GHG'E$ et $HOH'I$. De même pour les rapports des aires des rectangles $JHL'L$ et $KOM'M$. Grâce à la loi opératoire des proportions $a/b = c/d = (a+c)/(b+d)$,

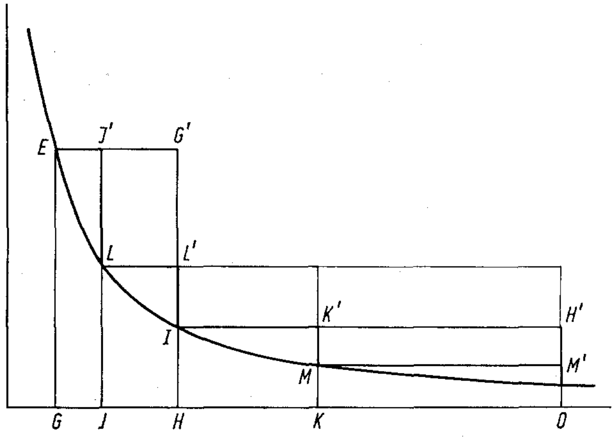

Fig. 7

on déduit l'égalité des rapports des aires $GHL'LJ'E$ et $HOM'MK'I$ au rapport des aires des rectangles $GHG'E$ et $HOH'I$. On en déduit (A) selon la méthode d'exhaustion. Puis, sans aucun passage à la limite, le résultat de FERMAT. Ce dernier préfère court-circuiter cette preuve en faisant tendre la raison vers 1, c'est-à-dire H vers G, ce qui *a priori* pose un problème sur la manière "*uniforme*" dont les rectangles lointains tendent vers les aires de la courbe. Ce problème est facilement résoluble grâce à la propriété fondamentale des aires des rectangles, cette permanence d'une progression géométrique. Or FERMAT préfère un énoncé géométrique et il passe ainsi à côté de l'explication de la relation fonctionnelle qui n'est que le constat de cette permanence. Il *ne cite nullement* GRÉGOIRE DE SAINT-VINCENT.

L'aspect fonctionnel n'est pourtant pas perdu. Vers 1667, I. NEWTON (1642–1727) utilisant l'hyperbole équilatère $y = 1/(x+1)$, note verbalement[29] ce qui correspond à l'équation fonctionnelle (11) et il en déduit une table des logarithmes des entiers. La quadrature de l'hyperbole, liée à cette propriété fonctionnelle (11), passionne J. GREGORY (1638–1675), et N. MERCATOR (1620–1687) en particulier dans sa *Logarithmotechnia* de 1668. Ce dernier établit le passage d'un système de logarithmes à un autre par proportionnalité et donc fournit toutes les solutions (régulières) de l'équation (11).

Cependant NEWTON et MERCATOR parviennent à un développement en série entière de $\operatorname{Log}(1+x)$. Ceci marque un tournant. En effet, l'accent est mis sur *les séries infinies* dont on découvre de nombreux exemples et la souplesse d'emploi en la dernière partie du 17$^{\text{ème}}$ siècle. Il suffit de songer au titre de l'ouvrage de

[29] Voir *The mathematical papers of I. Newton*, D. T. WHITESIDE (ed.), Cambridge Univ. Press, Vol. II, p. 184–189.

NEWTON *Methodus fluxionum et serierum infinitarum* où les fluxions apparaissent à égalité avec les séries infinies.[30] Cette nouvelle direction fit oublier les équations fonctionnelles et on n'en connaît pas d'utilisation nouvelle jusque vers la fin du 17ème siècle.

5. Détermination pratique des fonctions de Leibniz à Euler

Le calcul différentiel fondé dans la seconde moitié du 17ème siècle par NEWTON (1642–1627) et LEIBNIZ (1646–1716) vint fournir une considérable impulsion fonctionnelle dont faute de place nous ne parlerons pas car cela est mieux connu et s'éloigne de notre sujet[31]. Il faut noter que les mathématiciens anglais, après NEWTON, restèrent plus fidèles à la méthode des séries et ne prirent guère le biais fonctionnel nouveau, ce qui renforce notre appréciation du rôle inhibiteur des séries à ce stade, quant au développement fonctionnel. Mais l'essentiel pour notre propos est qu'un être mathématique nouveau, *la fonction*, soit susceptible d'un calcul et d'une recherche quand bien même les contours significatifs de sa définition restent flous ou variables d'un auteur à l'autre. LEIBNIZ dans un manuscrit de 1673, vraisemblablement le premier où le mot fonction intervient, marque bien la problématique. Il s'agit au moyen de propriétés différentielles, sur la tangente par exemple, de *déterminer* une courbe géométrique, c'est-à-dire une certaine relation entre une abscisse et une ordonnée. Le titre du manuscrit est explicite: "*Methodus tangentium inversa, seu de functionibus*", "*de la méthode inverse des tangentes, ou des fonctions*". Et le mot fonction s'inscrit donc plus dans le sens d'équation fonctionnelle, de recherche d'une relation satisfaisant certaines conditions, par l'étymologie latine, que dans une définition formelle de correspondance. Fleurissent en effet nombre de problèmes géométriques de caractérisation de courbes ou de surfaces au moyen de leurs éléments différentiels. On peut citer par exemple l'étude des développées des courbes par C. HUYGENS. S'y déploie toute l'habileté des géomètres de la fin du 17ème et du début du 18ème siècle, premiers utilisateurs du calcul différentiel. Et ceux-ci sont conduits lentement à une classification des équations différentielles selon les méthodes utilisées pour les résoudre. Ce sont les méthodes qui comptent, beaucoup plus que les applications géométriques. Ces méthodes touchent entre autres les équations différentielles linéaires du premier ordre, les cas de séparation des variables, les équations de JACQUES BERNOULLI résolues par changement de variable et ramenées au cas linéaire *etc*. Contentons-nous d'un exemple. Dans une lettre du 24 Octobre 1676 adressée à LEIBNIZ par l'intermédiaire d'OLDENBOURG, NEWTON décrit une

---

[30] L'ouvrage de NEWTON était écrit en 1671, mais ne fut pas publié. Il ne parut qu'en 1736 et en anglais, après la mort de NEWTON. Une traduction française, *Méthode des fluxions et des séries infinies*, fut effectuée par BUFFON en 1740. (*Cf.* réédition, librairie Blanchard, Paris, 1966.)

[31] Mais nous renvoyons le lecteur aux ouvrages classiques d'histoire du calcul différentiel et du calcul intégral, comme C. B. BOYER, *The History of the Calculus and its Conceptual Development*, New York, Réédition Dover, 1959; M. E. BARON, *The origins of the Infinitesimal Calculus*, London, Pergamon, 1969. Pour le concept de fonction à proprement parler, voir le texte très documenté cité en [2].

courbe pour laquelle le segment (*cf.* figure 8) déterminé sur la tangente par le point de tangence M et l'intersection T avec une droite fixe est de longueur constante a. C. HUYGENS (1629–1695) étudie en 1692 cette courbe dont l'équation différentielle est $y^2(1 + y'^2) = a^2y'^2$. Il y montre toute son habileté (détermination par exemple de l'aire, *etc.*) et trouve une signification mécanique comme lieu des points décrits par l'extrémité d'un fil inflexible dont l'autre extrémité se déplace sur une droite. Il s'appelle "*tractoria*" ainsi que LEIBNIZ. L'aîné des BERNOULLIS, JACQUES BERNOULLI (1655–1705), parle aussi de cette courbe dès 1691, puis son frère JEAN BERNOULLI (1667–1748) établit que la développée de la tractrice est une chaînette[32]. Cette dernière courbe est représentée par un fil pesant attaché à deux extrémités comme LEIBNIZ, JEAN BERNOULLI et HUYGENS l'avaient établi juste auparavant dans les Acta Eruditorum de 1691. Il y a mélange, on le voit, de problèmes mécaniques, de connexions géométriques, et de prouesses de calcul. *Ces dernières dégagent l'aspect fonctionnel.*

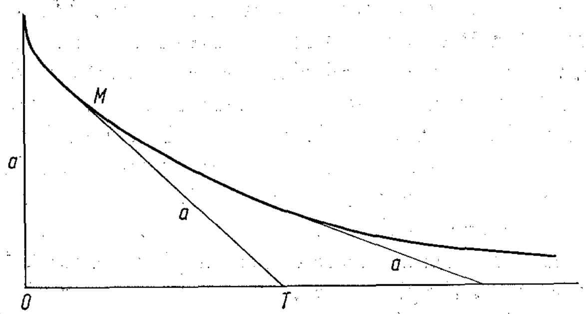

Fig. 8

Le problème des isopérimètres, d'une nature plus générale encore, popularise cet objet imprécis qu'est une fonction[33] et, ce qui est souvent essentiel en mathématique, J. BERNOULLI propose de noter "*la caractéristique d'une fonction*", reprenant un vocabulaire leibnizien et indiquant ϕx. Il s'agit avec le problème des isopérimètres de déterminer une courbe continue, de longueur donnée, délimitant entre deux ordonnées égales une aire maximale avec la droite parallèle à l'axe des abscisses et passant par la valeur commune des ordonnées. Il s'agit par exemple de réaliser un maximum de $\int_{x_1}^{x_2} \phi(x)\, dx$ sachant que $\phi(x_1) = \phi(x_2)$ et que $\int_{x_1}^{x_2} \sqrt{1 + \phi'^2(x)}\, dx$ est constante (*cf.* figure 9). L'objet de la recherche est la fonction ϕ qui intervient sous le signe intégral.

Le commentaire de l'Académie des Sciences de 1718 est instructif quant à l'évolution du concept de fonction (c'est nous qui soulignons dans le texte):

[32] Voir les articles de JEAN BERNOULLI, LEIBNIZ et HUYGENS dans les Acta Eruditorum de 1691. Citons également quant à l'exemple choisi de la tractrice: LEIBNIZ, *Opera*, 1768, t III, p. 294; HUYGENS, *Oeuvres*, tome X, p. 418; JEAN BERNOULLI, *Lectiones mathematicae*, in Opera, tome III, p. 497–499.

[33] JEAN BERNOULLI, Remarques sur ce qu'on a donné jusqu'ici de solutions des problèmes sur les isopérimètres, Mémoires de l'Académie Royale des Sciences de Paris, 1718, p. 100–139, (*cf.* aussi p. 48–54).

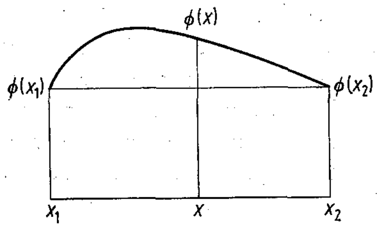

Fig. 9

"*Il faut donc concevoir une infinité d'arcs appartenant à différentes Courbes qui ont tous les deux mêmes Ordonnées extrêmes, et un axe égal compris entre ces Ordonnées, et tous Isopérimètres, ou de la même longueur; et entre cette infinité d'arcs il faut déterminer celui dont les Ordonnées sont telles qu'étant élevées à des puissances parfaites ou imparfaites, ou plus généralement à des fonctions que l'on déterminera comme on voudra, elles remplissent un espace plus grand ou plus petit que celui que rempliraient les Ordonnées de tout autre arc élevées aux mêmes fonctions*"[33]. Et c'est d'ailleurs l'image d'un fil, attaché à deux points fixes, qui est préconisée pour comprendre la nature du problème. Par ailleurs, le texte même de BERNOULLI est très ambitieux quant à la définition d'une fonction: "*On appelle fonction d'une grandeur variable une quantité composée de quelque manière que ce soit de cette grandeur variable et de constantes*". Le mot "composé" est "ambigu", et on aurait tort de trop attribuer à cette définition.

En 1748, le livre fameux d'EULER, *Introductio in analysin infinitorum*, marque la prise de conscience que l'analyse est la science générale des grandeurs et des fonctions[34]. Dans sa préface EULER indique:

"*Je me suis surtout étendu dans le premier Livre sur les fonctions de variables, parce qu'elles sont l'objet de l'Analyse infinitésimale*". Mais comment EULER lui-même applique-t-il la généralité du concept fonctionnel? Prenons un exemple au lieu de citer une définition. Au chapitre XVII de son *Introduction à l'Analyse Infinitésimale*, il s'agit de "*la manière de trouver les courbes en vertu d'autres propriétés*". Et EULER d'indiquer que la méthode fonctionnelle n'est pas restreinte à la seule représentation en coordonnées rectangulaires. Ainsi EULER veut envisager les courbes "*coupées en un seul point par les droites tirées d'un point C*" (en dehors éventuellement du point C). EULER indique comme détermination la longueur algébrique $CM = z$ sur un axe faisant un angle ϕ avec une droite fixe CA (figure 10). Et il considère z comme "*une fonction quelconque du sinus de l'angle* ϕ". Puis EULER fait semblant de croire que le problème d'unicité de l'intersection serait résolu en prenant une "*fonction uniforme, parce qu'il ne répond à l'angle*

[33] JEAN BERNOULLI, Remarques sur ce qu'on a donné jusqu'ici de solutions des problèmes sur les isopérimètres, Mémoires de l'Académie Royale des Sciences de Paris, 1718, p. 100–138 (*cf.* aussi p. 48–54).

[34] L. EULER, *Introduction à l'Analyse infinitésimale*, traduction française de J. B. LABEY, 1797, Paris (2 volumes). C'est dans cette traduction que nous prendrons les citations d'EULER, dont l'original parut en latin en 1748 (Lausanne, Bousquet; voir *Opera Omnia*, Pars I, Vol. 8–9).

$ACM = \phi$ qu'une seule valeur de la droite CM". Mais aussitôt le passage de ϕ à $(\phi + \pi)$ lui indique qu'à tout le moins z doit être une fonction *impaire* (uniforme) de $\sin\phi$. Même remarque quant à l'imparité si z est une fonction uniforme de $\cos\phi$. Sans aucune précaution oratoire, EULER conclut que: "*Toutes les courbes que les droites menées de C couperont en un seul point seront donc renfermées dans cette équation $z = P$, si P est une fonction impaire du sinus ou du cosinus de l'angle $ACM = \phi$*". Quelques lignes plus loin, on passe du "*ou*", cas d'une variable, à un "*et*", cas de deux variables pour P, puisque la fonction P est seulement assujettie à devenir négative "*lorsqu'on prend négativement le sinus et le cosinus de l'angle ϕ*". En revenant aux notations cartésiennes, cela donne z comme fonction impaire de x/z et de y/z. Puis EULER se contente alors d'une écriture en série, d'ailleurs particulièrement compliquée puisqu'il s'agit de tenir compte de tous les produits ou quotients possibles de x/z et y/z fournissant une fonction impaire:

$$z = \alpha\frac{x}{z} + \beta\frac{y}{z} + \gamma\frac{z}{x} + \delta\frac{z}{y} + \varepsilon\frac{x^3}{z^3} + \zeta\frac{x^2y}{z^3} + \eta\frac{xy^2}{z^3} + \Theta\frac{y^3}{z^3} + i\frac{x^2}{yz} + \varkappa\frac{y}{x}\frac{y}{z} + \lambda\frac{y}{x}\frac{z}{x} + etc. \tag{12}$$

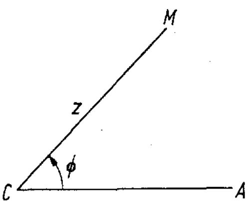

Fig. 10

Ce développement (12) *ne correspond donc pas à un développement analytique* en x/z et y/z, mais paraît plus riche *a priori*. L'important est qu'en divisant par z on n'obtient que des puissances paires de z ($z^2 = x^2 + y^2$) et, dit EULER: "*on fait disparaître toute irrationalité, et il restera une équation rationnelle entre x et y*". Toutefois une partie de la richesse du développement (12) est illusoire et le "etc." qui y figure ne semble pas signifier un développement infini. En fait le passage par ce développement ne fut effectué, semble-t-il, que pour établir que la courbe cherchée se représente comme l'égalité à une constante d'une fonction homogène en x et y de degré -1. Comme si l'écriture $z = \phi\left(\frac{x}{z}, \frac{y}{z}\right)$ où $\phi\left(-\frac{x}{z}, -\frac{y}{z}\right) = -\phi\left(\frac{x}{z}, \frac{y}{z}\right)$ impliquait $1 = \phi\left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2}\right) = \frac{1}{\Psi(x,y)}$ où $\Psi(\lambda x, \lambda y) = \lambda\Psi(x,y)$ pour tout λ! Et EULER d'écrire *a priori* la fonction homogène Ψ de degré 1 comme quotient de deux fonctions homogènes en x et y, le numérateur étant de degré n et le dénominateur de degré $(n-1)$. Mais il réduit encore le résultat en ne prenant que des polynômes et énumère enfin les premiers cas possibles quant aux courbes considérées:

$$\alpha x + \beta y = c, \qquad \text{droites}, \quad n = 1,$$

$$\alpha x^2 + \beta xy + \gamma y^2 = c(Ax + By), \qquad \text{coniques}, \quad n = 2,$$

$$\alpha x^3 + \beta x^2 y + \gamma xy^2 + \delta y^3 = c(Ax^2 + Bxy + Cy^2), \text{cubiques}, \quad n = 3, \ etc.$$

Certes la conclusion d'EULER nous apprend, *mais a posteriori*, la limitation de sa recherche. "*Nous avons donc fait une énumération complète de toutes les courbes algébriques que les lignes droites, menées par un point donné C, ne rencontrent qu'en un point*"[35]. Nous restons donc sur l'impression qu'EULER se situe à la fois en deçà et au delà de la limitation qu'il donnait, comme BERNOULLI, à la définition d'une fonction, au n° 4 du premier chapitre de *l'Introduction à l'analyse infinitésimale*: "*Une fonction de quantité variable est une expression analytique composée, de quelque manière que ce soit, de cette même quantité et de nombres, ou de quantités constantes*"[36]. (C'est nous qui soulignons le qualificatif "analytique".) Juste après cette définition, EULER distingue les fonctions algébriques et les fonctions transcendantes. Comme les moyens de calcul qualifiés de transcendants sont variés au chapitre IV de l'Introduction[34], EULER indique que la façon la plus convenable est l'obtention d'une série entière, voire d'une série infinie de puissances non nécessairement entières ("*praeter potestates ipsius z exponentes integros affirmativos habentes admitti debent potestates quaecunque*"). En somme, EULER spécifie le mode de calcul explicite d'une fonction et nous voyons avec l'écriture (12) que les puissances négatives sont acceptées.

Il peut être intéressant de préciser des restrictions pratiques chez EULER quant au concept de fonction, notamment en liaison avec des équations fonctionnelles et particulièrement les fonctions homogènes déjà envisagées ci-dessus. En 1755, dans son ouvrage *Institutiones calculi differentialis*[37], EULER entend caractériser

---

[35] Ouvrage cité en [34], 2ème volume, chapitre XVII, page 208 et suivantes (c'est nous qui soulignons).

[36] L'original latin de l'*Introduction à l'Analyse infinitésimale* d'EULER énonce

"*Functio quantitatis variabilis est expressio analytica quomodocunque composita ex illa quantitate variabili et numeris seu quantitatibus constantibus.*"

La difficulté d'interprétation de cette définition est dans la compréhension de l'adjectif "analytique". Dans l'exemple pratique choisi, on a pu voir qu'EULER ne prend pas seulement une série entière, mais admet des puissances négatives. Par ailleurs, au bout du compte il n'y a pas développement infini! Concluons plutôt à l'ambiguïté. Ambiguïté renforcée par la mention du mot uniforme par opposition au mot multiforme.

"*La fonction multiforme est celle qui, pour chaque valeur déterminée qu'on met à la place de la variable, donne plusieurs valeurs déterminées*".

(*Cf.* [34], page 6.) On a dès lors l'impression que *le mot "analytique" a l'acception de moyen explicite du calcul.* Nous allons voir que sur bien des cas concrets, telle semble être la position eulérienne. Mais celle-ci va évoluer.

[37] L. EULER, *Institutiones calculi differentialis*, Petropoli, 1755 (§ 222, Partis prioris Caput VII, voir *Opera Omnia*, Series I, vol. 10, Leipzig, 1913).

les fonctions homogènes V de degré n, disons en deux variables x et y, c'est-à-dire les fonctions pour lesquelles on a l'équation fonctionnelle

$$V(\lambda x, \lambda y) = \lambda^n V(x, y). \tag{13}$$

EULER remarque qu'en prenant $y = tx$, on transforme $V(x, y)$ en un produit de deux fonctions x^n et $T(t)\,(= V(1, t))$: "*Quare si ponentur* $y = tx$, *induet* V *huiusmodi formam* Tx^n *existente* T *functione ipsius* t". Pour EULER, T est *a priori* différentiable ("*sitque* $dT = H\,dt$") tout comme V. Nous sommes, il est vrai, dans le cadre d'un texte consacré au calcul différentiel! Calculant ensuite de deux façons la différentielle de la fonction supposée homogène V,

$$dV = P\,dx + Q\,dy = x^n\,dT + nx^{n-1}\,T\,dx,$$

EULER obtient $Px + Qy = nV$. Soit l'équation dite d'EULER,

$$\frac{\partial V}{\partial x} x + \frac{\partial V}{\partial y} y = nV. \tag{14}$$

("*quae aequatio relationem inter* P *et* Q *ita definit, ut, si altera sit cognita, altere facile inveniatur*"). Réciproquemment, en 1770, EULER[38] établit que la solution générale de (14) est $V = x^n T\left(\dfrac{y}{x}\right)$, c'est-à-dire la forme fonctionnelle déduite de (13). Il avait déjà fait antérieurement une telle remarque dans le cas où $n = 0$[39].

Les différents problèmes fonctionnels d'origine géométrique (courbes développées *etc.*), cinématique (brachystochrone), optique (caustiques) *etc.*, voire de classification mathématique comme noté chez EULER, conduisent à une multiplicité de fonctions, dont la régularité est toujours vérifiée, sauf en des points isolés et exceptionnels. Et du coup le besoin se fait sentir d'une investigation systématique des fonctions qualifiées d'élémentaires. Ce sont alors *les propriétés différentielles ou les propriétés intégrales*, par exemple $\dfrac{d(e^x)}{dx} = e^x$ ou $\displaystyle\int \frac{dx}{x} = \operatorname{Log} x$ qui apparaissent comme les plus naturelles pour fournir les propriétés de ces fonctions et non des propriétés fonctionnelles comme $e^{x+y} = e^x \cdot e^y$, lesquelles sont déduites et ne servent pas de propriété caractéristique. Il n'y a pas vraiment de démarche constructive, c'est-à-dire de présentation systématique de toutes les fonctions élémentaires à partir de quelques propriétés de base. Car est absente l'idée même d'une caractérisation pour de telles fonctions d'un usage courant. L'avantage des propriétés différentielles ou intégrales est de conduire, avec la formule de TAYLOR, aux développements en série considérés comme tellement importants. Aussi est-il naturel de voir figurer les relations fonctionnelles surtout lorsqu'il est possible d'en induire des développements en série entière. La démarche d'EULER, quant à

[38] L. EULER, *Institutiones calculi integralis*, III, Petropol, 1770, Pars I, Cap. VII, p. 92. (Voir *Opera Omnia*, Series I, Vol. 13, Berlin, 1925, p. 116–119.)

[39] L. EULER, *Institutiones calculi integralis*, I, Petropol. 1768. (*Opera Omnia*, Series I, vol. 11, p. 259–260, Leipzig, 1913).

"*Functio homogenea nullius dimensionis, quae ergo posito* $y = ux$ *abit in functionem ipsius* u".

l'exponentielle, est révélatrice dans l'ouvrage déjà cité[34] de 1748. Peut-être n'est-il pas inutile de rappeler cette démarche. Partant d'une constante $a > 1$, la définition de a^x est considérée comme allant de soi, ou plutôt comme relevant du calcul intégral. Une explication est ajoutée, laquelle indique que $a^{p/q}$ se trouve comme solution X de $X^q = a^p$, l'unique solution comprise entre $a^{[p/q]}$ et $a^{[p/q]+1}$ où $[p/q]$ désigne la partie entière de p/q, pour des entiers p et q. La propriété de monotonie de a^x semble suffire à EULER pour définir a^x pour toutes les valeurs de x. Et c'est seulement arrivé à ce point qu'EULER signale l'avantage *numérique* de la relation fonctionnelle

$$a^{x+y} = a^x a^y. \tag{15}$$

Le logarithme (de base a) est défini comme fonction réciproque de a^x et son équation fonctionnelle dûment indiquée (déduite de (15)). Une des formes de cette équation est soulignée par EULER comme *numériquement pratique*

$$\log \sqrt{xy} = \frac{\log x + \log y}{2}.$$

Mais n'oublions pas qu'EULER traite d'analyse infinitésimale, donc de développements en série. Il envisage, pour une valeur k ne dépendant que de a, l'écriture

$$a^w = 1 + kw + o(w), \tag{16}$$

lorsque w tend vers 0, n'indiquant ce que j'ai noté $o(w)$ que verbalement en précisant que k est le rapport $(a^w - 1)/w$ lorsque w est infiniment petit. Passant à la puissance d'ordre n, l'idée est de faire tendre simultanément w vers 0 et n vers l'infini de sorte que le produit nw tende vers z. Développant $(1 + kw + o(w))^n$ par la formule du binôme, dont on note l'importance pratique (*cf.* § 8, méthode fonctionnelle en mathématiques), et par un hardi passage à la limite sur chaque terme du développement du binôme, EULER obtient:

$$a^z = 1 + \frac{kz}{1} + \frac{k^2 z^2}{1 \cdot 2} + \dots \tag{17}$$

Puis par inversion, ce que nous notons

$$\log(1 + x) = \lim_{n \to \infty} \frac{n}{k} [(1 + x)^{\frac{1}{n}} - 1].$$

De nouveau, le passage à la limite dans chaque terme du développement du binôme le conduit à

$$\log(1 + x) = \frac{1}{k} \left(x - \frac{x^2}{2} + \frac{x^3}{3} + \dots \right). \tag{18}$$

Le choix de $k = 1$ détermine la base des logarithmes naturels (ou hyperboliques précise EULER pour faire référence au calcul intégral, c'est-à-dire à la propriété découverte par GRÉGOIRE DE SAINT-VINCENT). Les séries (17) et (18) permettent le calcul de l'exponentielle et du logarithme. Et l'on sent bien que c'est *seulement* une fois ce calcul rendu possible qu'EULER peut traiter l'exponentielle et le logarithme comme des fonctions. Ce qui donne, croyons-nous, une tonalité particulière à la définition initiale d'une fonction[40]. Reprenant (16), et par le même procédé introduisant n, *sans qu'il s'agisse en aucune façon d'une définition des fonctions en jeu*, EULER conclut (sans la notation de la limite mais pour un infiniment grand n):

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n,$$

et

$$\operatorname{Log}(1+x) = \lim_{n\to\infty} n\left((1+x)^{\frac{1}{n}}-1\right).$$

Pratiquement, dans ce qui précède, aucune utilisation n'est notable d'une équation fonctionnelle. Il n'en sera pas de même pour le traitement eulérien des fonctions de la trigonométrie. Celles-ci sont définies géométriquement à partir du cercle et l'on déduit les équations fonctionnelles:

$$\sin^2 z + \cos^2 z = 1, \tag{19}$$

$$\sin(y+z) = \sin y \cos z + \cos y \sin z, \tag{20}$$

et

$$\cos(y+z) = \cos y \cos z - \sin y \sin z. \tag{21}$$

EULER décompose (19) en $(\cos z + i \sin z)(\cos z - i \sin z) = 1$ et déduit de (20) et (21) la relation

$$(\cos z + i\sin z)(\cos y + i\sin y) = \cos(y+z) + i\sin(y+z).$$

Ce qui, pour un entier n, lui fournit la relation dont DE MOIVRE avait la substance dès 1730,

$$\cos nz = \frac{(\cos z + i\sin z)^n + (\cos z - i\sin z)^n}{2}.$$

Par le même procédé de passage à la limite, déjà utilisé pour déduire le développement de l'exponentielle, à partir d'une variable z qui tend vers 0 et d'une variable n qui tend vers l'infini de sorte que nz approche y, EULER déduit le développement en série

$$\cos y = 1 - \frac{y^2}{1\cdot 2} + \frac{y^4}{1\cdot 2\cdot 3\cdot 4} - \frac{y^6}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6} + \dots$$

Du coup, $\cos y$ devient une fonction *acceptable* car explicitement calculable. Il procède de même pour le sinus. La même technique, à partir de la relation (19),

[40] Et cette tonalité, déjà perçue quant à la signification du mot uniforme, est reprise par J. L. LAGRANGE dans sa *Théorie des fonctions analytiques* (Paris, 1797) quand il parle explicitement d'une fonction comme d'une expression de calcul:

"*On appelle fonction d'une ou de plusieurs quantités, toute expression de calcul dans laquelle ces quantités entrent d'une manière quelconque ...*"

étendue sans commentaire aux valeurs x éventuellement complexes, le conduit aux célèbres formules dites d'EULER:

$$\cos y = \frac{e^{iy} + e^{-iy}}{2} \quad \text{et} \quad \sin y = \frac{e^{iy} - e^{-iy}}{2i}.$$

Il faut dire que plusieurs précurseurs avaient fait le point sur les formules d'addition, de soustraction et de duplication en trigonométrie circulaire, essayant de mettre un ordre dans la logique des déductions en cascade des formules. Ces auteurs avaient ainsi acquis certaine habileté à tirer des conséquences d'équations fonctionnelles. On peut ainsi signaler THOMAS FANTET DE LAGNY (1660–1734) dont un mémoire fut commenté[41] à l'Académie des Sciences en 1703, signalant une réorganisation logique de la trigonométrie. En 1727, dans un Mémoire à l'Académie de Saint-Pétersbourg, FRÉDÉRIC-CHRISTIAN MAYER déduisait toutes les formules trigonométriques des seules équations d'addition et de soustraction. En 1748, EULER traite quasi définitivement des fonctions trigonométriques[42]. De même, c'est J. H. LAMBERT (1728–1777), dans un Mémoire paru à l'Académie de Berlin en 1770, qui clôt la question des fonctions de la trigonométrie hyperbolique[43]. La question de la caractérisation[44] des fonctions trigonométriques au moyen

[41] Hist. Acad. Sc. Paris, 1703.

[42] L. EULER, Inégalités des mouvements de Jupiter et de Saturne (*Opera Omnia*, Série II, Vol. 25, p. 45–157).

[43] J. H. LAMBERT, Observations trigonométriques, Mém. Acad. Berlin **24** (1768), p. 327–354 (*Opera Math.* 2, p. 245–269, Füssli, Zürich, 1948).

[44] JULES TANNERY, dans son *Introduction à la théorie des fonctions d'une variable* (Paris, 1886), signale explicitement la portée d'un traitement des fonctions trigonométriques par la méthode fonctionnelle (§ 96): celle d'éliminer la géométrie de l'analyse. Il est surprenant que ce point de vue n'ait pas été adopté plus tôt, disons entre EULER et CAUCHY. A tout le moins cela nous oblige à nous méfier des définitions trop générales et à surtout rechercher la pratique des mathématiciens. Mais citons TANNERY (dans l'édition de 1904, Vol. I, § 197, TANNERY supprime l'adjectif philosophique):

"*Il y a un interêt philosophique évident à introduire dans l'analyse le moins possible de données expérimentales, et il importe par conséquent de donner des fonctions sin x et cos x, une définition qui repose uniquement sur la notion de nombre et n'emprunte rien à l'idée d'espace*".

Et J. TANNERY se propose de déterminer toutes les fonctions continues ϕ et Ψ telles que:

$$\phi(a + b) = \phi(a)\,\phi(b) - \Psi(a)\,\Psi(b),$$

$$\Psi(a + b) = \Psi(a)\,\phi(b) + \Psi(b)\,\phi(a).$$

Il vérifie que $f(x) = \phi^2(x) + \Psi^2(x)$ satisfait l'équation fonctionnelle de l'exponentielle:

$$f(x + y) = f(x) f(y).$$

Soit, grâce à la continuité supposée, et comme CAUCHY l'avait établi (*cf.* § 8), $f(x) = e^{gx}$. TANNERY élimine explicitement la solution identiquement nulle. Il *pose* alors

$$\cos x = e^{-(g/2)x}\,\Phi(x) \quad \text{et} \quad \sin x = e^{-(g/2)x}\,\Psi(x),$$

de formules d'addition ne fut entreprise qu'au XIX[ème] siècle finissant, et cela en dépit de l'équation fonctionnelle du cosinus traitée dans un tout autre contexte par d'ALEMBERT (*cf.* § 7, la méthode fonctionnelle en physique mathématique), et en dépit du traitement de cette même équation par A. L. CAUCHY dans son cours d'Analyse algébrique de 1821 (*cf.* § 8, la méthode fonctionnelle en mathématiques).

Pourtant, ces études d'équations fonctionnelles dérivées les unes des autres, et non certes de résolution d'équations fonctionnelles, s'avéreront utiles lorsqu'il s'agira de traiter des fonctions elliptiques ouvrant alors un champ très riche pour les équations fonctionnelles. Il est impossible de reprendre ici le détail de cette évolution. Signalons qu'une première formule d'addition fut obtenue par EULER après FAGNANO[45] pour $\int_0^x \frac{dx}{\sqrt{P(x)}}$ lorsque $P(x) = 1 - x^4$. Cette formule fut généralisée par ABEL avec les fonctions sn u et $dn\, u$. De même que pour la trigonométrie, ce n'est que beaucoup plus tard, avec WEIERSTRASS, que l'on se servira des formules d'addition (ou de soustraction) pour caractériser *a priori* les fonctions elliptiques. WEIERSTRASS a établi en effet que toute fonction analytique f satisfaisant une relation rationnelle $F(f(x), f(y), f(x+y)) = 0$ pour tous x, y complexes était soit rationnelle, soit rationnelle en $\exp(\eta x)$, soit doublement périodique. Mais ceci sort des limites que nous nous sommes imposés dans cet article[46].

---

puis procède comme EULER, dont il s'inspire, mais en suivant la rigueur weierstrassienne. Il développe ainsi $\cos x/\cos^m \frac{x}{m}$ $\left(\text{et } \sin x/\cos^m \frac{x}{m}\right)$, fait tendre m vers l'infini et déduit le développement en série entière tant de $\cos x$ que de $\sin x$. D'où la caractérisation requise. Cependant, TANNERY a besoin au cours de sa démonstration de *l'hypothèse* supplémentaire (avec sa définition de la fonction sinus): $\lim\limits_{x\to 0} \frac{\sin x}{x} = 1$. Il pourra ensuite définir le nombre π sans intervention de la géométrie. Il faut signaler *toutefois* que LOBATCHEVSKI dans son *Algèbre ou le calcul des fonctions* (Kazan, 1834) présentait une définition analytique des fonctions trigonométriques à partir du développement en série entière de $\exp(ix)$. M. YOUSCHKEVITCH me signale que cette approche date au moins de 1825 d'après le programme du cours pour 1825–1826 de LOBATCHEVSKI.

La *définition* de l'exponentielle complexe par série entière ne deviendra familière qu'au 20[ème] siècle et ainsi la déduction par l'analyse des propriétés de $\sin x$ et $\cos x$, indépendamment de toute allusion géométrique.

[45] Les recherches de G. C. DI FAGNANO (1682–1766) réalisées entre 1714 et 1718 et republiées en 1750 (*Produzioni matematiche*, Pesaro) suscitèrent deux papiers de L. EULER: De integratione aequationis differentialis, Novi Comm. Acad. Sc. Petrop. **6**, 1756–57 (1761), p. 37–57; Observationes de comparatione arcuum curvarum ellipticarum, Novi Comm. Acad. Sc. Petrop. **6**, 1756–37 (1761), p. 58–84. Voir *Opera Omnia*, Series I, vol. **20**, p. 58–79 et p. 80–107.

[46] Des détails historiques sont fournis dans les deux références suivantes: Encyclopédie des Sciences Mathématiques pures et appliquées, Volume II_{11}, II_{13}, Paris, Leipzig, 1912.

J. DIEUDONNÉ (éd.), *Abrégé d'histoire des mathématiques* 1700–1900, Paris, Hermann, 1978 (article de C. HOUZEL: Fonctions elliptiques et intégrales abéliennes, Vol. II).

Vers l'époque où EULER écrivait son *Introduction à l'Analyse Infinitésimale*, une autre intervention des équations fonctionnelles apparaissait, qui fut appelée à bouleverser le champ de l'analyse. Dans deux articles fameux de 1747, D'ALEMBERT montrait que les sinusoïdales n'étaient pas les seules solutions du problème des cordes "*tendues mises en vibration*"[47]. Une étude de ces deux articles montre à quel point les dépendances fonctionnelles sont maîtrisées. Partant d'une analyse dynamique du mouvement d'une corde tendue, sous l'hypothèse de petites vibrations, et pour une corde "*uniformément épaisse dans toute sa longueur*", fixée à deux extrémités, D'ALEMBERT fournit des relations différentielles pour la fonction $y(t, s)$, laquelle marque la dépendance de la hauteur y de la corde à partir de l'abscisse s, laquelle varie de 0 à l, et du temps t. Posant $dy = p\,dt + q\,ds$, D'ALEMBERT prouve

$$\begin{cases} dp = \alpha\,dt + \nu\,ds, \\ dq = \nu\,dt + \alpha\,ds. \end{cases}$$

Il n'écrit pas explicitement l'équation équivalente qui est l'équation aux dérivées partielles des cordes vibrantes (devenue si familière, avec un coefficient)

$$\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial s^2},$$

car il procède directement à la résolution du système différentiel. Posant en effet $dp + dq = (\alpha + \nu)(dt + ds)$, D'ALEMBERT déduit sans commentaire que $\alpha + \nu$ est fonction de la seule somme $s + t$. De fait, quelques lignes plus haut, pour justifier l'apparition de ν tant dans l'expression de dp que dans celle de dq, D'ALEMBERT avait fait appel à un théorème d'EULER, assurant que si A $(= \partial Y/\partial t)$ et B $(= \partial Y/\partial s)$ sont les dérivées partielles d'une fonction Y de deux variables s et t, on doit avoir $\partial A/\partial s = \partial B/\partial t$. En particulier, ce résultat fournit l'égalité des dérivées partielles,

$$\frac{\partial(\alpha + \nu)}{\partial s} = \frac{\partial(\alpha + \nu)}{\partial t}.$$

Un changement de variables, avec $u = t + s$ et $v = t - s$, permet d'avoir la nullité $\dfrac{\partial(\alpha + \nu)}{\partial v} = 0$, ce qui assure que $\alpha + \nu$ est fonction seulement de $t + s$. De telles manipulations sont constamment suggérées comme évidentes dans le texte de D'ALEMBERT. En tout cas, il déduit de la même façon que $\alpha - \nu$ est seulement fonction de $t - s$. D'où l'intégration possible, avec des fonctions Φ et Δ générales, du système différentiel selon

$$\begin{aligned} dy &= \frac{\Phi(t + s) + \Delta(t - s)}{2}\,dt + \frac{\Phi(t + s) - \Delta(t - s)}{2}\,ds \\ &= \frac{\Phi(t + s)}{2}\,d(t + s) + \frac{\Delta(t - s)}{2}\,d(t - s), \end{aligned}$$

[47] J. D'ALEMBERT, Recherches sur la courbe que forme une corde tenduë mise en vibration, Hist. Acad. Berlin, 1747, tome **1**, p. 214–219 et p. 220–249; J. D'ALEMBERT, Addition au mémoire sur la courbe que forme une corde tenduë mise en vibration, Hist. Acad. Berlin, 1750, p. 355–360.

dont D'ALEMBERT tire aussitôt la forme générale du mouvement des cordes vibrantes par une semblable intégration avec des fonctions ψ et Γ,

$$y = \psi(t+s) + \Gamma(t-s).$$

Ici interviennent donc *deux* fonctions ψ et Γ. C'est cette conclusion qui est la plus marquante: "*Or il est aisé de voir que cette équation renferme une infinité de courbes*". Puis D'ALEMBERT étudie les conditions aux limites du problème. A l'origine des temps, la forme $y_0(s)$ de la corde fournit une première équation fonctionnelle $y_0(s) = \psi(s) + \Gamma(-s)$. D'ALEMBERT étudie le cas particulier d'une corde rectiligne le long des abscisses au départ, soit $y_0 \equiv 0$ et déduit donc $\psi(s) = -\Gamma(-s)$. Ce qui lui procure la forme suivante pour la fonction y des deux variables s et t,

$$y = \psi(t+s) - \psi(s-t). \tag{22}$$

Mais la corde reste fixée aux abscisses $x = 0$ et $x = l$. Donc deux autres équations surgissent de ces conditions aux limites,

$$\psi(t) = \psi(-t),$$

et

$$\psi(t+l) = \psi(l-t).$$

La fonction ψ est paire d'après la première équation. Du coup, D'ALEMBERT conclut "*ψs doit être une fonction de s dans laquelle il n'entre que des puissances paires, lorsqu'on l'aura réduite en séries*". Cette conclusion marque bien la limitation pratique apportée à une manipulation explicite des fonctions, *une réduction au développement en série.* Heureusement le mode analytique de calcul semble dissocié de la représentation géométrique. En effet, la deuxième condition aux limites et la parité de ψ impliquent que ψ est périodique, de période $2l$, ce qui invalide la possibilité d'un développement arbitraire en série entière partout calculable selon D'ALEMBERT[48]. Il introduit, selon l'usage de la géométrie, une construction auxiliaire.

Il envisage la "*courbe génératrice*", une fonction paire et de période $2l$

"*qui est composée, comme je l'ai fait voir, d'une infinité de portions semblables et égales, et dont les ordonnées répondantes aux abscisses s, sont égales à ψs*".

Naturellement, D'ALEMBERT détermine ψ par une deuxième condition initiale, à savoir la vitesse dont chaque point de la corde est animé à l'origine des temps:

$$\left(\frac{\partial y}{\partial t}\right)_{t=0} = 2\psi'(s) \quad \text{pour} \quad 0 \leqq s \leqq l.$$

---

[48] Pourtant, page 248 de l'article cité en [47], D'ALEMBERT annule la condition de périodicité pour une corde non tenue à deux extrémités et dit que dans ce cas la courbe génératrice pourra être "*géométrique*".

Mais à nouveau l'obsession du développement en série entière lui fait commettre un faux pas. Puisque ψ est en série entière de puissances paires, ψ' sera en série entière de puissances impaires et donc $\left(\frac{\partial y}{\partial t}\right)_{t=0}$ ne peut être une fonction arbitraire!

"*si la fonction de s, qui exprime cette vitesse initiale, n'était pas une fonction impaire de s, le problème serait impossible*".

La phrase suivante est surprenante car elle semble mettre en doute la possibilité, non bien sûr du mouvement sous une telle condition, mais de sa représentativité par une fonction! Du moins par une fonction d'un genre acceptable.

"*On ne pourrait pas assigner une fonction de t et de s, qui représentât en général la valeur des ordonnées de la courbe pour une abscisse s et pour un temps t quelconque*".

Grâce à sa méthode, et après B. Taylor (1685–1731), d'Alembert retrouve comme solution particulière possible et en choisissant une sinusoïde pour vitesse initiale, la forme

$$y(t, s) = A \sin \lambda s \sin \lambda t,$$

pour des constantes λ convenables. En effet, il lui suffit de prendre

$$\psi(s) = -\frac{A}{2}\cos \lambda s,$$

ce qui lui fournit $y(t, s)$ sous la forme suivante, où l'on a repris les notations de d'Alembert:

$$y = A\,\frac{\left(C^{\frac{n\sqrt{-1}}{l}(t+s)} + C^{\frac{-n\sqrt{-1}}{l}(t+s)}\right)}{-4} - A\,\frac{\left(C^{\frac{n\sqrt{-1}}{l}(t-s)} + C^{\frac{-n\sqrt{-1}}{l}(t-s)}\right)}{-4}.$$

"*Telle est l'équation de la corde vibrante, dans l'hypothèse qu'elle soit en ligne droite au commencement de son mouvement et que chacun de ses points reçoive l'impulsion convenable, pour qu'elle prenne la forme de la compagne de la cycloïde extrêmement allongée*".

D'Alembert constate que $y(t, s)$, apparaît dans ce cas comme produit de deux fonctions

$$y(t, s) = \psi(t + s) - \psi(t - s) = \Delta t \cdot \Gamma s, \tag{23}$$

avec ici des fonctions Γ et Δ particulières. Un autre cas le conduit aussi à cette forme produit. Du coup, d'Alembert change de point de vue et cherche à résoudre cette dernière équation fonctionnelle (23). Ce tournant, dont les motivations peuvent paraître le jeu un peu gratuit du mathématicien, est important. En effet, mettre la solution y sous forme de produit, c'est introduire précisément la méthode de séparation des variables si efficace pour les équations aux dérivées partielles dont traite le 18ème siècle. En outre apparaît une équation fonctionnelle typique (23). Symptomatiquement D'ALEMBERT interprète *a priori* le premier membre "*comme l'ordonnée d'une courbe que forme une corde vibrante*". Autrement dit, *la nature physique réalise toutes les fonctions envisageables par les mathématiciens*, un credo souvent adopté par les savants du 18ème siècle. D'ALEMBERT prouve par une sorte de géométrie différentielle que Γ est nécessairement une "*compagne de cycloïde fort allongée*", ce qui depuis ROBERVAL au moins, signifie une sinusoïde ($\Gamma(s) = \sin Ms$) tandis que $\Delta(t) = A \cos Mt$ ou $\Delta(t) = A \sin Mt$. Et D'ALEMBERT de conclure:

"*C'est ainsi que l'application de la Méchanique à la Géométrie aide quelquefois à découvrir des vérités purement géométriques, qu'il pourraît être assez difficile de trouver en se servant de méthodes directes*".

Cette attitude s'oppose à celle qui régira l'esprit de la méthode fonctionnelle que nous allons envisager aux sections 7 et 8. Le changement d'esprit ne provient-il pas du succès d'une approche fonctionnelle à laquelle D'ALEMBERT parvient trois ans plus tard? Il revient donc sur l'équation fonctionnelle (23) ([47], 2ème mémoire). Sans le détour par l'interprétation mécaniste, mais par double différentiation appliquée sans hésitation, D'ALEMBERT obtient (sous ses notations)

$$\frac{d\,d\Delta t}{dt^2}\,\Gamma s = \frac{d\,d\Gamma s}{ds^2}\,\Delta t.$$

L'indépendance des variables t et s, ce que représentait bien l'intuition d'une forme produit pour $y(t, s)$, fait que $\dfrac{d(d\Delta t)}{\Delta t\, dt^2}$ et $\dfrac{d\,d\Gamma s}{\Gamma s\, ds^2}$ sont des constantes. *C'est la technique de la séparation des variables clairement maîtrisée.* D'ALEMBERT déduit dès lors

$$\Delta t = Mc^{t\sqrt{A}} + gc^{-t\sqrt{A}} \qquad \text{et} \qquad \Gamma s = M'e^{s\sqrt{A}} + g'e^{-s\sqrt{A}}.$$

Et A doit être un imaginaire pur lorsqu'on reprend les conditions initiales du problème de la corde vibrante (corde fixée en 0 et en l). Ainsi D'ALEMBERT exploite jusqu'au bout sa méthode fonctionnelle.

La forme (22) donnée à la solution du problème des cordes vibrantes, en y ajoutant les deux conditions initiales, et pour ψ une fonction $2l$-prériodique et paire, ne semble exiger aucune "régularité" particulière autre pour ψ. Pourtant l'origine de cette forme (22), ce dont elle fournit la solution, est une équation de nature différentielle, celle des cordes vibrantes, donc exigeant des régularités. Nous avons vu que D'ALEMBERT, dans sa manipulation de l'équation (23), n'a aucun scrupule. Il est vrai que les restrictions ambiguës de régularité se nichent dans la notion même de fonction, avec le développement en série d'une part et une représentation graphique extensive d'autre part. La question se pose de savoir si l'on peut choisir une *courbe génératrice quelconque* ψ sur un segment de longueur l et obtenir ainsi *toutes* les solution du problème de la corde vibrante sous la forme (22)? Voilà l'origine de la controverse souvent aigre sur la *nature* des fonctions intervenant dans un tel problème, et qui va courir pendant des décennies. Nous n'insisterons pas sur cet aspect souvent théorique, et malgré son rôle essentiel, car de solides travaux l'ont bien fait connaître ([2], [31] et monographie de C. TRUESDELL citée en [51]). Cette controverse marque les profondes restrictions mentales apportées au définitions soi-disant larges de la notion de fonction, liée à celle de courbe, restrictions abandonnées progressivement et avec beaucoup de difficultés. Notons que D'ALEMBERT peut trouver par sa méthode une infinité de courbes possibles et donc établir que les équations aux dérivées partielles comme les équations fonctionnelles font intervenir une liberté de nature fonctionnelle, là où les équations différentielles ne faisaient apparaître que des constantes arbitraires, *à la difficulté près, très sensible au* 18ème *siècle, mais de liberté moins grande, des solutions singulières de ces équations différentielles.*

6. Une équation fonctionnelle à une variable chez Euler et le nouveau concept de fonction

EULER sent aussi qu'il existe bien une *différence de nature* entre la résolution d'une équation différentielle et celle d'une équation fonctionnelle générale. Pour les premières, un nombre fini de constantes détermine la solution, selon l'ordre de l'équation différentielle. Il n'en n'est pas de même pour les secondes. Cette remarque est faite dans un travail de 1761, à partir certes d'un problème d'origine géométrique[49]. Le texte est peu connu et non traduit, aussi en citerons-nous d'assez larges extraits en français. Ce travail fut lu à l'Académie de Berlin en 1761 et présenté à l'Académie de Saint-Pétersbourg l'année suivante. Le titre en est pompeux:

"*Un remarquable perfectionnement de la méthode inverse des tangentes*".

Et le résumé de l'article, en tête de celui-ci, commence par signaler que la véritable gloire de l'analyse infinitésimale ne réside pas tant dans l'établissement des propriétés différentielles que dans la résolution des problèmes inverses, c'est-à-dire "*la méthode inverse des tangentes*", que nous rangeons aujourd'hui sous la rubrique des équations différentielles, auxquelles le calcul intégral sert de vestibule.

"*On rapporte d'habitude à la méthode inverse des tangentes les problèmes où l'on cherche des courbes susceptibles, quant à leurs tangentes, d'une propriété assignée d'avance. Et, puisque grâce aux tangentes, la direction du tracé des courbes est déterminée et dépend d'une relation entre les différentielles des coordonnées, cette méthode s'ouvre très largement et embrasse tous les problèmes où une propriété*

[49] L. EULER, De insigni promotione methodi tangentium inversae, Novi Comm. Acad. Sc. Petrop. **10** (1764) 1766, p. 135–155 (*Opera Omnia*, Series I, vol. 27, p. 365–383). Nous remercions vivement Mr VIOLETTE, Professeur à l'Université de Nantes, pour son infatigable collaboration dans la traduction du texte d'EULER, dont nous donnons des extraits. Cette traduction est parue, avec un commentaire, dans J. DHOMBRES, sur un texte d'EULER relatif à une équation fonctionnelle: Archaïsme, pédagogie, et style d'écriture, Sciences et Techniques en perspective, 1986, vol. 8, p. 1–55.

assignée comprend des différentielles. Beaucoup de problèmes variés sont résolus par cette méthode et l'analyse des infinis en recueille les plus grands bénéfices. Mais les problèmes qui s'y rattachent semblent devoir être rapportés à des genres distincts selon que la propriété assignée concerne seulement les points d'une courbe pris un à un, ou des points pris deux à deux, ou à plusieurs, voire une infinité. Et jusqu'à présent, on ne trouve guère d'autres problèmes traités que ceux pour lesquels la propriété assignée se rapporte soit à des points de la courbe pris un à un, soit deux à deux, c'est-à-dire des problèmes qui devraient être rangés dans ceux du premier et du second genre"[49].

EULER avec son souci pédagogique coutumier, même dans une note savante, donne des exemples concrets des deux premiers genres. Il cite comme typique du premier genre une courbe (*cf.* figure 8) pour laquelle $MT/OT = m/n$ où m/n est mis pour une constante quelconque. D'où l'équation différentielle

$$m(x\,dy - y\,dx) = ny\,ds.$$

($ds = \sqrt{1 + y'^2}\,dx$, un cas voisin de celui de la tractrice déjà envisagée[32]). EULER cite comme exemples d'un problème du deuxième genre celui des trajectoires réciproques, puis celui de la courbe catoptrique. Dans ce dernier problème, il s'agit de déterminer une courbe, liée à un point 0 donné, telle que tout rayon lumineux issu de 0 retourne en 0 après deux réflexions successives sur la courbe. EULER avait d'ailleurs résolu ce problème antérieurement. En un paragraphe court, EULER expose le rôle symétrique des deux points corrélés, donc l'involution induite sur une courbe. Il indique alors qu'il faut s'arranger pour écrire cette involution de sorte que les variables x et $-x$ désignent les deux points corrélés, ce qui témoigne d'un souci fonctionnel précis. Puis, pour les autres genres à un nombre fini de points corrélés, EULER renvoie au problème isopérimétrique et au traitement qu'il lui donna dans un article datant de 1761. Enfin, EULER s'attaque au dernier genre, apparemment original, où une infinité de points est concernée:

"*Mais si la condition assignée concerne à la fois deux points, sous la réserve que le retour de l'un à l'autre n'ait pas lieu, mais qu'il y ait passage du second à quelque troisième point, de là à un quatrième et ainsi de suite à l'infini, alors ce ne sont pas deux points mais une infinité de points de la courbe qui doivent être considérés comme liés par la condition assignée. C'est pourquoi la méthode usuelle pour deux points seulement ne saurait réussir ici. Par conséquent, ne me souvenant même pas d'un quelconque problème de ce genre jamais proposé ou résolu, je ne pense rencontrer aucune opposition de la part de ceux qui cultivent l'Analyse en exhibant quelques problèmes de ce genre et en offrant une méthode de résolution. Toutefois tant s'en faut que d'après moi cette méthode puisse résoudre tous les problèmes de ce genre que l'on puisse imaginer et je reconnaîtrais d'autant plus volontiers mon insuffisance. Et parce que l'exploration d'un genre tout à fait nouveau de questions n'est pas avare de temps, j'espère bien que personne ne me reprochera ces premières tentatives de mon cru dans l'investigation de cette partie vierge de l'Analyse*"[49].

Le premier exemple que fournit EULER provient du problème géométrique suivant, visiblement concocté pour les seuls besoins de la classification: "*Sur* *l'axe AV construire une courbe EO telle que, quand on mêne la normale IQ à un point quelconque I l'ordonnée QK levée du point Q soit égale à cette même normale QI*" (figure 11). La liaison entre I et K entraîne une liaison simultanée en une infinité d'autres points. "*C'est en cela même que consiste la difficulté de ce problème*".

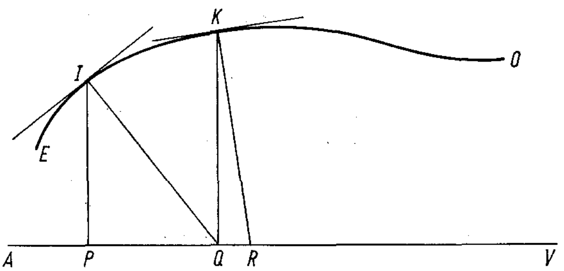

Fig. 11

En notations actuelles, l'équation fonctionnelle que l'on déduit s'écrit

$$(f(x+f(x)f'(x)))^2 = (f(x))^2\,(1+f'(x)^2)\,. \tag{24}$$

EULER est bien conscient d'une telle relation fonctionnelle dont l'existence, précise-t-il, tient au *principe de continuité* ("*per principium continuitatis talis aequatio inter x et y desideratur*"). Et ce principe indique une liaison $y=f(x)$, ou $F(x, y)=0$, c'est-à-dire une *liaison fonctionnelle.* Mais toute l'ambiguïté persiste par la qualification même du principe quant aux propriétés de régularité implicitement supposées pour f. Il est clair que dans l'exemple envisagé la fonction f est au moins supposée différentiable. Cependant, chez EULER, le mot continu associé à une courbe, le plus souvent, signifie bien autre chose. A savoir que la loi $y=f(x)$, celle qui permet le calcul effectif de y connaissant x, est de *même nature* pour toutes le valeurs de la variable x. Autrement dit, la connaissance locale de la courbe, implique celle de la "*loi*" donc entraîne la connaissance de la courbe en totalité. Lorsqu'il y a changement de loi d'un point à l'autre, on est dans le cas discontinu. C'est encore l'aspect explicite de *calcul* qui régit la manipulation d'une fonction, par l'équation en jeu; cet aspect pouvant être "*analytique*" dans l'acception large accordée à ce qualificatif par EULER. La meilleure explication du principe de continuité est fournie par L. EULER dans un texte écrit vers 1763 (et consacré à l'utilisation des fonctions discontinues en analyse). EULER précise: "*Iam vero notissimum est, in Geometria sublimiori alias lineas curvas considerari non solere, nisi quarum natura certa quadam relatione inter coordinatas per quampiam aequationem expressa definiatur, ita ut omnia eius puncta per eandem aequationem tanquam legem determinentur. Quae lex cum principium continuitatis in se complecti censeatur, quippe qua omnes curvae partes ita vinculo arctissimo inter se cohaerent, ut nulla in illis mutatio salvo continuitatis nexu locum invenire possit, hanc ob rem istae linae curvae continuae appellantur, nihilque interest, sive aequatio illarum naturam continens sit algebraica sive transcendens, sive cognita sive etiam-*

num incognita, dummodo intelligamus dari quandam aequationem, qua natura huiusmodi linearum curvarum exprimatur"[50].

"*Or il est très notable qu'en géométrie sublime il ne soit pas habituel de considérer d'autres courbes planes que celles dont la nature est d'être définie par une relation précise exprimée entre les coordonnées en sorte que tous ses points soient déterminés par la même équation, comme par une loi. Et cette loi est alors censée contenir en elle le principe de continuité: plus exactement toutes les parties de la courbe se tiennent entre elles par un lien tellement étroit que ne peut prendre place en elles aucun changement sans que ne cesse le lien de continuité. Pour cette raison, on appelle ces lignes des courbes continues. Et peu importe que l'équation contenant la nature de celles-ci soit algébrique, transcendante, connue ou même inconnue, pourvu que nous comprenions que soit donnée une certaine équation par laquelle la nature des courbes de ce genre s'exprime*".

On aura remarqué l'utilisation du mot "*loi*", là où l'on attendrait le mot "*fonction*", lequel figure par contraste dans le titre de l'article. Le mot fonction couvrant désormais un champ plus grand que celui de loi de calcul (le cas "*discontinu*" au sens d'EULER). Et cette extension de 1763 faisait suite à une évolution nette depuis la définition de 1748. En témoigne une définition très large du concept de fonction qui figure dès 1755:

"*Quae autem quantitates hoc modo ab aliis pendent, ut his mutatis etiam ipsae mutationes subeant, eae harum functiones appellari solent; quae denominatio latissime patet atque omnes modos, quibus una quantitas per alias determinari potest, in se complectitur*". "*Si des quantités dépendent d'autres de telle sorte que si ces dernières sont changées, les premières sont également changées, alors les premières sont appelées fonctions des dernières. Ce vocabulaire est de la nature la plus large et inclut toute méthode par laquelle une quantité se trouve déterminée par d'autres*"[37].

Cependant, il arriva bien souvent à EULER, après cette définition, de n'en pas faire un usage aussi large (sauf toutefois à y rester fidèle pour éliminer comme fonction une fonction constante!). Nous en avons déjà vu des exemples avec les fonctions homogènes. Nous allons voir un autre exemple avec une solution de l'équation fonctionnelle (24) et l'on pourrait multiplier les cas. Où se situe donc la ligne de démarcation entre tel et tel emploi du mot fonction? On peut certes affecter de n'y voir que des incohérences ou des oublis ou une désinvolture d'EULER! Er si la ligne de partage se situait au niveau de la mathématique en jeu, c'est-à-dire du domaine auquel l'application d'une fonction s'effectue? Des indications d'EULER lui-même conduisent à ce point de vue. EULER indique explicitement dans l'article cité en[50] que les courbes non continues (en son sens) interviennent valablement en mathématiques dans les théories jeunes, comme celle des équations aux dérivées partielles, mais non dans les théories traditionnelles comme la géométrie. Il faut rappeler que la date de l'article le situe en pleine controverse déjà évoquée

[50] L. EULER, De usu functionum discontinuarum in analysi, Novi Comm. Acad. Sc. Petrop. **11** (1765), 1767, p. 67–102 (*Opera Omnia*, Series I, vol. 23, p. 74). Commentaire et traduction à paraître dans Sciences et Techniques en perspective.

sur la nature des fonctions permises, tant comme conditions initiales que comme solutions de l'équation des cordes vibrantes, controverse impliquant D. BERNOULLI, D'ALEMBERT, J. L. LAGRANGE, L. EULER bien sûr et plus tard G. MONGE, P. S. LAPLACE *etc.* On constate en tout cas avec EULER la description d'une situation épistémologique différente du concept de fonction selon la maturité de la théorie en jeu[51].

Après cette longue parenthèse sur le concept de fonction chez EULER, venons-en à titre d'exemple au traitement de l'équation fonctionnelle (24). Remarquons qu'avec le problème considéré, nous sommes en "*géométrie sublime*" et que le principe de continuité s'applique donc aux fonctions envisagées. (Une même "nature" de la loi fonctionnelle permettant le calcul d'une ordonnée à partir d'une abscisse donnée.)

EULER *n'écrit pas l'équation fonctionnelle* (24). Il préfère avec $y = f(x)$ et $t = f(x)f'(x)$, écrire une "*équation différentielle infinie*" afin de manifester en effet le nombre "*infini*" de points reliés entre eux sur la courbe et la difficulté analytique du problème le situant bien au-delà des équations différentielles ordinaires. Naturellement, cette équation différentielle ne résout rien. EULER l'obtient en développant $f(x + t)$ par la formule de TAYLOR, de sorte que l'analyse de (24) fournit:

$$\sqrt{y^2 + t^2} = y + \frac{t}{1!}y' + \frac{t^2}{2!}y'' + \ldots + \frac{t^n}{n!}y^{(n)} + \ldots \tag{25}$$

Pourtant une remarque d'inspiration géométrique, et non fonctionnelle, sauve tout: "*Mais ce n'est pas une petite aide qu'apporte à la solution de ce problème la propriété ajoutée à la condition assignée, qui veut que toutes ces sous-normales PQ, QR etc, soient égales entre-elles*". (Pour la démonstration d'EULER, voir la note[52] où nous avons tenu à détailler cette géométrie oubliée des éléments diffé-

[51] C. TRUESDELL signale l'utilisation par EULER de fonctions valant 0 partout sauf en un point (editor's introduction, L. EULER, *Opera Omnia*, sér. II, **13**, 1956, p. 56–57). Par ailleurs, C. TRUESDELL a consacré un long traité à l'aspect mécanique du problème des cordes vibrantes (jusqu'en 1788) et a détaillé la controverse relative aux fonctions acceptables comme solutions d'équations aux dérivées partielles. Naturellement nous renvoyons le lecteur à cet ouvrage fondamental: *The rational mechanics of flexible or elastic bodies*, 1638–1788. (In L. EULER, *Opera Omnia*, sér. II, vol. II_2, Turici 1960.) Il faut aussi consulter l'article de A. P. YOUSCHKEVITCH[2] et nous n'en dirons pas plus ici sur cet usage des fonctions dans les équations aux dérivées partielles, ayant choisi de traiter *d'autres* équations fonctionnelles. Voir aussi le livre de J. LÜTZEN. *Cf.* Note [145], p. 197–198.

[52] Ce n'est pas en différentiant l'équation fonctionnelle (24) qu'EULER obtient le résultat énoncé mais au terme d'une manipulation géométrique par triangles semblables (grâce au triangle caractéristique de LEIBNIZ). La méthode fonctionnelle n'est pas allée au bout de ses potentialités! EULER bouge le point P d'une quantité infinitésimale (du 1er ordre) et porte horizontalement (*cf.* figure B) $K\nu = Qq$, le point k étant sur la courbe. Les triangles IQP et $Qq\mu$ sont semblables puisqu'ayant leurs trois angles égaux chacun à chacun. Soit

$$\frac{QI}{PQ} = \frac{Qq}{q\mu}.$$

De même les triangles QKR et νKk sont semblables. Ils ont en effet leurs trois angles rentiels, géométrie encore très imprégnée de la géométrie de l'ultime inventée par NEWTON.) Ayant remarqué que cette dernière propriété n'impliquait toutefois par la constance de la sous-normale à la courbe (cas de la parabole laquelle ne résout pas le problème posé), EULER considère seulement la fonction $t(x)$, une procédure à laquelle il est très habitué, et d'ailleurs $t(x)$ est la valeur de la sous-normale PQ,

$$t(x) = f(x) f'(x).$$

Et ainsi EULER explique par une phrase l'équation fonctionnelle satisfaite par la fonction t, une équation quelque peu surprenante,

$$t(x + t(x)) = t(x). \tag{26}$$

égaux si, au second ordre près, on place k à l'intersection de qv avec la tangente en K à la courbe, tangente orthogonale à la normale KR.

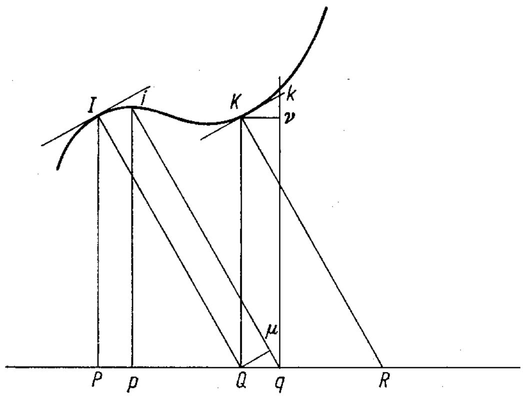

Fig. B

La similitude de ces deux triangles est une évidence pour EULER. On déduit:

$$\frac{QK}{QR} = \frac{vK}{vk}.$$

Or $vK = Qq$ par construction et $vk = q\mu$, grâce à la propriété géométrique assignée à la courbe. Donc

$$\frac{QI}{PQ} = \frac{QK}{QR}.$$

Or $QI = QK$, donc $PQ = QR$, ce qu'il fallait démontrer. EULER ajoute que sa démonstration tombe en défaut lorsque $Qq = 0$ subsiste. Ce qui revient à dire que toutes les normales concourent au même point. La courbe dans ce cas est un cercle, solution particulière du problème posé et écartée par EULER. Cas singulier qui correspond au facteur $1 + (ff')'$ (annulé) obtenu en différentiant l'équation (24). LAPLACE jouera de la même géométrie des éléments différentiels (*cf.* [61]).

Nous savons aujourd'hui que les seules solutions continues, au sens moderne, de (26), sont les fonctions constantes[53]. Naturellement EULER ne dit rien de tel, puisque le concept actuel de continuité est absent. Pour attaquer la solution de (26), EULER propose hardiment une paramétrisation à partir des fonctions cosinus et sinus. Il agit ainsi par analogie avec la résolution du problème de la courbe

[53] Le problème soulevé par EULER, par cette résolution d'équations fonctionnelles, ne semble avoir évoqué aucune recherche similaire chez les contemporains ou successeurs immédiats, pourtant avides de dépouiller l'oeuvre du maître ("*Lisez, relisez Euler*" ratiocinait LAPLACE). On trouve seulement une note parue dans les Annales de Mathématiques, un remarquable journal fondé à Nîmes par GERGONNE en 1810. (Note de TÉDENAT, Tome 12, 1821, p. 285–288 et additions pages 378–379 du même auteur.) Avec TÉDENAT, on vérifie l'évolution de la technique fonctionnelle: dérivation, changement de variable, *etc*. Toutefois TÉDENAT ne retrouve que le cercle et la droite comme courbes géométriques solutions du problème d'EULER. Mais il laisse "*à la sagacité du lecteur*" le soin de trouver toutes les solutions de (26). Pour les résultats obtenus au 20$^{\text{ème}}$ siècle, et des généralisations, voir:
M. KUCZMA, *Functional equations in a single variable*. Monografie Mat., Warszawa, 1968; J. DHOMBRES, *Some aspects of functional equations*, Lecture notes, Chula Univ., 1979; J. DHOMBRES, Itération linéaire d'ordre deux, Publ. Math., Debrecen (24), 1977, p. 277–287.

Par son intérêt pour le propos de cet article, il faudrait étudier les équations fonctionnelles à une variable développées par l'école analytique anglaise des premières décennies du 19$^{\text{ème}}$ siècle. Voir par exemple C. BABBAGE, An essay towards the calculus of functions, Phil. Trans., 1815, vol. **105**, p. 389–423. Ce dernier signale l'extension donnée au mot fonction par l'emploi de la méthode fonctionnelle.

"*The term function has long been introduced into analysis with great advantage, for the purpose of designating the result of every operation that can be performed on an equality. This extent of signification has rendered it of essential use, but the various applications of which it admits, and the question to which it gives rise, do not appear to have met with sufficient attention*".

BABBAGE indique la spécificité de la résolution des équations fonctionnelles:

"*From this and other causes, I am still inclined to think that the evolution of functional equations must be sought by methods peculiarly their own*".

GERGONNE est très intéressé et il traduit pour les Annales de Mathématiques (Tome 12, n° 3, Septembre 1821, des équations fonctionnelles), un texte de C. BABBAGE annexé au livre de son collègue J. F. W. HERSCHEL (Cambridge, 1820). Mais l'analyse des travaux de GERGONNE, de SERVOIS, et du trio "analytique", HERSCHEL, BABBAGE et PEACOCK, alourdirait trop le présent article. Il faudrait aussi citer: A. DE MORGAN, Calculus of functions, article paru dans *Encyclopaedia Metropolitana* **2**, (1835), Londres. p. 305–392. La raison principale de l'exclusion de l'étude de ces travaux du début du XIX$^{\text{ème}}$ siècle est de nature mathématique. Nous avons évité, à l'exception près de l'équation d'EULER, de mentionner des équations fonctionnelles à une seule variable (*cf.* [91] et [72]). Celles-ci, en effet, nous entraînaient vers des problèmes d'itération, bien distincts des problèmes évoqués dans le présent texte, et qui connaissent aujourd'hui un extraordinaire renouveau après les travaux de FATOU ou de DENJOY. Nous aimerions les étudier ailleurs plus amplement dans un contexte historique.

catoptrique. Là, l'existence d'une involution, traitée par les variables x et $-x$, imposait une fonction paire. Dans le cas du problème présent, la périodicité des fonctions cosinus et sinus, et de fonctions de celles-ci, doit représenter l'égalité des sous-normales pour des abscisses correspondant à des angles égaux modulo 2π. *Le principe de continuité*, qui accompagne de près la méthode fonctionnelle, se traduit ici par une périodicité de la représentation paramétrique, périodicité dont EULER suppose *a priori* qu'elle implique une fonction du sinus et du cosinus (naturellement sans qu'il soit question de développement en série de FOURIER[54]). Il y a au départ une idée d'inversion: au lieu de t fonction de x, EULER considère x comme fonction (mais multiforme) de t et exprime t à partir d'une autre variable. De fait EULER considère t comme fonction de $\sin\phi$ et de $\cos\phi$, où ϕ parcourt l'axe réel et dans ce cas x devient également une fonction du paramètre ϕ, sous la forme

$$x = \frac{\phi t}{2\pi} + T,$$

où T désigne une fonction, *a priori* quelconque, de $\cos\phi$ et de $\sin\phi$. Revenant au problème initial $(t(x) = f(x)f'(x))$, EULER déduit pour $y\ (= f(x))$ la relation

$$y^2 = \frac{1}{2\pi}\phi t^2 + \frac{1}{2\pi}\int t^2\, d\phi + 2\int t\, dT.$$

Il y a donc *a priori* supposition que T est dérivable. Naturellement des conditions supplémentaires sont nécessaires quant aux fonctions t et T (en tant que fonctions de $\cos\phi$ et de $\sin\phi$):

"*la nature tout entière du problème n'est pas épuisée puisque l'égalité des sous-normales n'implique pas nécessairement l'égalité de chaque ordonnée et de la normale précédente*".

Pour satisfaire le problème posé, un calcul montre que l'expression

$$\frac{1}{2\pi}\int t^2\, d\phi + 2\int t\, dT,$$

est "*fonction des quantités* $\sin\phi$ *et* $\cos\phi$ *et précisément une fonction uniforme pour qu'à la même abscisse ne correspondent pas plusieurs ordonnées*". Il est intéressant de constater le rôle pratique de l'adjectif *uniforme* accolé à fonction[36].

---

[54] Bien avant FOURIER (1768–1830), les mathématiciens avaient introduit des développements infinis en séries trigonométriques, ne serait-ce que pour résoudre le problème de la vibration des cordes. Voir par exemple la monographie de TRUESDELL mentionnée en [51]. C'est en 1777 qu'EULER constate les relations d'orthogonalité pour les fonctions cosinus et sinus, et en tire aussitôt profit pour obtenir le coefficient qualifié de coefficient de FOURIER, pour une fonction que l'on suppose développable en série trigonométrique. (L. EULER, Disquisitio ulterior super seriebus secundum multiplo cuiusdam anguli progredientibus (29 Mai 1777), Nova Acta Acad. Sc. Petrop. II (1793), 1798, p. 114–132 (*Opera Omnia*, Series I, vol. 16, 1, p. 333–355)). FOURIER dépasse le stade du calcul pour inaugurer une théorie de la représentation des fonctions périodiques en général.

Mais EULER, au bout du compte, ne vérifiera nullement "l'uniformité" de $y = f(x)$, et celle-ci est d'ailleurs inexacte[55]. Le principe de continuité d'EULER dépasse l'écriture uniforme $y = f(x)$ et inclut les relations implicites de la forme $F(x, y) = 0$, au fond des fonctions multiformes. EULER au terme d'une intégration par parties vérifie que l'expression

$$2V = \frac{1}{2\pi} \int t^2 \, d\phi - 2 \int T \, dt,$$

est également fonction (uniforme) de $\cos\phi$ et de $\sin\phi$. Il peut indiquer alors que les formules

$$(27) \qquad \begin{cases} x = \dfrac{t\phi}{2\pi} + \dfrac{t^2}{4\pi} \dfrac{d\phi}{dt} - \dfrac{dV}{dt}, \\[2ex] y^2 = \dfrac{t^2\phi}{2\pi} + \dfrac{t^3}{2\pi} \dfrac{d\phi}{dt} - 2t \dfrac{dV}{dt} + 2V, \end{cases}$$

"*contiennent la solution la plus générale*" du problème, c'est-à-dire de (24), du moins si t et V sont des fonctions quelconques uniformes exprimées à partir de $\sin\phi$ et de $\cos\phi$. L'introduction de deux fonctions *arbitraires* t et V pour la résolution de (24) ou (26) donne donc la mesure de la différence entre la résolution d'une équation différentielle et celle d'une équation fonctionnelle. En outre, EULER souligne qu'hormis le cercle, aucune courbe algébrique ne satisfait le problème géométrique posé. EULER donne explicitement quelques exemples de courbes solutions en choisissant des cas particuliers pour t et V. Ainsi

$$(28) \qquad \begin{cases} x = \dfrac{a}{2\pi} (\phi(1 - \cos\phi) + \frac{3}{2} \sin\phi), \\[2ex] y^2 = \dfrac{a^2}{2\pi} (1 - \cos\phi) (\phi(1 - \cos\phi) + \sin\phi), \end{cases}$$

courbe pour laquelle la sous-normale vaut $a(1 - \cos\phi)$. EULER a-t-il conscience des "discontinuitités" de la fonction y exprimée en x ? C'est en tout cas la première résolution d'une équation fonctionnelle (à une variable) conduisant à des solutions non régulières. Se donnant le prétexte de problèmes géométriques voisins, EULER va résoudre de façon analogue, c'est-à-dire par une représentation paramétrique convenable, des équations fonctionnelles comme

$$(29) \qquad f(x + f(x)) - f(x) = a(x), \; a(x + f(x)) = a(x),$$

et

$$(30) \qquad f(x + f(x))/f(x) = a(x), \; a(x + f(x)) = a(x).$$

EULER de conclure:

"... *je crois avoir découvert une méthode, à l'aide de laquelle on peut aussi résoudre*

[55] L'équation "cartésienne" d'une solution proposée par EULER pour l'équation (26) peut s'écrire $y = \sin(2\pi \, x/y)$.

des problèmes du genre de ceux dans lesquels il faut tenir compte simultanément d'une infinité de points de la courbe à étudier. La méthode en question consiste à prendre en guise de variable première un angle ϕ qui, à chaque fois qu'il augmente de quatre droits, convient successivement à chacun de ces points pris un à un. En effet puisque tous ces angles ϕ, $2\pi + \phi$, $4\pi + \phi$ ont le même sinus ($= \sin\phi$) *et le même cosinus* ($= \cos\phi$) *des fonctions uniformes quelconques de ces deux quantités conviennent également à tous ces points mis en corrélation. Or, par une composition convenable de telles fonctions, et de l'angle ϕ lui-même, on doit trouver les solutions dont ces problèmes sont susceptibles ...*"[49].

On peut conclure à la grande souplesse d'emploi de la méthode fonctionnelle chez EULER: une habileté à changer de variables, à faire intervenir des fonctions auxiliaires, dont l'origine peut être une heureuse interprétation géométrique, et surtout à utiliser le calcul différentiel sur des fonctions multiformes. Cette méthode fonctionnelle a de plus la souplesse de s'adapter à différentes "*rigueurs*": celle de la géométrie étant plus exigeante que celle de la théorie de l'élasticité ou de la vibration des cordes. Naturellement on doit constater le flou de la régularité présupposée des fonctions (multiformes) en jeu encore que l'on sente qu'EULER ait été de plus en plus sensible à la nécessité d'apporter des précisions. Et l'on pressent le danger présenté par une telle méthode entre les mains de mathématiciens moins habiles ou plus audacieux. Commençons par examiner, à partir de la remarque d'EULER sur la moindre rigueur exigée ailleurs qu'en géométrie, ce qui fut fait de la méthode fonctionnelle en physique mathématique.

7. La méthode fonctionnelle en physique mathématique

Comme nous l'avons vu avec LEIBNIZ, les BERNOULLIS, EULER, l'idée de fonction apparaît à la fin du $17^{\text{ème}}$ siècle et au $18^{\text{ème}}$ siècle, comme liée à la recherche d'une solution à un problème géométrique, à partir d'équations différentielles principalement. Assez vite toutefois, une seconde démarche vient généraliser la précédente, démarche à laquelle nous donnons le nom de *méthode fonctionnelle en physique mathematique*.

Il s'agit de quantifier, par les mathématiques, un phénomène d'origine physique, ou mécanique, ou astronomique. L'observation et "*le bon sens guidé par la raison*" indiquent un certain nombre de faits, que la tradition d'ailleurs peut déjà avoir mathématisés sous forme d'une loi, laquelle se traduit par une relation entre des variables. La méthode fonctionnelle consiste à réduire par analyse les conditions variées portant sur le phénomène, conditions de nature plus souvent rationnelle qu'expérimentale, de sorte que la forme traditionnelle de la loi soit obtenue de façon *nécessaire*, au terme de la résolution explicite d'une ou de plusieurs équations "fonctionnelles". Dans ce cadre, les équations fonctionnelles jouent le rôle clef dans "*l'axiomatisation*" mathématique des sciences physiques.

Un des exemples les plus anciens, et certainement le plus typique, est celui de la composition des forces. Cette composition revient en notre langage contemporain à l'obtention de la résultante de deux vecteurs selon la règle du parallélogramme des forces (figure 12). Comme il n'y a bien sûr pas d'identification

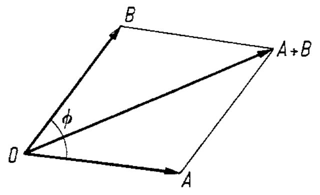

Fig. 12

entre la notion de force et le concept de vecteur, lequel ne sera élaboré que beaucoup plus tard, ce principe de construction de la résultante nécessite au moins une explication. Réfléchissant sur une tradition fort ancienne quant aux mouvements, tradition rénovée par GALILÉE, NEWTON (ainsi d'ailleurs que P. VARIGNON) place cette composition en fondement de la mécanique, mais la déduit seulement de celle expérimentalement constatée des mouvements produits[56]. En 1726, dans un mémoire incisif, DANIEL BERNOULLI réfute la nature de la preuve de NEWTON car il la considère comme non satisfaisante de par son origine expérimentale, liée à l'observation naturelle. Et il oppose classiquement les vérités contingentes ("*contingente verae*", celles provenant de l'expérience) aux vérités nécessaires ("*necessario verae*", comme celles fournies par les démonstrations géométriques). Or le but est de fonder la statique en vérité nécessaire ("... *theoremata statica non minus necessario vera esse, quam sunt geometrica*")[57]. La diatribe de D. BERNOULLI s'attaque à la composition des forces et donc à la proposition de NEWTON-(VARIGNON): "*minime tamen rigore geometrico munitum est, propositionemque non aliter quam contingenter veram reddit, uti inferius demonstrabo; nequaquam autem dubitandum, quin viri modo allegati ipsi defectum istum optime perspexerint, maluerint tamen aliqualem, eamque facilem et brevem, quam nullam vel valde prolixam magnoque propositionum geometricarum apparatu intricatam demonstrationem addere; Haec ut monerem necesse duxi, ne quis credat me nil nisi actum agere, imo per longos circuitus rem per se obviam petere*". "*Il n'est pourtant pas muni d'une rigueur géométrique, et ne rend la proposition que susceptible d'une vérité contingente comme je le montrerai ci-dessous; or il est indubitable que les maîtres dont on a parlé tout à l'heure ont très bien vu eux-mêmes ce défaut, mais ils ont cependant préféré ajouter n'importe quelle démonstration, à condition qu'elle soit facile et courte, plutôt qu'une démonstration prolixe et embarrassée d'un grand appareil géométrique. J'ai fait cela pour montrer avec nécessité, afin qu'on ne croie pas que je ne fasse rien que de redire du déjà fait, ou même de rechercher une chose évidente d'elle-même par de longs détours*". Il n'y a donc pas seulement diatribe

[56] I. NEWTON, *Principia mathematica philosophiae naturalis*, Londres, 1687 (510 p.); P. VARIGNON, *Projet d'une Nouvelle Mécanique*, Paris, 1687; P. VARIGNON, *Nouvelle Mécanique ou Statique*, Paris, 1725 (Oeuvre posthume) (Section 1ème, Lemme XVI).

Il conviendrait aussi de citer B. LAMY, Nouvelle manière de démontrer les principaux théorèmes des Eléments des Mécaniques, appendice à son *Traité de Méchanique*, Paris, 1687.

[57] D. BERNOULLI, Examen principiorum mechanicae, et demonstrationes geometricae de compositione et resolutione virium, Commentarii Acad. Petrop. 1726, tome 1, p. 126–142. C'est nous qui soulignons.

mais preuve nouvelle. BERNOULLI conclut après avoir fourni la preuve de NEWTON: "*Nihil in illa, ut falsum reiicio, sed quaedam ut obscura, quaedam ut non necessario vera*". (*Rien n'est à rejeter comme faux dans cette preuve, sauf que certains points sont obscurs et certains non vrais par nécessité.*) Et il s'attache à fournir une preuve "*géométrique*" *nécessaire*. Nous n'argumenterons pas, à la suite de E. MACH, sur l'origine physique donc expérimentale de la notion de force. La logique de la méthode fonctionnelle, sans doute des mathématiques en général, est de partir d'un donné, mais les relations de ce donné sont axiomatiquement fixées. Certes le donné initial provient fréquemment d'une inspiration physique, mais cette origine n'intervient plus pour la traitement mathématique ultérieur. La preuve de BERNOULLI est analytique en ce sens que cette preuve réduit la composition des forces au seul cas de forces égales et orthogonales. Pour ce cas, un passage à la limite, certes dans le style de la méthode d'exhaustion, est esquissé, mais non précisé. S'il n'y a pas d'équation fonctionnelle en tant que telle, de telles équations sont sous-jacentes. Par exemple l'intensité de la résultante de deux forces est une fonction homogène de degré 1 des intensités de ces forces. D'ailleurs cette homogénéité est considérée comme une évidence lorsque l'on double les intensités des deux forces et donc déduite pour les seuls nombres entiers, d'où(!) sa validité admise pour tout multiple. Plus tard, D'ALEMBERT suit la même voie analytique, et sans faire explicitement intervenir d'équations fonctionnelles, on le sent fort proche de cette démarche, la fonction étant seulement mentionnée par sa signification physique[58] mais une "*démonstration*" de continuité était établie. En effet, ayant obtenu la résultante pour deux forces d'intensités égales écartées d'un angle A, écartées d'un angle b ou écartées d'un angle $A - b$, D'ALEMBERT en déduit la résultante selon la diagonale pour des angles $A + b$ ou $2A - b$. Comme D'ALEMBERT constate astucieusement (figure 13) l'équilibre de forces d'intensités égales et écartées les unes des autres de 120°, il peut en déduire par itération le résultat selon la diagonale pour tous les angles de la forme $\frac{P}{2^n}$ 120°, où p et n sont des entiers. Utilisant explicitement la densité de $\frac{P}{2^n}$ 120° parmi tous les angles (proposition V), il a tout pour conclure grâce à un argument du pur style de la méthode d'exhaustion (proposition VI où figurent deux raisonnements par l'absurde à partir d'une inégalité dans un sens, puis dans l'autre) et non un argu-

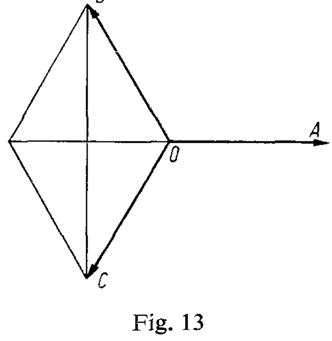

Fig. 13

ment de passage à la limite. Car D'ALEMBERT critiquait précisément un tel passage à la limite chez D. BERNOULLI:

"*D. Bernoulli démontre bien en rigueur la Proposition dont il s'agit pour tout rhombe dont l'angle sera à 90 degrés comme nombre à nombre; mais il ne le démontre pour le cas où cet angle serait incommensurable, qu'en supposant la division à l'infini*"[58].

En outre, le pas fonctionnel est franchi par D'ALEMBERT en 1769 dans un très court texte[59] dont nous allons maintenant faire l'analyse et qui contient plusieurs mémoires. Montrons d'abord l'argument fonctionnel de D'ALEMBERT dans un des mémoires à propos d'une proposition clef de la statique, celle de l'équilibre du levier. D'ALEMBERT réduit cet équilibre au problème suivant: Trouver une fonction ϕ telle que pour tous x, z on ait

$$2\phi(x) = \phi(x+z) + \phi(x-z). \tag{31}$$

C'est une équation fonctionnelle où interviennent deux variables x et z. C'est le calcul différentiel qui est aussitôt sollicité pour résoudre cette équation. D'ALEMBERT procède par différentiation de (31). D'abord en x, la première variable, ce qui lui fournit,

$$2\phi'(x) = \phi'(x+z) + \phi'(x-z).$$

Puis D'ALEMBERT différentie (31) en z, la deuxième variable, ce qui lui procure

$$0 = \phi'(x+z) - \phi'(x-z).$$

Reportant ce résultat dans l'équation précédente, et divisant par 2, il obtient,

$$\phi'(x+z) = \phi'(x). \tag{32}$$

Mais D'ALEMBERT ne conclut pas immédiatement de cette relation (32). Il applique à nouveau son procédé de double différentiation, tant en x qu'en z, de sorte que deux nouvelles équations apparaissent: c'est la méthode systématique déduite de la séparation des variables:

$$\phi''(x+z) = \phi''(x), \tag{33}$$

et

$$\phi''(x+z) = 0. \tag{34}$$

Reportant (34) dans (33), il déduit $\phi''(x) = 0$. Donc $\phi'(x) = a$ et donc $\phi(x) = ax + b$. Et D'ALEMBERT de noter que toute fonction ϕ de la forme précédente est solution de (31). L'application au levier est immédiate. Il n'y a certes aucune

[58] J. D'ALEMBERT, *Opuscules mathématiques*, Paris, Tome 1, 1761 (cinquième mémoire, p. 169–179, démonstration du principe de la composition des forces); *Opuscules mathématiques*, Paris, Tome 6, Mémoire LI, p. 360–370.

[59] J. D'ALEMBERT, Mémoires sur les principes de la mécanique, Mémoires de l'Acad. Royale des Sc., 1769, p. 278–286.

objection de faite quant à la régularité supposée de la fonction ϕ introduite pour représenter un phénomène physique. Toutefois, un souci de rigueur est latent et témoigne d'une nette évolution par rapport aux articles de 1747 car D'ALEMBERT signale que la résolution de (31), parue plus tôt à la Société Royale de Turin[86], utilisait un développement en série infinie "*dont nous n'avons pas eu besoin*" et que cette autre manière lui semble "*moins rigoureuse et moins simple*". Si la méthode fonctionnelle était dans l'air avec D. BERNOULLI et D'ALEMBERT, c'est en fait l'article de Turin par DAVIET DE FONCENEX qui fut la révélation pratique. Cet article, sans doute inspiré par LAGRANGE, avait entre autres pour but de prouver "en nécessité" la composition des forces. Mais l'article de DE FONCENEX pèche en de nombreux endroits et ce sont ces manques qui ont éveillé l'attention de D'ALEMBERT. Nous verrons plus loin au paragraphe 9 combien, au contraire, un auteur à peine plus tardif comme LEGENDRE se laissera prendre à la naïveté fonctionnelle de DE FONCENEX. Mais revenons à la composition des forces.

A ce propos, toujours dans l'article de 1769, D'ALEMBERT introduit explicitement une fonction ϕ à laquelle l'analyse du problème impose de satisfaire à l'équation fonctionnelle suivante, pour toutes valeurs des angles δ et m',

$$\phi(\delta - m') + \phi(\delta + m') = \phi(m')\,\phi(\delta). \tag{35}$$

Disons que $\phi(\delta)$ est l'intensité de la résultante de deux forces égales unitaires écartés d'un angle 2δ. Ainsi que nous l'avons déjà noté, D'ALEMBERT dans le mémoire de 1750, avait envisagé une équation fonctionnelle très voisine[47]. La méthode de double différentiation de cet article le conduit aussitôt, par séparation des variables, à la forme suivante quant à la fonction ϕ (sous ses notations)

$$\phi\alpha = c^{d\sqrt{A}} + c^{-d\sqrt{A}},$$

en ignorant la fonction partout nulle puisque d'après l'origine physique, ϕ dépend de l'angle 2α des deux forces d'intensité égale dans la composition des forces. Cela impose que le premier zéro de ϕ soit en $\alpha = 90^\circ$ et donc la conclusion,

$$\phi\alpha = 2\cos\alpha. \tag{36}$$

D'ALEMBERT en déduit alors *complètement* la loi du parallélogramme. Cette résolution lui permet de corriger une erreur de DE FONCENEX quant à l'équilibre du levier. Celui-ci, en effet, avait introduit pour déterminer cet équilibre l'équation fonctionnelle à une seule variable,

$$(\phi(x))^2 = 2 + \phi(2x). \tag{37}$$

Mais DE FONCENEX concluait à la constance de la solution de (37), ce que récuse D'ALEMBERT, fort de sa solution de l'équation (35). En effet comme $\phi(0) = 2$ dans (35), cette même équation avec $\delta = m' = x$ fournit (37). D'ALEMBERT pourtant ne s'applique pas à la résolution complète de (37), plus générale que (35). Il se contente de manifester l'"*analogie analytique singulière entre les démonstrations du parallélogramme des forces, et celle du levier*". Il reviendra d'ailleurs sur cette analogie au tome 6 de ses Opuscules Mathématiques[58]. Dans cet article, D'ALEMBERT fait intervenir systématiquement plusieurs équations fonctionnelles sur lesquelles nous ne pouvons insister faute de place. Citons un seul exemple,

$$\frac{\phi(z) + \phi(n)}{1 - \phi(z)\phi(n)} = \phi\left(\frac{z + n}{1 - nz}\right).$$

L'analogie "fonctionnelle" entre la composition des forces et l'équilibre du levier est reprise géométriquement par J. L. LAGRANGE (1736–1813) dans sa *Mécanique Analytique* de 1811:

"*Mais on peut établir une liaison immédiate entre ces deux principes par le théorème que Varignon a donné dans sa nouvelle Mécanique* (*Section 1ère, lemme XVI*), *et qui consiste en ce que si, d'un point quelconque pris dans le plan d'un parallélogramme, on abaisse des perpendiculaires sur la diagonale et sur les deux côtés qui comprennent cette diagonale, le produit de la diagonale par sa perpendiculaire est égal à la somme des produits des deux côtés par leurs perpendiculaires respectives, si le point tombe hors du parallélogramme, ou à leur différence, s'il tombe dans le parallélogramme*"[60]. Le titre même de l'ouvrage de LAGRANGE est bien révélateur du nouvel esprit fonctionnel devenu particulièrement efficace. En outre, la méthode fonctionnelle, chez LAGRANGE, est portée à une très grande fécondité grâce au principe des puissances virtuelles sur lequel tout repose. Ce dernier principe est d'abord justifié à partir du principe "*expérimental*" des poulies. Dans la preuve, il y a un mélange de calcul infinitésimal et de recours quasiment incantatoire à la méthode d'exhaustion, un recours dangereux dans le cadre de la méthode fonctionnelle. On lit en effet:

"*Le principe des puissances virtuelles étant ainsi démontré pour des puissances commensurables entre elles, le sera aussi pour des puissances quelconques incommensurables, puisqu'on sait que toute proposition qu'on démontre pour les quantités commensurables, peut se démontrer également par la réduction à l'absurde lorsque ces quantités sont incommensurables*"[60]. Ceci noté, l'essentiel est que la composition des forces, *sous forme uniquement analytique*, soit déduite par LAGRANGE du principe des puissances virtuelles. D'autres auteurs attaquèrent de même le problème de la composition des forces. Ainsi G. MONGE (1746–1818) dans son *Traité Elémentaire de Statique*, paru en 1788. Mais citons tout particulièrement l'emploi de la méthode fonctionnelle par P. S. LAPLACE (1749–1827)[61]. Ce dernier sait, après D'ALEMBERT et D. BERNOULLI, qu'il suffit de réaliser la démonstration dans le cas de forces perpendiculaires. La méthode fonctionnelle est utilisée par LAPLACE avec économie. Le calcul différentiel intervient pour

[60] J. L. LAGRANGE, *Mécanique Analytique*, Paris, 1811. Rappelons pour expliquer la citatation que LAGRANGE se pique de ne donner aucune figure dans son ouvrage, se fiant à la seule Analyse. L'influence de la méthode fonctionnelle est grande. Une première édition du texte de LAGRANGE parut en 1788 sous le titre *Méchanique Analitique*. Un travail en cours en collaboration avec P. RADELET se propose de retracer l'histoire fonctionnelle de la démonstration de la résultante des forces et de l'équation fonctionnelle (35) où seront évoqués les rôles de FOURIER, POISSON, FRANCOEUR et jusqu'à DARBOUX.

[61] P. S. LAPLACE, *Traité de Mécanique Céleste*, Paris, 1799, Volume 1, chapitre premier, p. 1–6 (*Oeuvres Complètes*, 1, Paris, 1878).

l'obtention des conditions portant sur la fonction, *ce qui évite le recours explicite à une équation fonctionnelle. "Pour cela, considérons deux forces x et y agissantes à la fois sur un point matériel M, et formant entre elles un angle droit. Soit z leur résultante, et θ l'angle qu'elle fait avec la direction de la force x; les deux forces x et y étant données, l'angle θ sera déterminé, ainsi que la résultante z, en sorte qu'il existe entre les quantités x, y et θ, une relation qu'il s'agit de connaître".*

LAPLACE cherche à améliorer l'argument d'homogénéité de DE FONCENEX (*cf.* § 9) en passant aux différentielles, mais l'abus reste le même. De ce que $n\,dx$ et $n\,dy$, pour n entier, fournissent $n\,dz$ il déduit que dx/dz est constant, c'est à dire ne dépend que de l'angle θ. Ici x, y et z sont des grandeurs scalaires. Donc

$$x = z\phi(\theta).$$

Il faut maintenant déterminer la forme de cette fonction ϕ dont on s'est assuré l'existence. Et LAPLACE joue de la flexibilité de la méthode fonctionnelle, avec plus d'acuité que ses prédécesseurs, en remarquant (figure 14),

$$y = z\phi\left(\frac{\pi}{2} - \theta\right). \tag{38}$$

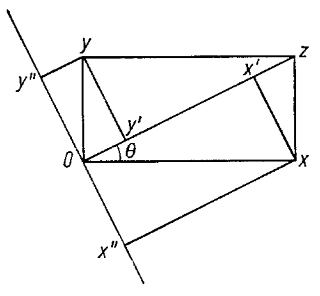

Fig. 14

Puis il considère à son tour x comme la résultante d'une force suivant z et d'une autre suivant une droite perpendiculaire à z. Il écrit astucieusement

$$x' = x\phi(\theta) = \frac{x^2}{z},$$

et

$$x'' = x\phi\left(\frac{\pi}{2} - \theta\right) = \frac{xy}{z}.$$

Une écriture analogue en y le conduit à constater la nullité de la somme $x'' + y''$, et donc le seul reste $\dfrac{x^2}{z} + \dfrac{y^2}{z}$ devient égal à z. D'où l'intensité de la résultante

$$z^2 = x^2 + y^2.$$

Scalairement, on peut mesurer z par la diagonale du rectangle construit sur y et x. Il convient maintenant de déterminer l'angle θ. Ce qui précède fournit l'équation fonctionnelle:

$$(\phi(\theta))^2 + \left(\phi\left(\frac{\pi}{2} - \theta\right)\right)^2 = 1. \tag{39}$$

Une équation que LAPLACE n'écrit pas explicitement. Puisque D'ALEMBERT a fourni une méthode de résolution des équations fonctionnelles par différentiation, autant, juge LAPLACE, travailler tout de suite avec les différentielles plutôt qu'avec l'équation fonctionnelle. Aussi garde-t-il d'abord y et change-t-il x en $x + dx$. De qui provoque "*une diminution d'une quantité infiniment petite $d\theta$*" de l'angle d'inclinaison de la résultante vis à vis de x. Il décompose dx en une composante le long de z, la résultante, et une autre orthogonale et de valeur $y/z\, dx$ grâce à (38) convenablement appliquée. Et LAPLACE explicite à nouveau (38) sous forme géométrique au niveau des éléments différentiels, négligeant en outre les infiniment petits d'ordre supérieur. Ce qui revient à écrire $\frac{y}{z}\,dx = z\phi\left(\frac{\pi}{2} - d\theta\right) = z\left(\phi\left(\frac{\pi}{2}\right) - \phi'\left(\frac{\pi}{2}\right)d\theta\right)$. Or $\phi\left(\frac{\pi}{2}\right) = 0$. D'où

$$d\theta = \frac{-y}{kz^2}\,dx, \quad \text{où } k \text{ est une constante.}$$

Par analogie, en échangeant les rôles de x et y, on obtient $d\theta = \frac{x}{kz^2}\,dy$. Sans vergogne, si simultanément x et y varient, LAPLACE additionne. D'où

$$k\,d\theta = \frac{x\,dy - y\,dx}{z^2}.$$

Ce que LAPLACE écrit sous forme d'une équation différentielle à variables séparées grâce à la valeur de $z^2 = x^2 + y^2$

$$k\,d\theta = \frac{d\left(\frac{y}{x}\right)}{1 + \frac{y^2}{x^2}}.$$

Par intégration

$$\frac{y}{x} = \operatorname{tg}(k\theta + \varrho), \quad \text{où } \varrho \text{ est une constante.}$$

LAPLACE en déduit

$$x = z\cos(k\theta + \varrho).$$

La nullité de x n'a lieu que pour $\theta = \pi/2$ (mod. π). D'où finalement

$$x = z\cos\theta.$$

Donc $\phi(\theta) = \cos\theta$. Ce qui termine la démonstration. Une démonstration relativement difficile pour l'époque et pourtant placée aux toutes premières pages de la *Mécanique Céleste*[61], publiée en 1799. Il est vrai que l'ouvrage de vulgarisation, à savoir l'*Exposition du système du Monde*, était paru auparavant en 1796.

On constate dans le traitement de LAPLACE une grande habileté dans le maniement de la méthode fonctionnelle mais une réticence à calculer par différentiation automatique d'une équation, comme eût pu le faire D'ALEMBERT et une préférence pour la géométrie vieillote des éléments différentiels. *Au fond, sa démarche ne conduit pas à un progrès.*

Dans sa *Mécanique Rationnelle*, en 1805, S. D. POISSON (1781–1840) pour la démonstration du parallélogramme des forces réintroduit l'équation fonctionnelle (35) de D'ALEMBERT en ajoutant au second membre un facteur 2 (forme désormais classique):

$$\phi(x+y) + \phi(x-y) = 2\phi(x)\phi(y). \tag{40}$$

En 1821, A. L. CAUCHY en fournit toutes les solutions continues ([75], page 106 et suivantes) sans doute inspiré par l'argument ancien de dichotomie de BERNOULLI. Mais avec cette résolution systématique d'équations fonctionnelles par CAUCHY un autre chapitre de l'analyse est ouvert, chapitre que nous allons explorer.

8. La méthode fonctionnelle en mathématiques

Puisque l'axiomatisation de la statique, par exemple, puis de la dynamique, conduisaient à des succès éclatants, couronnés par la publication de la *Méchanique analitique* de LAGRANGE (1736–1813) en 1788, la méthode fonctionnelle acquérait ainsi un prestige considérable en physique mathématique. Et cette même méthode pouvait se transposer dans l'étude axiomatique de certaines propriétés mathématiques. Telle était bien la démarche d'EULER en 1755 se donnant pour but de caractériser les fonctions homogènes. On peut ajouter que l'étude d'EULER de l'équation fonctionnelle (24) relève de la volonté d'éclairer la méthode fonctionnelle, ce qu'EULER appelle "*principe de continuité*". Un exemple beaucoup plus frappant d'utilisation de la méthode fonctionnelle en mathématiques est fourni par la démonstration de la *formule du binôme*. Plus encore que la formule de TAYLOR, cette formule du binôme est l'une des formules clefs des mathématiques au cours du 18$^{\text{ème}}$ siècle. Elle énonce,

$$(1+x)^{\alpha} = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}x^n + \dots, \tag{41}$$

et la démonstration, combinatoire pour α entier, remonte bien avant PASCAL[62]. Pour les valeurs non entières, en particulier $\alpha = \pm\frac{1}{2}$, cette formule est essentielle car elle permet l'analyse locale de toutes les courbes algébriques dont s'occupe le 17ème et le 18ème siècle. C'est le rôle que NEWTON manifeste. Ceci explique qu'on lui attribue la paternité de cette formule. Mais on ne peut pas dire que l'on dispose d'une démonstration convaincante de (41) pour toutes les valeurs, éventuellement complexes, de α, lorsqu' EULER s'y attaque en 1755 dans un livre qui fit date, *Institutiones calculi differentialis*[37]. Malheureusement pour sa démonstration, EULER utilise les dérivées alors que précisément c'est la formule (41) qui sert de base à l'obtention de la dérivée de x^α. Il y a cercle vicieux[63].

En 1774, EULER[64] revient à la charge et tente une toute autre démonstration basée sur la loi fonctionnelle (9) avec $a = 1 + x$, à savoir $(1 + x)^{n+m} = (1 + x)^n (1 + x)^m$ pour tous entiers relatifs n et m. Pour bien marquer l'importance de la relation (9), EULER note

$$[n] = 1 + nx + \frac{n(n-1)}{2!}x^2 + \dots \tag{42}$$

---

[62] Pour l'histoire de la formule du binôme chez les Grecs, les Chinois, les Arabes et les Indiens, nous renvoyons M. KLINE, *Mathematical Thought from Ancient to Modern Times*, Oxford Univ. Press, New York, 1972.

BLAISE PASCAL donne une démonstration de la formule du binôme pour des puissances entières dans son *Traité du triangle arithmétique*, paru à Paris en 1665 et dont la composition remonte vraisemblablement à 1654, Voir *les Oeuvres complètes*, la Pléiade, Paris, Ed. J. CHEVALIER, 1962. Le cas d'une puissance fractionnaire apparaît sans preuve chez J. GREGORY en 1670, puis avec des coefficients comme $\frac{1}{2}$, $\frac{3}{2}$ chez NEWTON. Par la suite, deux sortes de preuves sont fournies:

– soit l'identification terme à terme dans des séries infinies. Ainsi G. SALVEMINI dans les Philosophical Transactions de 1742–1743 part du développement suivant, où r et n sont des entiers

$$(p + q)^{r/n} = Ap^{r/n} + Bp^{(r/n)-1}q + \dots$$

Puis élève à la puissance n,

$$(p + q)^r = p^r(A + Bp^{-1}q + Cp^{-2}q^2 + \dots)^n$$

et applique la formule du binôme à puissance entière aux deux membres.

– soit la méthode des fluxions, naturellement appliquée à une série infinie avec identification des termes par la dérivation de $(1 + x)^\alpha$. Tel fut le cas de COLSON (*The method of fluxions by the inventor, Sir Isaac Newton*, 1736, London, p. 308), de MACLAURIN (*Treatise of fluxions*, 1742, London, p. 607–608, vol. 2). COLSON remarque le cercle vicieux de la démonstration puisque l'on y utilise la fluxion d'une puissance. Or l'expression de cette fluxion est précisément elle-même établie par la formule du binôme! EULER revient sur cette pétition de principe [64].

[63] Dejà souligné par COLSON (note [62]). Comme nous l'avons établi, l'importance de la formule du binôme de NEWTON est très notable dans la déduction des propriétés de l'exponentielle par EULER (*cf.* § 5).

[64] L. EULER, Demonstratio theorematis neutoniani de evolutione potestatum binomii pro casibus quibus exponentes non sunt numeri integri, Novi Comm. Acad. Sc. Petrop. 19 (1774), 1775, p. 103–111 (*Opera Omnia*, Series I, vol. 15, p. 207–216). Traduction et commentaire dans J. DHOMBRES: Les préjugés d'Euler dans l'emploi de la méthode fonctionnelle, Sciences et Techniques en perspective, vol. **10**, 1986.

Mais dans cette notation, n ne désigne pas nécessairement un entier. Il n'y a aucun souci de convergence de la série (42). Remarquons bien que pour un entier n positif ou nul, la série entière en x notée $[n]$ prend la valeur

$$[n] = (1+x)^n. \tag{43}$$

Formellement, par produit, EULER déduit une équation fonctionnelle quant à $[n]$:

$$[m]\,[n] = [m+n]. \tag{44}$$

EULER résout cette équation fonctionnelle sur le corps des rationnels. Il part de la remarque $[m]^a = [am]$ pour a entier. Dès lors $\left[\frac{m}{2}\right]^2 = [m]$. Lorsque m est entier, le second membre est $(1+x)^m$, d'où la valeur explicite de $\left[\frac{m}{2}\right]$ dans ce cas. Ce qui prouve (41) pour $\alpha = \frac{m}{2}$, m entier. De la même façon il obtient $\left[\frac{m}{a}\right] = (1+x)^{\frac{m}{a}}$ pour tous entiers m et a. Envisageant ensuite $[0] = 1$, il déduit de (44), $[-m] = (1+x)^{-m}$. D'où la relation (41) établie "*rationnellement*" pour toute valeur "*rationnelle*" de α:

"*atque adeo hoc theorema nunc quidem firmissimis rationibus est confirmatum*",

"*de sorte que ce théorème se trouve désormais confirmé par des preuves très solides*".

En 1776, EULER reprend le même problème[65] et veut cette fois obtenir une démonstration valable pour toutes les valeurs de α, rationnelles, négatives, ou transcendantes, *etc.*[66]. On peut s'étonner qu'il ne reprenne pas la méthode fonctionnelle précédente conduisant à la résolution de (44), désormais avec des valeurs m et n aussi variées que celles souhaitées pour α. Sagesse instinctive? Prémonition qu'il faudrait acquérir d'autres propriétés fonctionnelles *a priori* de $[m]$? La preuve cette fois est basée sur une restriction de (9), à savoir

$$(1+x)^{n+1} = (1+x)^n\,(1+x). \tag{45}$$

Mais on va constater combien la nouvelle preuve néglige les vraies difficultés. Ici encore n n'est pas nécessairement un nombre entier, mais bien une variable quelconque, même complexe. En notations modernes, EULER part de la série

[65] L. EULER, Nova demonstratio quod evolutio potestatum binomi newtoniana etiam pro exponentibus fractis valeat, Nova Acta Acad. Sc. Petrop. 5, (1787), 1789, p. 52–58 (*Opera Omnia*, Series I, vol. 16, p. 112–121). (Mais le papier d'EULER date de 1776.)

[66] EULER emploie une tournure dubitative pour décrire les démonstrations antérieures d'autres géométres notamment pour les valeurs non rationnelles de l'exposant.

"*Quod autem eadem expressio veritati sit consentanea, quando exponens n est vel numerus fractus vel negativus vel adeo transcendens, plures Geometrae ostendere sunt conati, quorum demonstrationes autem vel nimis sunt abstrusae vel etiam nimis longe petitae, quam ut in limine Analyseos locum invenire queant*".

entière[67]

$$(1 + x)^n = 1 + \sum_{k=1}^{\infty} f_k(n)\, x^k,$$

et déduit de (45) les équations fonctionnelles en cascade quant aux f_k,

$$f_k(n+1) - f_k(n) = f_{k-1}(n), \quad \text{pour } k \geqq 1, \text{ et avec } f_0(n) = 1. \tag{46}$$

Implicitement ou inconsciement, compte-tenu de sa conception pratique d'une fonction, *et en contradiction avec ses définitions abstraites*, EULER présuppose que f_k est un polynôme en n et même de degré k. Comme f_k s'annule pour tout entier positif n strictement inférieur à k, EULER écrit:

$$f_k(n) = n(n-1)\dots(n-k+1)\, a_k. \tag{47}$$

La résolution en cascade du système (46) d'équations fonctionnelles le conduit à la valeur explicite du coefficient $a_k = 1/k!$. Une fois encore, la méthode fonctionnelle est couronnée de succès (mais à quel prix!) et EULER conclut à sa généralité:

"*Prorsus superfluum foret hos casus ulterius prosequi, cum iam luce meridiana clarius appareat pro singulis litteris sequentibus eosdem plane valores necessario prodire debere, quos evolutio NEWTONIANA docuit, atque haec demonstratio naturae rei tam apprime accommodata videtur, ut illi etiam in primis Analyseos elementis locus denegari nequeat. Quin etiam universum ratiocinium, quo hic usi sumus, omnem vim retinet, etiamsi adeo exponens n ut imaginarius spectaretur*".[68]

"*Il serait tout à fait superflu de poursuivre plus longtemps l'étude des cas successifs, puisqu'il apparaît plus clairement qu'à la lumière de midi, que pour les lettres suivantes prises une à une viendront exactement les valeurs mêmes qu'a montrées la formule du développement de Newton. Et cette démonstration semble tellement appropriée à la nature de la question qu'on ne saurait lui refuser place même dans les premiers éléments de l'Analyse. Bien plus, le raisonnement universel ici utilisé garde toute sa force quand bien même l'on considérerait des valeurs imaginaires de l'exposant n*".

Un peu plus de vingt ans plus tard, S. F. LACROIX (1765–1843) reprend la démonstration de cette formule du binôme dans son monumental *Traité du Calcul*

---

[67] Dans le texte d'EULER, il n'y a pas utilisation de la notation indexée. La série entière $(1 + x)^n$ s'écrit en faisant usage des lettres de l'alphabet dans leur ordre. Cependant, la notation indexée moderne ne rajoute guère à l'idée fonctionnelle, explicite chez EULER dans le présent texte.

[68] La citation d'EULER tirée de[65] joue sur l'ordre des lettres quant à la formule (47). Cette démonstration d'EULER n'est certainement pas satisfaisante et il était intéressant de mettre en évidence la limitation *a priori* apportée à la variété des fonctions f possibles. Sur le plan des équations fonctionnelles toutefois, la considération du système (46) marque un progrès.

différentiel et du calcul intégral[69]. LACROIX fut un mathématicien particulièrement informé et un auteur prolifique de manuels mathématiques à la fin de la Révolution française et sous l'Empire. Et de manuels influents, réédités avec persistance pendant tout le XIX^{ème} siècle, traduits en anglais, en allemand, en russe, en espagnol, *etc.* Son souci est encyclopédique – présenter toutes les méthodes utiles de l'analyse et leurs avantages variés – plus que de rigueur linéaire, c'est-à-dire de construction pierre à pierre de l'édifice mathématique[70]. Mais voyons LACROIX à l'eouvre quant à la démonstration de la formule du binôme. Elle figure dans l'Introduction, long préalable au Traité, et dans cette Introduction les dérivées n'interviennent pas puisqu'il s'agit d'analyse algébrique, c'est-à-dire d'un traitement de séries comme des polynômes de degré infini. Dans la démonstration, LACROIX fait un usage systématique d'équations fonctionnelles au lieu d'équations différentielles. Car LACROIX insiste sur la méthode fonctionnelle, une de ces nouvelles idées "*amenées par les progrès de l'analyse*". Il définit une fonction d'une façon très générale, sans même insister sur la façon explicite dont on peut calculer celle-ci en un point donné:

"*Toute quantité dont la valeur dépend d'une ou plusieurs autres quantités est dite fonction de ces dernières, soit qu'on sache ou qu'on ignore par quelles opérations il faut passer pour remonter de celles-ci à la première*"[71].

---

[69] S. F. LACROIX, *Traité du calcul différentiel et du calcul intégral*, Paris, 1ère édition, 1797.

Un traité écrit presque intégralement avant 1795, dont une seconde édition à partir de 1810 constituera, en analyse, le texte essentiel de la 1ère moitié du XIX^{ème} siècle avec le cours de CAUCHY. C'est par LACROIX que K. MARX apprendra le calcul différentiel! Nous citerons ici la deuxième édition. Ajoutons que la Révolution Française et le Consulat virent sortir une grande quantité de manuels mathématiques. Pourtant trois auteurs au plus émergèrent du lot. E. BÉZOUT (1730–1783), S. F. LACROIX déjà mentionné et A. M. LEGENDRE dont nous parlerons plus loin. Pour une étude de l'édition de manuels mathématiques français, nous renvoyons à l'analyse quantitative suivante:
J. DHOMBRES, French Mathematical Textbooks from Bézout to Cauchy, Hist. Scient., 1985, No. 28, p. 31–137.

[70] Pour étayer cette affirmation, citons quelques lignes de la préface du traité de LACROIX (*cf.* [69]), tirées de la seconde édition.

"*On était à l'une de ces époques, lorsque je publiai (en 1797) la première édition du Traité que je mets de nouveau sous les yeux du public: la réunion des nombreux matériaux, relatifs au Calcul différentiel et au Calcul intégral, épars dans les collections académiques, pouvait seule faire connaître toutes les richesses de cette branche importante de l'Analyse, et réduire à un petit nombre de méthodes générales, une foule de procédés particuliers qui tenaient à l'enfance de ces calculs; mais une simple compilation n'aurait pas atteint ce but*".

LACROIX explique naturellement qu'il a fait des choix parmi les méthodes, et mis en évidence "*les rapports par lesquels elles se lient les unes aux autres*". L'accent sur les méthodes est typique de l'état d'esprit des mathématiciens français de la fin du 18ème siècle. Mais il n'y a pas encore choix pour un unique mode d'exposition du calcul différentiel et du calcul intégral. CAUCHY fera un tel choix.

[71] La définition de LACROIX a le grand mérite, par rapport à la définition d'EULER fournie dans *l'Introduction à l'analyse infinitésimale*, de ne pas s'appuyer sur un procédé

LACROIX part donc de $(p + x)^m$. Comme EULER, et appuyé par la théorisation de LAGRANGE sur les fonctions, LACROIX suppose que le développement en série entière en x de $(p + x)^m$ existe. Il s'agit de trouver le coefficient général. LACROIX introduit une variable supplémentaire sous la forme $(p + x + u)^m$, ce qui est désormais une idée fructueuse pour les équations fonctionnelles (*cf.*, par exemple,[72]). Il développe d'abord les puissances de $(x + u)$, puis développe chacune de ces puissances et identifie en puissances de u avec le développement en u de $(p + x + u)^m$. Avec une notation où pointe l'idée d'une notation indicielle, LACROIX établit que tous les coefficients se déduisent en cascade du premier coefficient (que nous notions $f_1(n)$ en (46)). LACROIX s'applique désormais à calculer explicitement ce premier coefficient.

"*Quel que soit le coefficient de x dans le second terme du développement de* $(1 + x)^m$, *il est nécessairement une fonction du seul exposant m, et je le désignerai en conséquence par f(m), la lettre f étant l'abréviation du mot fonction. D'après cette convention, il vient*":

(49) $$(1 + x)^m = 1 + xf(m) + \text{etc.}$$

Grâce à (9), LACROIX déduit

(50) $$f(m + z) = f(m) + f(z),$$

"*équation qui exprime la propriété caractéristique de la fonction représentée par f*".

Il établit aussitôt $f(m + z + s + t) = f(m) + f(z) + f(s) + f(t)$. Comme $f(1) = 1$, et $f(m + 1) = f(m) + 1$, il déduit $f(m) = m$ pour tout entier m. Puis $f(i/k) = i/k$ pour des entiers i et k. Enfin $f(-m) = -f(m) = -m$. Finalement, à tout le moins pour toutes les valeurs rationnelles de m, la formule du binôme (sans véritable considération de convergence encore que x soit supposé petit) est établie puisque $f(m) = m$. LACROIX promet la démonstration dans le cas général:

---

de calcul. Et nous avions noté que, pratiquement, EULER par l'utilisation du mot analytique, requérait une démarche calculatoire. Nous avons déjà signalé la définition plus large dans les *Institutiones calculi differentialis* [37] et l'ambiguïté qui s'y rattachait selon les utilisations envisagées. La définition de 1755 fut reprise, et en fait précisée sur le point qui nous occupe, dans le deuxième *Traité du Calcul Intégral* de J. A. N. DE CONDORCET (1743–1794), traité débuté vers 1778, dont l'impression commença en 1786 mais fut interrompue par la Révolution. LACROIX, dans sa préface [69] dit que "*les vingt quatre premières feuilles*" circulaient à Paris chez des mathématiciens et qu'elles "*contiennent toute l'exposition des principes du Calcul différentiel*". On peut imaginer qu'il fut influencé. CONDORCET, dans son texte, prend soin de signaler qu'une fonction est une relation qui à des variables x, y, z associe des valeurs F "*quand même je ne connaîtrais ni la manière d'exprimer F en x, y, z, ni la forme de l'équation entre F et x, y, z; je saurai que F est fonction de x, y, z*".

[72] *Cf.* en particulier le chapitre 13 de J. ACZÉL & J. DHOMBRES, *Functional equations in several variables*, à paraître, in Encycl. of Math. and its Applications, Cambridge University Press.

"*Je montrerai dans la suite, que quand l'exposant m serait irrationnel et même imaginaire, la proposition précédente serait toujours vraie*".

Mais LACROIX débute par une digression combinatoire, puis il envisage le développement en série de l'exponentielle et du logarithme. Et c'est cette fois l'équation fonctionnelle de l'exponentielle, c'est-à-dire (15)

$$a^{x+u} = a^x a^u, \tag{15}$$

qui lui fournit aussitôt le coefficient général du développement de a^x sous la forme $A_1^n/1 \cdot 2 \cdot 3 \ldots n$ et donc e^x [73]. Grâce à la formule générale du binôme, LACROIX en déduit le développement de $\operatorname{Log} x$ en série des puissances de $(x - 1)$ et, fidèle à la tradition du 18$^{\text{ème}}$ siècle ne séparant jamais calcul numérique et théorie, il envisage des accélérations de convergence de la série obtenue. LACROIX retrouve ensuite le développement de $\operatorname{Log}(1 + x)$ à partir de la *seule équation fonctionnelle du logarithme*, c'est-à-dire (10). Ceci est indispensable pour éviter un cercle vicieux car, maintenant, LACROIX peut établir la formule du binôme dans le cas général de coefficients α non nécessairement rationnels. En effet la relation $\operatorname{Log}(1 + x)^\alpha = \alpha \operatorname{Log}(1 + x)$ fournit par identification du premier coefficient l'égalité

$$f(\alpha) = \alpha,$$

pour tous α, éventuellement complexes[74]. On aura remarqué que ce résultat n'est pas obtenu par passage à la limite à partir de ce qui fut obtenu pour les nombres rationnels, *mais résulte d'une part d'une considération d'équation fonctionnelle, et de développement en série de puissances d'autre part*. D'ailleurs, c'est l'équation fonctionnelle de D'ALEMBERT-POISSON,

$$\cos(x + u) + \cos(x - u) = 2 \cos x \cos u, \tag{51}$$

qui fournit de la même façon le développement en série de la fonction cosinus, puis celui de la fonction sinus.

En un sens, la *double* démarche de LACROIX pour prouver la formule de binôme manifesterait surtout l'inadéquation de la méthode fonctionnelle. Que faut-il de plus pour passer de la détermination de $f(\alpha) = \alpha$ pour α rationnel, à la même relation pour α quelconque? LACROIX est muet sur ce point, mais ne commet

[73] C'est à cette occasion que LACROIX introduit la notation indexée

$$A_0 + A_1x + A_2x^2 + A_3x^3 + \ldots$$

"*Cette notation, qui d'abord paraît un peu compliquée, est néanmoins très commode, et très propre à faire reconnaître la loi qui règne entre les valeurs des coefficients*".

[74] Le passage au cas complexe exige la connaissance des logarithmes complexes. LACROIX ne soulève aucune difficulté, se contentant d'une identification formelle. Signalons ici une résolution de l'équation (50) chez L. B. FRANCOEUR, *Traité de mécanique élémentaire*, Paris, an IX. Cette fois l'équation provient de la démonstration du parallélogramme des forces pour déduire "analytiquement" la dépendance homogène par rapport à l'intensité de la résultante de deux forces égales. Il développe f en série entière et constate la nullité de la dérivée. Plus loin (pages 19–22), FRANCOEUR attaque l'équation fonctionnelle (40), dite de D'ALEMBERT aujourd'hui.

aucune généralisation présomptueuse: il change de méthode pour passer au cas général, d'ailleurs avec α complexe, ce qui paraît compliquer le problème. Quelle que soit l'idée que l'on ait de la liaison "continue" des points d'une fonction, chacun sent qu'un passage "à la limite" doit, sous certaines conditions, être possible! Le "principe de continuité" d'EULER[50], selon lequel il y a *même loi* (analytique) par laquelle une fonction (de la géométrie) doit être explicitement calculable, donne le résultat $f(\alpha) = \alpha$ pour toutes valeurs de α, éventuellement complexes. Pourtant ni EULER, ni LACROIX, ne proposent un tel raisonnement. A l'opposé, dans sa Préface au *Traité*[69], LACROIX n'hésite pas à considérer comme une évidence que toute fonction soit dérivable.

"*La propriété commune à toutes les fonctions d'admettre une limite, dans le rapport de leurs accroissements à ceux de la variable dont elles dépendent, limite différente pour chaque fonction, mais constamment la même pour une même fonction et toujours indépendante des valeurs absolues des accroissements, est un fait analytique bien constaté*".

Devrait-on donc la belle démonstration de LACROIX quant à la formule du binôme au seul fait qu'il la donne dans son Introduction, et donc sans faire usage des limites? Au contraire, la généralité de sa définition d'une fonction serait-elle illusoire? Nous verrons d'autres auteurs, pourtant de haut calibre, oublier tout scrupule et s'empêtrer.

En tout cas, il est difficile de penser que la démonstration de LACROIX, jointe à celle d'EULER, n'ait pas donné à réfléchir à A. L. CAUCHY (1789–1857). Et avec ce mathématicien, les équations fonctionnelles vont occuper une des premières places en Analyse dans le cadre d'une totale réorganisation de celle-ci. Longtemps préparé, le nouvel exposé est publié dans le *Cours d'Analyse Algébrique* de 1821[75]. Cette réorganisation, qu'on a souvent qualifiée à juste titre de rigorisation, part d'un petit nombre de concepts et vise à une économie de moyens en évitant tant des cercles vicieux du raisonnement que des mélanges de méthodes distinctes. Aussi les méthodes algébriques sont-elles nettement séparées des méthodes relatives aux séries entières. Le *Cours* ne concerne que ces deux méthodes. Celles-ci sont donc distinguées très explicitement des méthodes différentielles qui font l'objet d'un cours différent[76]. Les concepts de base sont ceux de fonction et de limite de quantités, concepts auxquels est intimement lié celui de continuité

---

[75] A. L. CAUCHY, *Cours d'Analyse de l'Ecole Royale Polytechnique*, Paris, 1821, Vol. I, Analyse algébrique (Oeuvres complètes, série 2, vol. 3, Paris, 1897).

[76] A. L. CAUCHY, *Résumé des leçons sur le calcul infinitésimal*, Paris, 1823. Voir aussi les *Leçons sur le calcul différentiel*, Paris, 1829. Pour la séparation des méthodes algébriques et des méthodes liées aux séries, CAUCHY est explicite dans sa préface [75]:

"*Quant aux méthodes, j'ai cherché à leur donner toute la rigueur qu'on exige en géométrie, de manière à ne jamais recourir aux raisons tirées de la généralité de l'algèbre. Les raisons de cette espèce, quoique assez communément admises, surtout dans le passage des séries convergentes aux séries divergentes, et des quantités réelles aux expressions imaginaires, ne peuvent être considérées, ce me semble, que comme des inductions propres à faire pres-*

d'une fonction. Avec ces limites, un style mathématique est mis au point, par manipulation d'encadrements et d'inégalités, par jeu de stabilité des limites sous l'effet d'une fonction continue et donc par une technique d'approximation portée à un niveau d'investigation théorique. Ainsi nous voyons à l'oeuvre la définition de la continuité chez CAUCHY pour résoudre l'équation

$$\phi(x+y)=\phi(x)+\phi(y). \tag{52}$$

Cette résolution intervient assez tôt, au chapitre V, et elle ne constitue pas un exercise sans importance. En fait deux motifs principaux doivent être signalés. D'une part, cette résolution est la preuve pratique du *savoir-faire* de CAUCHY par utilisation rigoureuse de sa définition de la continuité. CAUCHY se sait novateur par rapport à un EULER, quand bien même il use, mais dans un autre sens, du même vocabulaire de continuité. Cette exhibition d'efficacité le pose en maître. D'autre part cette démonstration prépare, sans intervention du calcul différentiel, une démonstration de la formule du binôme de NEWTON, formule qui est *la clef* de l'ouvrage selon nous. Quand on sait le soin mis par CAUCHY à l'ordonnance logique de ses cours écrits, en quoi consiste peut être sa plus exigeante forme de rigueur mathématique, laquelle tranche sur ses contemporains et rejoint la rigueur euclidienne, on voit bien l'importance qu'il accorde aux équations fonctionnelles dans sa mise en place de l'Analyse. C'est donc à très bon droit qu'on qualifie l'équation (52) d'équation de CAUCHY. L'évolution des résultats de CAUCHY sur (52) au cours du livre est révélatrice.

D'abord l'énoncé du problème, très moderne:

"*Déterminer la fonction $\phi(x)$ de manière qu'elle reste continue entre deux limites réelles quelconques de la variable x, et que l'on ait pour toutes les valeurs réelles des variables x et y,*

$$\phi(x+y)=\phi(x)+\phi(y)\text{''}. \tag{52}$$

Comme LACROIX, CAUCHY obtient $\phi(\varrho)=\frac{m}{n}\phi(\alpha)$ où $\varrho=\frac{m}{n}\alpha$, avec m et n entiers positifs.

"*puis, en supposant que la fraction $\frac{m}{n}$ varie de manière à converger vers un nombre quelconque μ, et passant aux limites, on trouvera*

$$\phi(\mu\alpha)=\mu\phi(\alpha)\text{''}.$$

---

sentir quelque fois la vérité, mais qui s'accordent peu avec l'exactitude si vantée des sciences mathématiques".

S'il est question dans [75] de quantités infinitésimales (en fait de limites), la dérivation n'intervient pas. (Ce qui nous vaut de belles démonstrations des inégalités de convexité, note II de [75].) CAUCHY ne précise pas ce que l'on entend par variable ou quantité réelle. Il se fonde sur le livre V d'EUCLIDE (quantités mesurables) et la bijection opérée clairement par DESCARTES entre opérations géométriques et opérations algébriques. Sur ce point, voir les ouvrages signalés en [5].

Valable pour $\mu > 0$, CAUCHY étend le résultat à $\mu \leqq 0$, puis fait $\alpha = 1$. Il vérifie que, réciproquement, $\phi(x) = ax$, a quelconque, satisfait (52) et est continue[77]. Aussitôt après, CAUCHY résout, toujours pour des fonctions continues, l'équation fonctionnelle

$$(53) \qquad \phi(x + y) = \phi(x)\phi(y).$$

Par le même artifice après avoir constaté la positivité de ϕ puisque $\phi(2x) = (\phi(x))^2$, il déduit $\phi(x) = (\phi(1))^x$. Notons dans le cours de CAUCHY que l'exponentielle est définie, mais en Note seulement[78], par passage à la limite à partir de A^n, $A^{m/n}$, donc A^x. Le corps du texte parle de la continuité de A^x sans vraiment la prouver. Mais ultérieurement les *Leçons sur le calcul différentiel* de 1829 fourniront une preuve. L'attitude générale, sinon l'ordre, est cohérente. Il faut bien mesurer déjà le caractère novateur de CAUCHY *définissant* A^x. Au chapitre suivant, consacré aux séries divergentes et convergentes, le problème suivant est posé:

"*Développer, lorsque que cela se peut, la fonction* $(1 + x)^\mu$ *en série convergente entière suivant les puissances ascendantes et entières de* x".

[77] Continuité: il s'agit de la stabilité $\lim\limits_{x \to x_\alpha} f(x_\alpha) = f(x)$ et ici de la densité des rationnels dans les réels. On a pourtant beaucoup glosé sur la définition d'une fonction continue dans le Cours d'Analyse Algébrique (Chap. II, § 2).

"*Cela posé, la fonction* $f(x)$ *sera, entre les deux limites assignées à la variable* x, *fonction continue de cette variable, si, pour chaque valeur de* x *intermédiaire entre ces limites, la valeur numérique de la différence*

$$f(x + \alpha) = f(x)$$

décroît indéfiniment avec celle de x. *En d'autres termes, la fonction* $f(x)$ *restera continue par rapport à* x *entre les limites données, si, entre cas limites, un accroissement infiniment petit de la variable produit toujours un accroissement infiniment petit de la fonction elle-même*".

Certes, on doit noter que dans cette remarquable définition il n'est pourtant pas distingué entre la continuité en un point et la continuité sur tout un intervalle. Cette ambiguïté se marque sur la façon dont $f(x + \alpha) - f(x)$ dépend de x et toutes les erreurs de CAUCHY sur l'uniformité sont alors prévisibles. Mais une glose encore plus lourde fut entretenue parce que CAUCHY utilisa l'expression "infiniment petit" et on a voulu y voir un jeu sur les quantités infinitésimales, jeu post-leibnizien interprétable par l'analyse non-standard. J'ai beaucoup de mal à adhérer à une telle explication en examinant des démonstrations de CAUCHY comme celle que nous venons de voir sur l'équation fonctionnelle (52). D'autant que la définition du début du chapitre au § 1 est suffisamment claire:

"Soit α une quantité infiniment petite, c'est-à-dire une variable dont la valeur numérique décroisse indéfiniment".

Pour l'équation fonctionnelle (53), CAUCHY élimine sans le dire la fonction identiquement nulle comme solution possible, car il ne la considère pas comme une fonction, puisqu'elle ne "dépend" pas de la variable. EULER agissait de même (*cf.* [49]).

[78] Note I du Cours cité en [75], page 341–342 dans la reproduction des *Oeuvres Complètes*. Remarquons encore dans cette définition de l'exponentielle le passage à la limite, par continuité, des rationnels aux réels.

CAUCHY sait qu'il ne peut suivre la voie de LACROIX. Comment établir en effet la continuité de $f(\mu)$, valeur du coefficient en x, sans précisément faire intervenir la dérivée de $(1+x)^{\mu}$, ce que CAUCHY veut absolument éviter? Il reprend EULER et part alors directement de la série entière à laquelle on veut aboutir, ce qui est une technique analytique typique. CAUCHY pose:

$$\phi(\mu) = 1 + \mu x + \frac{\mu(\mu-1)}{2} x^2 + \dots \tag{54}$$

Il en assure la convergence pour x compris entre -1 et $+1$ grâce à une étude du rayon de convergence qu'il vient de mettre au point (à partir du critère dit de D'ALEMBERT u_n/u_{n+1} déduit du critère de CAUCHY en $\sqrt[n]{u_n}$). Puis, pour ces valeurs de x, CAUCHY assure la continuité de ϕ comme fonction de μ, pour toutes les valeurs de ϕ. Légitime assertion? CAUCHY se base sur un théorème prétendant établir que la continuité subsiste par passage à la limite[79]. Puis CAUCHY examine le produit $\phi(u)\phi(\mu')$. Il utilise pour cela le théorème du produit de deux séries entières, dit théorème de CAUCHY[80]. Il faut donc calculer l'expression

$$\frac{\mu(\mu-1)\dots(\mu-n+1)}{n(n-1)\dots 2\cdot 1} + \frac{\mu(\mu-1)\dots(\mu-n+2)}{(n-1)(n-2)\dots 2\cdot 1}\cdot\frac{\mu'}{1} + \dots + \frac{\mu'(\mu'-1)\dots(\mu'-n+1)}{n(n-1)\dots 2\cdot 1}.$$

CAUCHY dans un chapitre précédent (chap. IV, § III), et d'ailleurs par une méthode fonctionnelle, a montré l'égalité de cette expression à

$$\frac{(\mu+\mu')(\mu+\mu'-1)\dots(\mu+\mu'-n+1)}{n(n-1)\dots 2\cdot 1}.$$

---

[79] Il s'agit du théorème I du chapitre VI, § VI, du cours cité en [75].

"*Lorsque les différents termes de la série*

$$u_0, u_1, \dots, u_n, u_{n+1}, \dots \tag{1}$$

sont des fonctions d'une même variable x, continues par rapport à cette variable dans le voisinage d'une valeur particulière pour laquelle la série est convergente, la somme s de la série est aussi, dans le voisinage de cette valeur particulière, fonction continue de x".

On peut comprendre l'erreur de CAUCHY et chercher une explication du côté d'une uniformité cachée dans la définition de la continuité. Disons plus simplement qu'il y a ambiguïté de la définition et ignorance de l'uniformité dans la convergence. Mais on comprend surtout que CAUCHY, ayant placé la continuité et la convergence comme concepts fondamentaux de sa construction de l'Analyse, ait absolument besoin du théorème liant si simplement les deux concepts. Malgré un contre-exemple d'ABEL en 1826, les travaux de SEIDEL, de STOKES *etc.*, CAUCHY ne modifiera son énoncé qu'en 1853 en introduisant la convergence uniforme. Voir par exemple sur cette question: P. DUGAC, *R. Dedekind et les fondements de l'analyse*, VRIN, Paris, 1975.

[80] Théorème IV du chapitre VI, § IV du cours cité en [75]. Il faut noter *l'enchaînement* des théorèmes utilisés et l'ordre d'exposition. Forte ordonnance!

Donc il dispose de l'équation fonctionnelle qu'EULER avait résolue sur le corps des rationnels[64], et qu'on appelle aujourd'hui équation de CAUCHY exponentielle,

$$\phi(\mu + \mu') = \phi(\mu)\,\phi(\mu'). \tag{55}$$

La résolution de (55), déjà obtenue par CAUCHY pour une fonction continue ϕ, donne

$$\phi(\mu) = (\phi(1))^{\mu},$$

et donc fournit le résultat attendu, grâce à la valeur de $\phi(1)$ donnée par (54). CAUCHY obtient

$$\phi(\mu) = (1 + x)^{\mu};$$

ce qui établit la formule du binôme de NEWTON. Du moins pour toutes les valeurs réelles de μ. Ce résultat est aussitôt appliqué pour obtenir e comme $\lim\limits_{\alpha\to 0}(1 + \alpha)^{\frac{1}{\alpha}}$ et retrouver e^x par un passage à la limite. Il obtient l'équation fonctionnelle de l'exponentielle et il en déduit le logarithme par passage à la limite de $\dfrac{(1 + x)^{\mu} - 1}{\mu}$, selon la démarche eulérienne déjà expliquée. Mais CAUCHY ne s'arrête pas en un tel chemin et veut maintenant obtenir la formule du binôme pour les *valeurs imaginaires* de x, puisque sa véracité, depuis EULER au moins, était proclamée pour de telles valeurs. Pour pouvoir appliquer la même méthode, il faut désormais résoudre (55) pour une fonction prenant des valeurs imaginaires. Au chapitre VIII de son cours, CAUCHY résout aisément (52) pour une fonction à valeurs complexes. Mais une difficulté technique se présente pour le cas multiplicatif (55) puisque la méthode de résolution réelle consiste à prendre une racine n-ième, laquelle est multiforme dans le plan complexe. CAUCHY part donc de l'équation fonctionnelle

$$\overline{\omega}(x + y) = \overline{\omega}(x)\,\overline{\omega}(y), \tag{56}$$

x et y étant des variables réelles, mais où $\overline{\omega}(x) = \phi(x) + \sqrt{-1}\,\chi(x)$. Ignorant comme d'habitude la fonction identiquement nulle, CAUCHY déduit $\overline{\omega}(0) = 1$, donc, par continuité, la positivité de ϕ autour de l'origine, disons pour tout $|x| \leqq \alpha$. Pour cette valeur α, il pose avec $\varrho > 0$, $\overline{\omega}(\alpha) = \varrho(\cos\zeta + \sqrt{-1}\sin\zeta)$. Pour m entier, $\overline{\omega}(m\alpha) = \varrho^m(\cos m\zeta + \sqrt{-1}\sin m\zeta)$ grâce à la formule de DE MOIVRE. Grâce maintenant à la positivité de la partie réelle de $\overline{\omega}\left(\frac{\alpha}{2^n}\right)$, CAUCHY déduit *sans ambiguïté*

$$\overline{\omega}\left(\frac{\alpha}{2^n}\right) = \varrho^{\frac{1}{2}n}\left[\cos\left(\frac{1}{2^n}\zeta\right) + \sqrt{-1}\sin\left(\frac{1}{2^n}\zeta\right)\right],$$

et donc pour tous entiers m et n,

$$\overline{\omega}\left(\frac{m}{2^n}\alpha\right) = \varrho^{\frac{m}{2^n}}\left[\cos\left(\frac{m}{2^n}\zeta\right) + \sqrt{-1}\sin\left(\frac{m}{2^n}\zeta\right)\right].$$

CAUCHY peut désormais appliquer sa technique systématique de passage à la limite, base de la continuité et obtenir pour toute valeur réelle x, grâce à la continuité supposée de $\overline{\omega}$,

$$\overline{\omega}(\alpha x) = \varrho^x \left[\cos(\zeta x) + \sqrt{-1}\sin(\zeta x)\right].$$

Donc

$$\overline{\omega}(x) = A^x (\cos bx + \sqrt{-1}\sin bx),$$

où A est une constante positive et b une constante réelle. L'application de ceci à la formule du binôme survient au chapitre IX ([81]). Par le même procédé que dans le cas réel, c'est-à-dire par résolution de l'équation fonctionnelle (55) pour une fonction à valeurs complexes. CAUCHY obtient l'égalité pour z tel que $-1 < z < +1$, de

$$\sum_{n=0}^{\infty} \frac{\mu(\mu-1)\dots(\mu-n+1)}{n!} z^n (\cos n\Theta + \sqrt{-1}\sin n\Theta), \tag{57}$$

une série convergente, avec l'expression bien définie

$$r^{\mu}(\cos \mu t + \sqrt{-1}\sin \mu t), \tag{58}$$

où r et t dépendent de z et Θ. Avec $\mu = 1$, on dispose de la relation liant Θ, z, r et t,

$$1 + z(\cos\Theta + \sqrt{-1}\sin\Theta) = r(\cos t + \sqrt{-1}\sin t).$$

Faute d'une définition univoque de l'exponentielle complexe, il faut encore à CAUCHY deux pages de calcul et de raisonnement pour conclure à la validité générale de la formule du binôme. La difficulté réside en ceci. Pour a rationnel, CAUCHY a pu vérifier que toutes les valeurs complexes assignables à $(\alpha + \beta\sqrt{-1})^a$, ensemble noté $((\alpha + \beta\sqrt{-1}))^a$, où $\alpha = \varrho\cos\zeta$ et $\beta = \varrho\sin\zeta$, sont de la forme $((1))^a$ multiplié par l'expression $\varrho^a(\cos a\zeta + \sqrt{-1}\sin a\zeta)$ où $((1))^a$ désigne toutes les valeurs $\cos 2ka\pi + \sqrt{-1}\sin 2ka\pi$ lorsque k parcourt les entiers relatifs. Il y *a* un ensemble fini de valeurs possibles. Choisissant la détermination 1, c'est-à-dire $k = 0$, CAUCHY définit explicitement

$$[\varrho(\cos\zeta + \sqrt{-1}\sin\zeta)]^a,$$

par

$$\varrho^a(\cos a\zeta + \sqrt{-1}\sin a\zeta)$$

sous la condition de la positivité de $\cos\zeta$ [82]. Puis il prolonge cette définition *par*

[81] Problème I du chapitre IX, page 243 de la reproduction dans les *Oeuvres Complètes*.

[82] On a déjà signalé que CAUCHY avait relégué en note la définition de A^x, $A > 0$ et x réel. Mais son propos restait cohérent par passage à la limite à partir de $A^{p/q}$. Lorsque A est complexe, CAUCHY commence par définir A^x pour les valeurs rationnelles de x, à partir de la loi opératoire $(A^{p/q})^q = A^p$. Mais la fonction apparaît multiforme, d'où le choix d'une détermination complexe. Il y a nettement plus de précaution chez CAUCHY que chez LACROIX. Toutefois, dans cette démonstration de la formule du binôme, la puissance μ utilisée reste réelle.

continuité au cas où a est irrationnel, *conservant la condition* $\cos\zeta \geqq 0$. Pour les valeurs négatives du cosinus, "*on ne voit plus, même en supposant fractionnaire la valeur numérique de a, quelle est celle de l'expression de* $((\alpha + \beta\sqrt{-1}))^a$ *que l'on pourrait distinguer des autres et désigner par la notation* $(\alpha + \beta\sqrt{-1})^a$." L'expression $[1 + (z\cos\Theta + z\sqrt{-1}\sin\Theta)]^\mu$ a donc un sens *précis* pour CAUCHY lorsque $-1 < z < 1$ puisque $1 + z\cos\Theta > 0$. C'est sur elle que porte la formule du binôme. Il convient donc de préciser dans (58) les valeurs réelles de $r (r > 0)$ et de t. On a $1 + z\cos\Theta = r\cos t$, $z\sin\Theta = r\sin t$ et $\cos t$ reste donc positif. Il est bien clair que $r^2 = 1 + 2z\cos\theta + z^2$, ce qui fournit r univoquement $(r > 0)$. Il reste t. Posant $s = \operatorname{Arctg}\dfrac{z\sin\Theta}{1 + z\cos\Theta}$, on déduit $t = s + 2k\pi$ où k est un entier relatif, lequel dépend *a priori* de Θ et de z. CAUCHY utilise la continuité en z pour déduire, d'ailleurs par contradiction, la continuité de t en z, celle de s étant acquise par composition de fonctions continues. Donc on obtient la continuité de l'entier k et "*il est clair que, pour satisfaire à la condition énoncée, la quantité s devra varier toute seule, et la quantité k demeurera constante*". Premier exemple, sans doute, d'utilisation d'un argument de connexité pour une fonction explicitement continue. En faisant $z = 0$, on vérifie en fait que $k = 0$. D'où la formule finale, sous les conditions édictées par CAUCHY pour l'exponentielle complexe,

$$1 + \frac{\mu}{1} z(\cos\Theta + \sqrt{-1}\sin\Theta) + \frac{\mu(\mu-1)}{1\cdot 2} z^2(\cos 2\Theta + \sqrt{-1}\sin 2\Theta) + \dots$$

$$= (1 + 2z\cos\Theta + z^2)^{\frac{\mu}{2}}(\cos\mu s + \sqrt{-1}\sin\mu s)$$

$$= [1 + z(\cos\Theta + \sqrt{-1}\sin\Theta)]^\mu,$$

où $-1 < z < +1$ et μ est un nombre réel quelconque. C'est la formule du binôme de CAUCHY.

Nous avons cru utile de citer longuement CAUCHY car ainsi on conçoit mieux ses soucis de démonstration, les présupposés de ses notations et l'association évidente des équations fonctionnelles à une volonté de rigueur, notamment en ce qui concerne les expressions imaginaires. Quelques années plus tard, ABEL (1802–1829) reprend magistralement le point de vue de CAUCHY, tente d'en corriger l'insuffisance qu'il manifeste Dans la preuve dans la continuité de la fonction $\phi(\mu)$ et pousse encore plus loin le souci de rigueur en appliquant systématiquement la méthode fonctionnelle[83]. Ce qui lui fait multiplier le nombre d'équations fonctionnelles traitées. En effet, ABEL résout (55) dans le cas $\phi: \mathbb{C} \to \mathbb{C}$ en posant

$$\varphi(k + k'\sqrt{-1}) = f(k, k')\,[\cos\psi(k, k') + \sqrt{-1}\sin\chi(k, k')].$$

---

[83] N. ABEL, Recherches sur la série $1 + \frac{m}{1}x + \frac{m(m-1)}{1\cdot 2}x^2 + \frac{m(m-1)(m\cdot 2)}{1\cdot 2\cdot 3}x^3 + \dots$ J. für die reine und angew. Math., Bd. I, Berlin, 1826 (en allemand, mais en français dans les *Oeuvres Complètes*, p. 219–250).

Il obtient deux équations fonctionnelles dont il fournit les solutions continues. Son analyse examine ensuite les conditions de convergence de (57) pour $z = \pm 1$. *C'est la première preuve sérieuse de la validité de la formule du binôme dans le cas des exposants complexes.*

Nous venons donc de donner quelques brillantes réussites de la méthode fonctionnelle, tant en mathématiques pures que dans certains domaines non mathématiques. Et nous avons noté l'intervention grandissante des équations fonctionnelles. Avec ABEL et CAUCHY, un tournant de rigueur est certainement pris quant à la manipulation des fonctions, et les définitions utilisées sont opératoires.

Toutefois l'enthousiasme naïf pour cette méthode pouvait conduire à de sérieux déboires et un exemple suffira pour les illustrer, tout en ayant l'avantage de manifester la distance qui sépare CAUCHY et ABEL de leurs prédécesseurs immédiats quant à la maîtrise des outils de l'analyse.

9. Les naïvetés fonctionnelles

La méthode fonctionnelle est avidement utilisée par A. M. LEGENDRE (1752–1833) dans ses *Elémens de Géométrie*[84], livre qui connaîtra un franc succès. Il faut rappeler que le but de LEGENDRE, en écrivant de tels Eléments, était de renouer avec la tradition euclidienne de l'enseignement de la géométrie, en particulier, avec un exposé axiomatique (et systématique). Certes, LEGENDRE réhabilitait aussi la recherche de preuves pour ce qui, traditionnellement, passait pour un axiome. Il s'agisait seulement d'améliorer EUCLIDE! Une telle démarche avait été condamnée au 17ème siècle par le cartésianisme et l'Ecole de Port-Royal, sous prétexte de l'immédiateté intuitive des axiomes[85]. Il est donc naturel que la recherche de LEGENDRE se centre sur une preuve du postulat d'EUCLIDE quant aux parallèles. Preuve que LEGENDRE fournit sous diverses formes et qu'il engrange dans des notes aux éditions successives de sa géométrie, en se rétractant des précédentes. Son effort dévie aussi vers des démonstrations de propriétés géométriques évitant, autant que faire se peut, le recours au postulat d'EUCLIDE. Ainsi son essai de preuve directe que la somme des trois angles d'un triangle quelconque est égale à deux angles droits. Une proposition clef en géométrie euclidienne dont LEGENDRE prouve d'ailleurs l'*équivalence* au postulat d'EUCLIDE. Et LEGENDRE découvre avec ravissement la méthode fonctionnelle pour les mathématiques, mentionnant un travail antérieur sur le parallélogramme des forces[86] sur lequel nous reviendrons.

"*Nous observons, au reste, que la considération des fonctions, qui fournit ainsi une*

[84] A. M. LEGENDRE, *Elémens de géométrie*, 1ère édition, F. Didot, Paris, 1794, 335 pages. Une deuxième édition succède en l'an VIII avec 471 pages. Plusieurs notes sont ajoutées (ou modifiées) au fil des éditions successives.

[85] Voir J. DHOMBRES, Esquisse historique du rôle du raisonnement par l'absurde dans la pratique des mathématiciens, Sciences et Techniques en perspective, à paraître (Univ. de Nantes).

[86] LEGENDRE cite les Mémoires de la Société Royale de Turin, tome II, 1760 (c'est-à-dire Miscellanea Taurinensia). L'article en question est du à DAVIET DE FONCENEX. Sur les principes fondamentaux de la mécanique, p. 299–325.

démonstration très simple des propositions fondamentales de la Géométrie, a déjà été employée avec succès pour la démonstration des principes fondamentaux de la Mécanique"[87].

Bien que LEGENDRE contre une fois de plus l'influence de Port-Royal en réhabilitant l'emploi du raisonnement par l'absurde, il apprécie l'aspect direct des démonstrations issues de la méthode fonctionnelle[88]. Examinons sa démarche pour les angles d'un triangle. Soit donc un triangle ABC et p la longueur du côté AB.

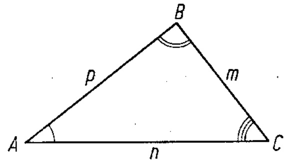

Fig. 15

LEGENDRE note A et B les deux angles adjacents et C le troisième (Fig. 15). L'angle C "*doit être une fonction déterminée des trois quantités A, B, p; ce que j'exprime ainsi,* $C = \phi : (A, B, p)$", puisque deux triangles sont égaux lorsqu'ils ont un côté égal adjacent à deux angles égaux chacun à chacun. LEGENDRE argumente pour éliminer p de la fonction ϕ:

"*Soit l'angle droit égal à l'unité, alors les angles A, B, C, seront des nombres compris entre 0 et 2; et puisque $C = \phi : (A, B, p)$, je dis que la ligne p ne doit point entrer dans la fonction ϕ. En effet, on a vu que C doit être entièrement déterminé par les seules données A, B, p, sans autre angle ni ligne quelconque, mais la ligne p est hétérogène avec les nombres A, B, C; et si on avait une équation quelconque entre A, B, C, p, on en pourrait tirer la valeur de p en A, B, C; d'où il résulterait que p est égal à un nombre, ce qui est absurde: donc p ne peut entrer dans la fonction ϕ, et on a simplement* $C = \phi : (A, B)$".

De ce résultat, il est facile de déduire que deux angles aigus à côtés deux à deux perpendiculaires sont égaux. Par décomposition, la somme des angles d'un triangle, puis celle d'un triangle quelconque, est établie égale à deux droits. Dans cette démonstration de LEGENDRE, l'hypothèse implicite d'existence de deux triangles semblables mais inégaux, est omise. La naïveté, ici, tient à une imprécision. Pourtant le mot "*hétérogène*" peut faire penser à une naïveté d'origine fonctionnelle. Et, effectivement, un deuxième exemple vient établir cette naïveté liée à l'homogénéité d'une fonction.

---

[87] Livre cité en [84], 12ème édition, 1823, p. 287.

[88] Voir la note [85]. Le rôle du raisonnement par l'absurde n'est pas neutre dans cette recherche de démonstrations du postulat d'EUCLIDE et dans la naissance des géométries non euclidiennes. On connaît la démarche de G. SACCHERI (1667–1733) partant d'une assertion contraire au postulat et tentant d'aboutir directement à une contradiction (*Euclides ab omni naevo vindicatus*, 1733). Cette démarche directe influencera K. SCHWEIKART et GAUSS.

Il s'agit pour LEGENDRE de prouver que deux triangles ayant tous les angles correspondants égaux sont semblables. Avec les notations précédentes, outre m pour la longueur du côté opposé à l'angle A, et n pour celle du côté opposé à l'angle B, LEGENDRE pose

$$\frac{m}{p} = \chi : (A, B, p).$$

"*Mais m/p est un nombre, ainsi que A et B; donc la fonction* ψ *ne doit point contenir la ligne p*". Le point soulevé est ici que le membre de gauche satisfaisant $\lambda m/\lambda p = m/p$, pour tout réel λ, on en déduit

$$\psi : (A, B, \lambda p) = \psi : (A, B, p). \tag{59}$$

Malheureusement sur la base d'EUCLIDE ou de LEGENDRE, mais sans l'axiome des parallèles, il n'a y pas suffisamment de λ distincts pour lesquels (59) reste vraie. Et quand bien même (59) resterait vraie pour tous les rationnels λ, passer à la limite présuppose une régularité non prouvée de ψ. Certes si l'indépendance de la fonction ψ en p est acquise, on est conduit aussitôt à la proportionnalité des côtés opposés aux angles égaux,[89], c'est-à-dire à la similitude de deux triangles ayant leurs angles correspondants égaux.

L'erreur d'homogénéité était déjà commise amplement par DAVIET DE FONCENEX dans le mémoire[86] auquel LEGENDRE fait référence. Il s'agissait alors de trouver la résultante de deux forces d'intensités égales a, faisant entre elles un angle ϕ. La valeur de la résultante z est fonction de a et de ϕ. Mais FONCENEX raisonne "par homogénéité" et écrit z/a comme fonction du *seul* angle ϕ. FONCENEX poursuit

"*Il suit de là que, l'angle* ϕ *demeurant constant, z est toujours proportionnel à a: on pourrait de même démontrer par cette méthode, d'une manière directe et fort naturelle, plusieurs théorèmes sur la proportionnalité des côtés des figures, et un grand nombre d'autres propositions de la Géométrie et de la Mécanique*".

Effectivement, cette erreur quant à l'homogénéité semble réduire à des truismes les propositions fortes de la géométrie euclidienne. C'est l'intuition physique de dimension. Elle réduit la Proposition 2 du livre XII d'EUCLIDE énonçant que "*les surfaces des cercles sont comme les carrés des rayons*" à une banalité ... mais c'est une intuition! Signalons enfin un dernier exemple où l'erreur d'homogénéité est explicite chez LEGENDRE.

---

[89] Il est intéressant de remarquer que LEGENDRE utilise encore le langage des proportions et note:

$$m : m' :: n : n' :: p : p'$$

pour $m/m' = n/n' = p/p'$ en désignant par des primes les éléments d'un triangle dont les angles sont égaux deux à deux à celui du premier triangle. La proportion étant déduite de $m/p = \Psi : (A, B)$ et $n/p = \Psi : (B, A)$. A l'actif de LEGENDRE, il faut indiquer qu'il prend soin de vérifier que sa méthode ne convient plus pour les triangles sphériques, et il fait remarquer qu'il faut alors faire intervenir le rayon r de la sphère, ce qui bloque la déduction par homogénéité.

Il entend prouver que l'aire d'un rectangle, dont les longueurs des côtés sont p et q est proportionnelle au produit pq. LEGENDRE use de la magie fonctionnelle en notant $\phi:(p, q)$ l'aire en question. L'additivité supposée des aires (*cf.* Fig. 16) lui indique l'équation fonctionnelle

$$\phi(p+p', q) = \phi(p, q) + \phi(p', q).$$

Soit pour un nombre entier k quelconque

$$\phi(kp, q) = k\phi(p, q).$$

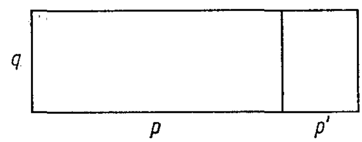

Fig. 16

L'homogénéité pour les entiers suffit selon LEGENDRE:

"*Il résulte de là que $\phi(p, q)/p$ est une telle fonction de p, qu'elle ne change pas en mettant à la place de p un multiple quelconque kp. Donc cette fonction est indépendante de p et ne doit renfermer que q*".

Ceci acquis, $\phi(p, q)/pq$ ne dépend ni de p, ni de q et l'on a une constante! Ici la naïveté tranche sur la prudence d'un LACROIX et sur la méthode rigoureuse de CAUCHY. Rappelons que l'erreur figure encore dans une édition de l'ouvrage de LEGENDRE datée de 1823, bien après la publication de LACROIX, édition postérieure au cours d'analyse de CAUCHY[90].

10. Les équations fonctionnelles de Cauchy: évolution jusqu'à nos jours

Faute de place, nous ne pouvons plus suivre linéairement l'évolution des idées quant aux fonctions et leurs incidences (ou vice versa) sur les équations fonctionnelles. A l'orée du vingtième siècle, l'analyse fonctionnelle se constitue en discipline autonome et cet essor fut préparé par des travaux se succédant à une cadence rapide sur les équations différentielles, intégrales, aux différences, aux dérivées partielles et sur les équations fonctionnelles, à une ou à plusieurs variables. Et pour le XIX$^{\text{ème}}$ siècle, même en se limitant aux seules équations fonctionnelles au sens restreint, il faudrait citer et commenter N. H. ABEL (1802–1829), N. I. LOBATCHEVSKI (1793–1856), J. BOLYAI (1802–1860), C. F. GAUSS (1774–1855), H. GRASSMANN (1809–1877), A. CAYLEY (1821–1899), B. RIEMANN (1827–1866), K. WEIERSTRASS (1825–1897), H. HANKEL (1814–1899), *etc.*[91].

[90] Et LEGENDRE balaie les objections à sa méthode fonctionnelle faites par LESLIE, les estimant réfutées.

[91] Voir par exemple la bibliographie de l'ouvrage cité en [72] ou celle de J. ACZÉL, *Lectures on functional equations and their applications*, Academic Press, 1966. Dans les

En guise de conclusion, nous nous contenterons de suivre historiquement les travaux sur l'équation de CAUCHY (52), en particulier dans son rôle dès l'introduction de la nouvelle théorie de l'intégration par H. LEBESGUE et nous citerons, chemin faisant, quelques perspectives récentes à propos de cette équation de CAUCHY.

Si de nombreux mathématiciens utilisèrent cette équation fonctionnelle et ses solutions continues découvertes par CAUCHY, il faut cependant attendre 1875 pour que G. DARBOUX (1842–1917) établisse que toute solution de (52) qui est non négative sur les réels positifs est nécessairement partout continue[92]. Ce travail de DARBOUX s'inscrit dans la lignée de l'utilisation de la méthode fonctionnelle en physique mathématique puisque l'article en question consiste à proposer une démonstration originale de la composition des forces en utilisant habilement l'équation de CAUCHY. DARBOUX résume la philosophie de cette méthode fonctionnelle:

"*Je me suis proposé de reprendre l'étude de cette question en la traitant comme un problème de pure Géométrie, ce qu'elle est au fond, et en tâchant de conduire la démonstration de manière à bien mettre en évidence quelles sont les hypothèses qu'on peut rejeter et celles qu'il est nécessaire d'admettre*"[92].

Et DARBOUX de conclure à l'intérêt pédagogique d'une telle démarche. Plusieurs auteurs vont, comme DARBOUX, chercher des conditions sur une solution de (52) pour que celle-ci soit continue, donc de la forme $f(x) = ax$ $(a \in \mathbb{R})$ d'après le résultat de CAUCHY. Ainsi CH. DE LA VALLÉE POUSSIN remarque qu'une solution de (52) "*bornée dans un intervalle, si petit qu'il soit*" est nécessairement continue[93]. La question se pose de déterminer si de telles conditions sont superflues, autrement dit si toutes les solutions de (52) sont *a priori* continues. Une démarche intéressante est de prendre le problème à l'envers et de noter les conséquences de l'hypothèse de solutions non continues de (52). Car on peut en déduire, par contradiction, des critères de continuité. C'est ce que fait WILSON en 1899. Il montre que pour une solution non continue de l'équation exponentielle de CAUCHY, à savoir

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

les valeurs atteintes par f, au voisinage de toute valeur fixée de la variable, permettent d'approcher toute valeur positive donnée d'aussi près qu'on le souhaite. Il prend toutefois soin d'indiquer que l'existence de solutions discontinues n'a

---

paragraphes précédents nous n'avons pas cherché à donner un sens restreint à la notion d'équation fonctionnelle. Cependant la plupart des exemples choisis se rapportaient à la définition restreinte établie par J. ACZÉL dans l'ouvrage mentionné ci-dessus. Nous avons très peu donné d'exemple à propos des équations fonctionnelles à une seule variable, surtout développées, il est vrai, après GERGONNE et ABEL au 19ème siècle. Désormais, pour la fin de cet article, nous nous restreignons aux équations fonctionnelles à *plusieurs variables* au sens d'ACZÉL.

[92] G. DARBOUX, Sur la composition des forces en statique, Bull. Sc. Math. [1], 1875, **9**, p. 281–288.

[93] CH. J. DE LA VALLÉE POUSSIN, *Cours d'analyse infinitésimale*, Tome 1, Paris, 1903, p. 30.

pas encore été prouvée et ne donne aucune indication supplémentaire[94]. Une démarche directe de construction de solutions discontinues de l'équation (52) fut tentée par R. VOLPI en 1897. Ce dernier remarque[95] que pour des nombres rationnels $\alpha_1, \ldots, \alpha_n$, si l'on écrit

$$x = \sum_{i=1}^{n} \alpha_i x_i, \tag{60}$$

alors une solution de (52) satisfait

$$f(x) = \sum_{i=1}^{n} \alpha_i f(x_i). \tag{61}$$

Réciproquement, soient des nombres $x_1, \ldots, x_n$ donnés pour lesquels toute relation $\sum_{i=1}^{n} \alpha_i x_i = 0$ avec des α_i rationnels, implique que $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Fixons des valeurs $z_1, \ldots, z_n$. On pose pour tout x de la forme (60),

$$f(x) = \sum_{i=1}^{n} \alpha_i z_i. \tag{62}$$

La fonction f ainsi définie par (62) satisfait (52) pour tous x et y de la forme (60). Comme le choix des $z_i, \ldots, z_n$ est arbitraire et que $f(x_i) = z_i$, on constate bien qu'en général f n'est pas de la forme $f(x) = ax$. En particulier, f serait discontinue du moins si une écriture telle que (60) était possible pour tout x réel. La construction de VOLPI s'avère insuffisante sur ce dernier point d'écriture de tout nombre réel. Quelques années plus tard, et apparemment motivé par l'addition des vecteurs, c'est-à-dire à nouveau la composition des forces, HAMEL (1877–1954) résout le problème[96]. Il construit une famille (x_i) d'éléments de $\mathbb{R}$, telle que pour tout entier $n \geqq 1$, $\sum_{i=1}^{n} \alpha_i x_i = 0$, avec des α_i rationnels, implique $\alpha_1 = \ldots = \alpha_n = 0$ et telle que (60) ait lieu pour tout x réel avec des α_i rationnels convenables. C'est ce que l'on appelle depuis une base de HAMEL, en ce sens qu'il ne s'agit que d'une base de l'espace vectoriel $\mathbb{R}$ considéré sur le corps des nombres rationnels. La force de HAMEL est d'avoir immédiatement compris le parti que l'on pouvait tirer, pour cette construction, de la découverte de ZERMELO, parue en 1904, c'est-à-dire de l'existence d'un bon ordre sur tout ensemble[97]. C'est d'ailleurs une très belle illustration du résultat de ZERMELO. HAMEL montre donc que *toutes* les solutions de (52) sont obtenues par un choix (quelconque) de valeurs sur la base en application de (62). Et HAMEL déduit en particulier que le graphe de toute solution discontinue de l'équation de CAUCHY est dense dans le plan. Ce qui lui procure un critère simple entraînant la continuité d'une solution de l'équation (52).

---

[94] E. B. WILSON, Note on the functions satisfying the functional relation $\phi(u)\,\phi(v) = \phi(u+v)$, Annals of Math., (2), 1899, 1, p. 47–48.

[95] R. VOLPI, Sulle funzione a variabile reale che godono della proprietà distributiva, Giorn. Math. (2), **4**, 1897, p. 104–111.

[96] G. HAMEL, Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung: $f(x+y) = f(x) + f(y)$, Math. Ann. **60**, 1905, p. 459–462.

[97] Le travail de E. ZERMELO (1871–1953) est de 1904 (Math. Annalen 59, p. 514).

"Weiß man daher, daß in einem ganz beliebigen Bereiche der Variabeln x eine Lösung f(x) der in Rede stehenden Funktionalgleichung irgend einem Werte B nicht beliebig nahe kommt, so darf man schließen, daß f(x) stetig ist und die Form .A.x hat"[96].

(*"Si l'on sait que dans un voisinage donné de la variable x une solution f(x) de l'équation fonctionnelle donnée n'approche pas d'une valeur quelconque B, on doit alors conclure que f(x) est continue et a la forme Ax".*)

Le critère de continuité introduit par HAMEL, déjà envisagé par WILSON pour l'équation de CAUCHY exponentielle, sera prouvé directement en 1946, c'est-à-dire sans recours à l'axiome de ZERMELO et donc en évitant la construction de toutes les solutions de l'équation fonctionnelle[98]. Car on sait que le bon ordre de ZERMELO souleva une polémique intellectuelle considérable chez les mathématiciens et les logiciens[99]. On cherchait donc à éviter son recours dans toute la mesure du possible. Un autre développement avait eu lieu entre temps et provenait encore de DARBOUX. En 1880, celui-ci à nouveau utilise l'équation de CAUCHY afin de trouver toutes les bijections du plan sur lui-même qui transforment une droite en une droite. Il conclut (théorème fondamental de la géométrie projective réelle) qu'une telle bijection est, à une translation près, une transformatlon linéaire non singulière[100]. Sa preuve repose sur la résolution du système fonctionnel,

$$\text{(63)} \qquad \begin{cases} f(x+y) = f(x) + f(y) \\ f(xy) \quad\;\; = f(x)f(y) \end{cases}, \quad (x, y \in \mathbb{R}).$$

La deuxième équation du système (63) fournit bien la non négativité de la fonction f sur les réels positifs, en prenant $x = y$ puisqu'alors $f(x^2) = f(x)^2$, ce qui permet à DARBOUX d'utiliser son résultat de 1875 (*cf.* [92]).

Le problème analogue en géométrie complexe conduit egalément au système (63) mais pour tous x, y dans le plan complexe $\mathbb{C}$. Il est soulevé par C. SÈGRE (1863–1924) en 1890[101]. Dans une lettre adressée à ce mathématicien, H. LEBESGUE (1874–1941) explique les grandes lignes d'une construction possible d'une solution non continue de (63) en faisant usage de l'axiome de ZERMELO[102]. LEBESGUE

---

[98] R. SAN JUAN, An application of Diophantine approximation to the functional equation $f(x_1 + x_2) = f(x_1) + f(x_2)$, Publ. Inst. Math. Univ. Litoral **6**, 1946, p. 221–224.

[99] Au sujet de cette polémique, on pourra consulter la thèse et les documents traduits de J. CASSINET, et M. GUILLEMOT, L'axiome du choix dans les mathématiques de Cauchy (1821) à Gödel (1940), 2 vol., Thèse, Université de Toulouse (P. Sabatier), Juin 1983. Voir aussi le livre en russe de F. MEDVEDEV (1982).

[100] G. DARBOUX, Sur le théorème fondamental de la géométrie projective, Math. Ann. **17**, 1880, p. 55–61.

[101] C. SÈGRE, Un nuovo campo di ricerche geometriche, Atti Accad. Torino, **25** (1889/90), p. 276–301, p. 430–457, p. 592–612 et Atti Accad. Torino, **26** (1890/91). Voir aussi Math. Annalen Vol. **40**, 1892, p. 413–467: Le rappresentazioni reali delle forme complesse e gli enti iperalgebrici.

[102] H. LEBESGUE, Sur les transformations ponctuelles transformant les plans en plans

profite de cette construction pour opposer, en mathématiques, le point de vue "*empiriste*" au point de vue "*idéaliste*", à la suite de DU BOIS-REYMOND. Il commente à partir de la construction d'une solution non continue du système (63), basée sur la possibilité d'un bon ordre de tout ensemble:

"*Cela prouve pour les Idéalistes l'existence d'une infinité de solutions différentes pour votre problème. Le raisonnement qui précède n'a au contraire guère de valeur pour un Empiriste, tout au plus — et seulement pour ceux des Empiristes qui admettent qu'on peut utiliser dans une définition des conditions telles qu'on ne sache ni vérifier qu'elles sont remplies, ni vérifier qu'elles ne le sont pas — montre-t-il que si l'on savait bien ordonner le continu, on saurait nommer des solutions de votre problème autres que les solutions connues*".

Et LEBESGUE de s'interroger sur les solutions de (63) qui sont "*exprimables analytiquement*", c'est-à-dire celles "*dont la valeur s'obtient en effectuant, suivant une loi déterminée, un nombre fini ou dénombrable d'opérations élémentaires (addition, soustraction, multiplication, division, extraction de racine)*" ou "*de passages à la limite à partir de variables ou des constantes*". Utilisant les brillants résultats de la thèse de R. BAIRE (1874–1932), sur les ensembles de 1ère et de 2ème catégorie[103], LEBESGUE conclut que de telles solutions exprimables analytiquement sont nécessairement continues. Une technique de mesure des pathologies est ainsi née à propos d'équations fonctionnelles. LEBESGUE donne à la notion de loi déterminée "*son sens le plus large*" ("*celui de l'Idéaliste*"). Comme il parvient à établir qu'il n'y a pas d'autres solutions du système (63) exprimables analytiquement que les solutions continues, LEBESGUE assure: "*ma conclusion se trouver légitimée pour l'Empiriste*"[102]. A tout le moins l'on constate qu'à cent cinquante ans de distance, après les controverses entre EULER, D'ALEMBERT, *etc.* les équations fonctionnelles cristallisent des affrontements sur le concept de fonction, au sujet cette fois du sens à donner à une construction mathématique et dans le cadre de la crise des fondements liée à la théorie des ensembles et à l'axiome du bon ordre.

Les indications de LEBESGUE quant à la construction d'une solution non continue, ou non régulière, du système fonctionnel (63), sont un peu sommaires. Certes, cette construction peut résulter de théorèmes généraux d'extension des corps commutatifs, tels que ceux fournis par E. NOETHER ou E. ARTIN. En 1947, B. SEGRE fournit pourtant une preuve élémentaire[104], de même que H. KESTELMAN en 1951, lequel commente avec pédagogie l'histoire du problème[105]. Beaucoup plus tard, J. MIKUSINSKI réduit d'une autre façon le problème géométrique à l'origine de (63), ainsi qu'un problème d'optique théorique, à la résolution de la

---

qu'on peut définir par des procédés analytiques, Accad. Reale delle Scienze di Torino (1906/70) p. 219–226.

[103] R. BAIRE, Sur les fonctions de variables réelles (thèse), Annali di Matematica, 1900.

[104] B. SEGRE, Gli automorfismi del corpo complesso, ed un problema di Corrado Sègre, Atti Accad. Naz. Lincei, Rend. **8**, 1947, 3, p. 414–420.

[105] H. KESTELMAN, Automorphisms in the field of complex numbers, Proc. London Math. Soc. 2, **53**, 1951, p. 1–12.

seule équation de CAUCHY (52), sans la seconde équation fonctionnelle du système (63). Toutefois il doit ajouter une condition supplémentaire:

$$(64) \qquad f(x+y) = f(x) + f(y) \quad \textit{lorsque} \quad f(x+y) \neq 0.$$

En 1973, il est démontré que toute solution de (64) pour $f:\mathbb{R}\to\mathbb{R}$ est en fait solution de l'équation (52) de CAUCHY, et ceci sans faire aucune hypothèse de régularité quant à la solution envisagée[106]. Autrement dit la condition $f(x+y)\neq 0$, est superflue. Ce premier résultat inaugure tout une série de travaux sur les *équations de Cauchy conditionnelles*, c'est-à-dire l'équation (52) soumise à une condition, une restriction, portant sur les variables x et y (comme la restriction $f(x+y)\neq 0$). De façon étonnante, ces travaux ont trouvé des applications variées, par exemple dans le cadre de la géométrie des espaces de BANACH, de la théorie cinétique des gaz, de l'économie mathématique, de la théorie des nombres ou de la théorie de l'information, *etc.* ce qui témoigne de l'actualité contemporaine de l'équation de CAUCHY[113]. On peut citer à titre d'exemple le problème posé par P. ERDÖS[107] et résolu en 1965 par W. B. JURKAT[108] et par de BRUIJN[109] (généralisé plus tard par R. GER[110]). Trouver les fonctions $f:\mathbb{R}\to\mathbb{R}$ telles que

$$(65) \qquad f(x+y) = f(x) + f(y)$$

pour tous (x, y) appartenant à un ensemble donné de $\mathbb{R}^2$ dont le complémentaire est de mesure de Lebesgue nulle dans $\mathbb{R}^2$?

On peut montrer que pour toute solution f de (65), il existe une solution g de l'équation de CAUCHY (52) telle que $f(x) = g(x)$ pour tout x de $\mathbb{R}\setminus E$ où E est un ensemble de mesure nulle.

On peut aussi citer le problème le plus ancien, également posé par P. ERDÖS: Trouver $f: N\to\mathbb{R}$ telle que

$$(66) \qquad f(mn) = f(m) + f(n),$$

lorsque m et n sont des entiers premiers entre eux? En 1946, P. ERDÖS[111] établissait le premier résultat de régularité selon lequel une solution monotone de (63) est un logarithme et inaugurait toute une branche de la théorie additive des nombres. Par ailleurs, le résultat d'ERDÖS intervient en théorie de l'information pour caractériser l'entropie[112].

Contentons-nous d'un dernier exemple de travaux sur les équations de CAUCHY conditionnelles. Cette fois nous prenons des variables x sur un espace vectoriel normé réel E. Si la norme dérive d'un produit scalaire, comme c'est le cas de l'es-

---

[106] L. DUBIKAJTIS, C. FERENS, R. GER & M. KUCZMA, On Mikusinski's functional equation, Annales Polonici Math. **28** (1973), p. 39–47.

[107] P. ERDÖS, Coll. Math. **7** (1966), p. 310–311.

[108] W. B. JURKAT, On Cauchy's functional equations, Proc. Amer. Math. Soc. **16** (1965), p. 683–686.

[109] N. G. DE BRUIJN, On almost additive functions., Colloq. Math. **15** (1966), p. 59–63.

[110] R. GER, Note on almost additive functions, Aeq. Math. **17** (1979), p. 73–76.

[111] P. ERDÖS, On the distribution function of additive functions, Ann. of Math. **47** (1946), p. 1–20.

[112] J. ACZÉL et Z. DARÓCZY, On measures of information and their characterizations, Math. in Science and Eng., Vol. 115, Academic Press, 1975.

pace euclidien de la géométrie usuelle, en posant $f(x) = \|x\|^2$, on a une écriture possible du théorème de PYTHAGORE sous la forme

$$(67) \qquad f(x+y) = f(x) + f(y) \quad \text{lorsque } \langle x, y \rangle = 0.$$

C'est-à-dire qu'il existe une solution non linéaire de l'équation de CAUCHY conditionnelle (67), la condition n'étant autre que l'orthogonalité des vecteurs x et y. Il est possible d'établir que cette situation caractérise les espaces dont la norme dépend d'un produit scalaire, ce qui est une belle réciproque du théorème de PYTHAGORE ([113] pour des références sur les équations de CAUCHY conditionnelles).

Mais les équations de CAUCHY conditionnelles nous entraînèrent trop près de nous et il nous faut revenir au début de ce siècle pour évoquer l'aspect lié à la théorie de l'intégration et au début de la topologie moderne.

Dès l'introduction de nouvelles notions de régularité pour les fonctions, à savoir la mesurabilité au sens de LEBESGUE et l'intégrabilité, la question se pose de sélectionner parmi toutes les solutions de (52) fournies par HAMEL celles qui possèdent une telle régularité[114]. Une nouvelle fois, l'équation fonctionnelle sert de faire-valoir à un concept nouveau quant aux fonctions. Le premier à attaquer le problème est M. FRÉCHET (1878–1973) en 1913 lequel prouve qu'une solution mesurable au sens de LEBESGUE de l'équation (52) est continue[115]. FRÉCHET indique que ce résultat contient le cas des solutions de (52) exprimables analytiquement au sens de LEBESGUE puisque celles-ci sont mesurables. Dans la dynamique qui les conduit à se spécialiser sur la topologie naissante, une pléiade de mathématiciens polonais, vers 1920, examine la même question, essayant d'éliminer l'utilisation de l'axiome de ZERMELO, ou de généraliser le résultat de FRÉCHET à plusieurs dimensions. Ainsi, W. SIERPIŃSKI (1882–1969) obtient le résultat de FRÉCHET sans l'axiome de ZERMELO, mais par utilisation d'un lemme astucieux assurant que, deux sous ensembles mesurables au sens de LEBESGUE étant donnés sur la droite réelle, et chacun de mesure strictement positive, il existe au moins un point dans l'un et un point dans l'autre dont la distance s'exprime par un nombre rationnel[116]. Il élargit d'ailleurs le résultat aux fonctions $f\colon \mathbb{R}^2 \to \mathbb{R}$

---

[113] Deux travaux en 1978 font le point sur les équations fonctionnelles conditionnelles et fournissent en particulier les références historiques: M. KUCZMA, Functional equations on restricted domains, Aeq. Math. **18** (1978), p. 1–34. J. DHOMBRES et R. GER, Conditional Cauchy equations, Glasnik Math., 13, 33 (1978), p. 39–62.

Pour des références récentes, voir J. DHOMBRES, Recent applications of functional equations, p. 67–92, in Functional equations: History, applications and theory, J. ACZÉL (ed.), D. Reidel, Dordrecht, 1984. Voir aussi l'ouvrage cité en note [72].

[114] Le point sur l'équation de CAUCHY au début du XXème siècle est fait dans R. SCHIMMACK, *Dissertation*, Halle, 1908 et S. PINCHERLE, *Encyclopédie des Sciences Mathématiques*, Equations et opérations fonctionnelles, Vol. 115, 1, 11.26, p. 45 et suivantes, Paris, Leipzig, 1912.

[115] M. FRÉCHET, Pri la funkcia equatio $f(x+y) = f(x) + f(y)$, Enseignement Math. **15**, p. 390–393.

[116] W. SIERPIŃSKI, Sur l'équation fonctionnelle $f(x+y) = f(x) + f(y)$, Fund. Math. (1), 1920, p. 116–122.

satisfaisant l'équation de CAUCHY. SIERPIŃSKI montre aussitôt dans un papier succédant au précédent[117] que toute fonction mesurable $f:\mathbb{R}\rightarrow\mathbb{R}$ satisfaisant l'inégalité de JENSEN (1859–1925)

$$2f\left(\frac{x+y}{2}\right)\leq f(x)+f(y), \tag{68}$$

est continue. Cette fois la démonstration se fait par l'absurde. La discontinuité en un point x_0 entraîne, pour tout $k>0$ et pour tout $\sigma>0$, l'existence d'un sous-ensemble mesurable de $(x_0-\sigma, x_0+\sigma)$ et de mesure σ au moins, sur lequel f reste au moins égale à k. Ce qui contredit une propriété des fonctions mesurables, propriété établie par E. BOREL (1871–1956) et selon laquelle il existe un k tel que l'ensemble des x de $(x_0-\sigma, x_0+\sigma)$ où f est au moins égale à k, est de mesure strictement inférieure à σ. Qui peut le plus, peut le moins et comme toute solution de (52) est solution de (68), on retrouve le résultat de FRÉCHET. Hélas, comme le signale SIERPIŃSKI, le théorème utilisé de BOREL dépend quant à lui de l'axiome de ZERMELO! Il faut dire que SIERPIŃSKI était bien préparé par un article de 1918 pour traquer cet axiome (et en noter l'utilité) (*cf.* [118]). Intercalant son travail entre les deux travaux mentionnées de SIERPIŃSKI, S. BANACH (1892–1945) prouve la continuité des solutions mesurables de l'équation de CAUCHY (52) au moyen d'un théorème de N. LUSIN (1883–1950) assurant qu'une fonction mesurable, pour tout $\sigma>0$, coincide avec une fonction continue, sauf peut-être sur un ensemble de mesure strictement inférieure à σ [119]. Mais le théorème de LUSIN utilisé (1912) dépend lui aussi de l'axiome de ZERMELO. En 1924, SIERPIŃSKI en répondant à une question de O. NIKODYM prouve qu'une solution de (52), majorée partout par une fonction mesurable au sens de LEBESGUE, est nécessairement continue[120]. KAC en 1936 retrouve le résultat de FRÉCHET en passant à $e^{if(x)}$ et par intégration locale[121]. Il faut d'ailleurs attribuer l'originalité de ce procédé de régularisation des solutions de certaines équations fonctionnelles à J. ANDRADE qui l'avait élaboré pour résoudre l'équation de D'ALEMBERT[122].

---

[117] W. SIERPIŃSKI, Sur les fonctions convexes mesurables, Fund. Math. (1), 1920, p. 125–129.

[118] W. SIERPIŃSKI, L'axiome de Zermelo et son rôle dans la théorie des ensembles et l'analyse, Bull. Acad. Sc. Cracovie, Série A, Avril 1918, p. 97.

[119] S. BANACH, Sur l'équation fonctionnelles $f(x+y)=f(x)+f(y)$, Fund. Math. (1), 1920, p. 123–124.

[120] W. SIERPIŃSKI, Sur les propriétés des fonctions de M. Hamel, Fund. Math. (5), 1924, p. 334–336.

[121] P. KAC, Une remarque sur les équations fonctionnelles, Comm. Math. Helv. **9**, (1936/37), p. 170–171.

[122] La technique de régularisation par intégration locale des solutions d'une équation fonctionnelle remonte à J. ANDRADE, Sur l'équation fonctionnelle de Poisson, Bull. Soc. Math. France **28**, 1900, p. 58–63.

Ce dernier montre tout l'intérêt de la résolution de l'équation fonctionnelle qualifiée aujourd'hui d'équation de D'ALEMBERT sous l'hypothèse d'intégrabilité. D'une part, comme l'indiquait J. TANNERY, pour la définition des fonctions de trigonométrique circulaire (*cf.* [44]). D'autre part pour le parallélogramme des forces (voir aussi le travail de R. SCHIMMACK cité en [114]) et enfin pour les prioprétés métriques en géométrie non

Au fil des années de nombreux auteurs vont découvrir les résultats de FRÉCHET et SIERPIŃSKI ou proposer d'autres démonstrations. Par exemple A. ALEXIEWICZ et W. ORLICZ[123] ou H. KESTELMAN[124]. Une technique particulièrement habile repose sur un résultat de H. STEINHAUS (1887–1972), résultat paru précisément dans le premier numéro de Fundamenta Mathematica, celui-là même où publièrent BANACH et SIERPIŃSKI. On peut s'étonner que le résultat de STEINHAUS n'ait pas aussitôt été exploité pour l'équation de CAUCHY. STEINHAUS prouve que l'ensemble des $x + y$, lorsque x et y parcourent chacun le même ensemble E dont la mesure de LEBESGUE est (strictement) positive sur $\mathbb{R}$, contient nécessairement un intervalle non vide[125]. Comme un intervalle non vide contient toujours un nombre rationnel, on note la proximité avec le lemme utilisé par SIERPIŃSKI dans son travail de 1920[116]. Ce sera M. KORMES en 1926 qui le premier exploitera le résultat de STEINHAUS dans le cadre de l'équation de CAUCHY afin d'établir qu'une solution de (52), bornée supérieurement sur un ensemble de mesure de LEBESGUE (strictement) positive, est continue[126]. Le théorème de FRÉCHET et ceux de SIERPIŃSKI s'en déduisent aisément. A. OSTROWSKI établit le résultat analogue pour le cas des fonctions convexes au sens de JENSEN[127]. Il faut dire ici que les fonctions convexes au sens de JENSEN, ou les fonctions convexes usuelles, connurent un grand succès après 1920. Rappelons ici qu'une fonction convexe satisfait l'inégalité fonctionnelle

$$f(\lambda x + (1 - \lambda)\, y) \leqq \lambda f(x) + (1 - \lambda) f(y), \tag{69}$$

pour tous x, y du domaine de définition et tout λ appartenant à l'intervalle $[0, 1]$. Le cas de JENSEN est donc la restriction de l'inégalité (69) au seul cas $\lambda = \frac{1}{2}$ et était déjà envisagé par CAUCHY dans son cours de 1821. Une étape marquante pour l'intérêt porté à la convexité fut obtenue en 1922 lorsque H. BOHR et J. MOLLERUP[128], prouvèrent que la fonction $\Gamma(x)$, étudiée par EULER, GAUSS et LEGENDRE, était l'unique solution (strictement positive) de l'équation fonctionnelle

$$f(x + 1) = x f(x), \tag{70}$$

valant 1 en $x = 1$ et dont le logarithme est une fonction convexe sur $[0, +\infty[$. Moins de dix ans plus tard, dans un petit livre d'une magistrale élégance,[129], E. ARTIN part de cette caractérisation pour déduire l'essentiel des propriétés de

---

euclidienne (sur ce dernier point voir J. ACZÉL[91], ou J. DHOMBRES et J. ACZÉL[72], auquel nous référons pour les développements récents quant à l'équation de D'ALEMBERT).

[123] A. ALEXIEWICZ et W. ORLICZ, Remarque sur l'équation fonctionnelle $f(x + y) = f(x) + f(y)$, Fund. Math. **33**, 1945, p. 314–315.

[124] H. KESTELMAN, On the functional equation $f(x + y) = f(x) + f(y)$, Fund. Math. **34**, (1947), p. 144–147.

[125] H. STEINHAUS, Sur les distances des points des ensembles de mesure positive, Fund. Math. **1**, 1920, p. 93–104.

[126] M. KORMES, On the functional equation $f(x + y) = f(x) + f(y)$, Bull. Amer. Math. Soc. **32**, (1926), p. 689–693.

[127] A. OSTROWSKI, Über die Funktionalgleichung der Exponentialfunktion und verwandte Funktionalgleichungen, Jahresber. Deut. Math. Verein, 38 (1929), p. 54–62.

[128] H. BOHR et J. MOLLERUP, *Laerebog i matematisk Analyse* III, Copenhague, 1922.

[129] E. ARTIN, *Einführung in die Theorie der Gammafunktion*, Leipzig, Teubner, 1931.

la fonction Γ, cette fonction qui interpole la fonction factorielle aux valeurs entières de l'argument. Quelques années après, en 1934, l'intérêt de la convexité est à nouveau mis en valeur et cette fois pour longtemps grâce à un livre phare de G. H. HARDY, J. E. LITTLEWOOD et G. PÓLYA, un de ces livres par lesquels tout une génération de mathématiciens fut éduquée[130]. Cet ouvrage contient de nombreuses précisions historiques et son intérêt mathématique est de donner à des calculs combinatoires une justification d'analyse fonctionnelle (notion de moyenne aussi bien discrète qu'au sens intégral avec les inégalités de HÖLDER ou de MINKOWSKI, inégalités de réarrangement sur la transformée fonctionnelle de HILBERT *etc.*). De très nombreux mathématiciens vont concentrer leur attention sur les fonctions convexes à une ou à plusieurs variables. Nous n'insisterons pas sur cet aspect de la convexité puisqu'un aperçu historique fût dressé par S. MARCUS en 1959[131] et qu'un livre mathématique à paraître de M. KUCZMA, sur les équations fonctionnelles à une variable, contient une trame historique particulièrement fine[132].

L'énumération un peu sèche des résultats sur l'équation fonctionnelle (52) a pu faire oublier le moteur de cette recherche autour des années 1920. En fait, il s'agissait d'évaluer la force et la maniabilité des nouveaux outils que le début du XX^ème^ siècle venait de découvrir: intégration et topologie, le tout installé dans la théorie des ensembles de CANTOR. L'équation de CAUCHY sert de travaux pratiques. Citons quelques exemples. De nombreux papiers s'intéressent aux problèmes de la mesurabilité liés à une base de HAMEL. En 1920, SIERPIŃSKI montre[133] qu'il existe une base de HAMEL mesurable au sens de LEBESGUE (et de mesure nulle) mais déduit d'un résultat de C. BURSTIN[135] qu'il existe également une base de HAMEL non mesurable (mais nécessairement de mesure intérieure nulle). Il indique aussi que toute base de HAMEL n'est pas mesurable au sens de BOREL. De même de nombreux travaux considèrent la décomposition de la droite réelle en ensembles non mesurables à partir de l'article de BURSTIN et d'articles de LUSIN et de SIERPIŃSKI publiés vers 1918. S. RUZIEWICZ[134] utilise astucieusement l'équation de CAUCHY pour obtenir une décomposition de la droite en ensembles superposables non mesurables (c'est-à-dire déduits les uns des autres par translations

---

[130] G. H. HARDY, J. E. LITTLEWOOD et G. POLYA, *Inequalities*, Cambridge University, Press, 1934.

[131] S. MARCUS, Généralisation aux fonctions de plusieurs variables, des théorèmes de Alexander Ostrowski et Masuo Hukuhara concernant les fonctions convexes (J), Journ. Math. Soc. Japan **11**, (1959), n° 3, p. 171–175.

[132] Le livre de M. KUCZMA devrait paraître chez Addison-Wesley. Mais on trouvera déjà des indications dans M. KUCZMA, *Functional equations in a single variable*, Monografie Matematyczne, Warszawa, 1968.

[133] W. SIERPIŃSKI, Sur la question de la mesurabilité de la base de M. Hamel. Fund. Math. **1**, 1920, p. 105–111.

[134] C. BURSTIN, Die Spaltung des Kontinuums in c in Lebesgueschem Sinne nichtmessbare Mengen, Sitzungsber Akad. Wiss., Wien, Math. Nat. Kl. Abt. IIa, Bd. 125 (1916).

[135] S. RUZIEWICZ, Une application de l'équation fonctionnelle $f(x+y) = f(x) + f(y)$ à la décomposition de la droite en ensembles superposables, non mesurables, Fund. Math., 1923, p. 92–95.

ou symétries). Il utilise les ensembles

$$E_a = \{x \mid x \in \mathbb{R}, f(x) = a\},$$

où f est une solution discontinue de (52). Généralisés, ces problèmes conduisent au fameux paradoxe de BANACH-TARSKI (1923) sur les décompositions finies ou dénombrables de $\mathbb{R}^n$ ($n \geqq 3$). Une autre avancée est obtenue par SIERPIŃSKI dans un article paru en 1924 [120]. Reprenant la démarche de LEBESGUE[102], SIERPIŃSKI montre que l'on peut agir de façon quantitative en topologie comme en théorie de la mesure, sans changements notables, grâce à la notion des catégories de BAIRE[103]. Il indique ainsi que toute solution de (52) sujette à une condition de BAIRE, est nécessairement continue. Cette technique est aujourd'hui systématiquement employée et on trouvera de remarquables illustrations dans le livre d'OXTOBY[136]. Il est pourtant notable que l'analogue, en topologie, du théorème de STEINHAUSS énoncé en 1920 pour la théorie de la mesure ne sera trouvé par S. PICCARD que vers les années 1940 [137]. Il est également notable que pour mettre en évidence les propriétés pathologiques de la connexité, sujet pourtant fort à la mode vers les années 1920, on n'ait pas utilisé l'équation de CAUCHY avant 1942, alors que cette équation servait depuis plus de vingt années déjà à exhiber des pathologies des ensembles non mesurables[138].

Toujours à partir des années 20, un autre système d'équations fonctionnelles faisant intervenir l'équation de CAUCHY, celui caractérisant les opérations qualifiées de linéaires, prenait un rôle désormais fondamental en Analyse. Il s'agit du système:

$$\begin{cases} f(x+y) = f(x) + f(y), \\ f(\lambda x) = \lambda f(x). \end{cases} \tag{71}$$

Mais dans la mise en évidence de (71), le plus difficile sans doute fut d'éprouver le besoin de spécifier axiomatiquement le domaine de validité des x, des y et des λ. Car bien entendu on utilisait dans des cas particuliers et depuis fort longtemps, des solutions de (71), qualifiées de "substitutions linéaires". Cette utilisation atteint chez LEIBNIZ une précision certaine et se retrouve en arithmétique, en géométrie aussi bien que pour des équations différentielles chez EULER, LAGRANGE, GAUSS ou CAUCHY. Nous ne pouvons entrer dans le détail de l'histoire de l'algèbre linéaire, et renvoyons à la littérature historique[139]. Contentons-nous de dire qu'avec CAYLEY le calcul de cette algèbre linéaire est devenu systématique tandis que GRASSMANN dévelope son caractère intrinsèque. L'aspect structurel et axiomatique n'apparaît qu'avec G. PEANO lequel en 1888 définit un espace vectoriel E sur le

---

[136] J. OXTOBY, *Measure and category*, Springer-Verlag, 1972.

[137] S. PICCARD, Sur les ensembles de distances des ensembles de points d'un espace euclidien. Gauthier-Villars, Paris, 1939.

[138] F. B. JONES, Measure and other properties of a Hamel basis, Bull. Am. Math. Soc., **48**, (1942), p. 472–482.

[139] On peut citer les deux chapitres consacrés à l'histoire de l'algèbre dans J. DIEUDONNÉ, *Abrégé d'histoire des mathématiques*, Vol. 1. Hermann, Paris, 1978.
M. J. CROWE, *A history of vector analysis*, University of Notre Dame Press, 1967.
L. NOVY, *Origins of modern algebra*, Leyden, Noordhoff, 1973.

corps des réels. Sur un tel espace E, il y a donc une addition (opération interne; si x et y sont dans E, de même $x + y$) et une multiplication (opération externe; si x est dans E et λ un nombre réel, λx est dans E), opérations liées par quelques propriétés simples. Avec cela, PEANO peut définir les opérations linéaires comme celles satisfaisant (71) pour tous x, y de E et tout λ réel, la valeur de $f(x)$ se prenant dans un second espace vectoriel réel[140]. Son exposé s'adapte aussi bien aux manipulations de l'algèbre linéaire de la géométrie ordinaire, c'est-à-dire ce que nous appelons la dimension finie, qu'à celles issues des équations différentielles, en dimension infinie. Il faut noter à ce propos le rôle des équations différentielles du type de STURM-LIOUVILLE systématiquement étudiées[141] à partir de 1836, puis reprises sous forme intégrale par FREDHOLM à partir de 1903 et magistralement placées dans un cadre explicatif "géométrique" par D. HILBERT, cadre rendu abstrait vers 1910 par F. RIEZ et M. FRÉCHET[142]. A la même époque, O. TOEPLITZ réalise l'ambition de PEANO, celle de ne pas séparer les cas de dimension finie et infinie en indiquant les méthodes algébriques qui évacuent la nécessité des déterminants[143]. Il indiquait aussi la possibilité de remplacer sans changement notable le corps des nombres réels par tout corps commutatif. S. BANACH, H. HAHN et E. HELLY systématisent la théorie des opérations linéaires au profit de l'analyse et tous font de (71) le thème explicite de leur étude. C'est BANACH qui publie en 1932 un condensé de l'oeuvre très riche déjà accumulée, en indiquant à titre d'excellente publicité l'application à des problèmes anciens. Mais son livre est beaucoup plus un programme à réaliser de problèmes pendants[144]. Programme qui occupera les analytes jusqu'à aujourd'hui même si entre 1940 et 1960 l'accent sera mis sur un cadre plus abstrait, celui des espaces vectoriels topologiques, dans lequel venait naturellement s'inscrire la théorie des distributions[145]. Un retour aux idées de

---

[140] G. PEANO, Calcolo geometrico secondo l'Ausdehnungslehre di Grassmann, preceduto dalle operazioni della logica deduttiva, Torino, 1888.

[141] C. STURM, Sur les équations différentielles linéaires du second ordre, Journal de Math. **1**, t. 1, 1836, p. 106–186; J. LIOUVILLE, Sur le développement des fonctions ou parties de fonctions en séries dont les divers termes sont assujettis à satisfaire une même équation différentielle du second ordre contenant un paramètre variable, Journal de Math. 1, t. 1, 1836, p. 253–265.

[142] Nous ne pouvons pas citer ici toutes les publications concernées et nous nous contentons de renvoyer à la littérature historique et critique, principalement J. DIEUDONNÉ, *A history of functional analysis*, North-Holland, 1981, et au lendemain des travaux de HILBERT, RIESZ et FRÉCHET à l'analyse poussée de F. HELLINGER & O. TOEPLITZ, Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten, Enzy. der Math. Wiss. Teubner, 1923–27. Vol. **2**, Part. 3, p. 1335–1597. Voir M. BERNKOPF, The development of function spaces with particular reference to their origins in integral equation theory, Archive for History of Exact Sciences, **3**, 1966, p. 1–96.

[143] O. TOEPLITZ, Über die Auflösung unendlichvieler linearer Gleichungen mit unendlichvielen Unbekannten. Rend. Cir. Mat. Pal., 28, 1909, p. 88–96.

[144] S. BANACH, *Théorie des opérations linéaires*, Varsovie, 1932.

[145] Là-aussi nous ne pouvons citer toutes les oeuvres qui ont forgé la théorie des espaces vectoriels topologiques et la théorie des distributions. Nous renvoyons à la première référence citée en [142] et à J. LÜTZEN, *The prehistory of the theory of distributions*, Springer-Verlag, 1981.

Par ailleurs, beaucoup d'exposés des espaces vectoriels topologiques et de la théorie

BANACH se faisait vers 1970[146]. Ce que BANACH montrait, c'est la possibilité de récupérer certaines propriétés de E à partir de son dual - le vocabulaire est volontairement géométrique — c'est-à-dire de toutes les solutions définies sur E et à valeurs réelles de (71), voire même des seules solutions continues de (71) lorsque E est muni d'une structure normée[147]. Nous ne pouvons en dire plus ici car nous sortirions des limites que nous nous sommes imposées quant aux mathématiques décrites[148]. Toutefois signalons que H. HAHN, par une habile utilisation de l'axiome du choix, avait établi en 1927 une condition permettant l'extension d'une solution continue de (71) à un espace vectoriel normé englobant E[149]. Cette même technique, connue sous le nom du théorème de HAHN-BANACH, permit en 1970 de préciser remarquablement la structure des solutions discontinues de l'équation de CAUCHY[150].

L'équation fonctionnelle de CAUCHY apparaît encore dans le cadre d'un développement contemporain très puissant celui de l'étude des semi-groupes d'opérateurs linéaires. Il s'agit des solutions de l'équation fonctionnelle

$$f(x+y) = f(x) \circ f(y), \tag{72}$$

où les variables x et y appartiennent à un semi-groupe tandis que $f(x)$ pour chaque x est un opérateur linéaire continu sur un espace de BANACH. Dans (72), le signe $\circ$ au second membre désigne la composition de tels opérateurs. La théorie des semi-groupes d'opérateurs, depuis la monumentale monographie de E. HILLE et R. S. PHILLIPS en 1957 a connu une grande extension, notamment en vue des

---

des distributions contiennent des notices historiques fournies. On peut citer N. BOURBAKI, *Espaces vectoriels topologiques*, Hermann, 1964, J. L. KELLEY, *Linear topological spaces*, Van Nostrand, 1960, et surtout pour les opérateurs linéaires, l'encyclopédie monumentale N. DUNFORD et J. T. SCHWARTZ, *Linear operators, Vols. I, II and III*, Interscience Publishers, New York, 1958–1971.

[146] Notamment avec le contre-exemple apporté à une conjecture de BANACH sur l'approximation de tout opérateur compact par des opérateurs de rang fini. P. ENFLO, A counter-example to the approximation problem in Banach spaces, Acta Math. **130** (1973), p. 309–317.

[147] En 1911, F. RIESZ avait caractérisé la continuité d'une solution du système (71) au moyen d'une inégalité. F. RIESZ, Sur certains systèmes singuliers d'équations intégrales, Ann. Ecole Normale Sup. (3), t. **28**, 1911, p. 33–62.

[148] Mais il faut souligner qu'avec la notion de dualité et donc celle de distribution telle que développée aussi bien par SOBOLEV à partir de 1936 que par L. SCHWARTZ, après guerre, une nouvelle aventure du concept fonctionnel et des équations fonctionnelles fut instaurée.

[149] H. HAHN, Über lineare Gleichungssysteme in linearen Räumen, J. de Crelle, t. 157, 1927, p. 214–229.

Les titres des oeuvres voisines de BANACH et HELLY portent tous sur le système (71). E. HELLY, Über Systeme linearer Gleichungen mit unendlichvielen Unbekannten, Monatshefte für Math. und Physik, t. **31**, 1921, p. 60–91. S. BANACH, Sur les fonctionnelles linéaires, Studia Math., **1**, 1929, p. 211–216 et p. 223–239.

[150] MARCIN E. KUCZMA, On discontinuous additive functions, Fund. Math., **66**, 1970, p. 383–392.

équations d'évolution en théorie des équations aux dérivées partielles[151]. La perturbation de tels semi-groupes et leur stabilité a des conséquences pratiques importantes[151]. Contentons-nous d'un résultat de stabilité pour l'équation de CAUCHY (52), résultat prouvé en 1941 par HYERS et quelque peu étonnant à première vue[152].

Soit $\varepsilon > 0$ et $f: \mathbb{R} \to \mathbb{R}$ une fonction satisfaisant pour tous x, y dans $\mathbb{R}$ l'inéquation

$$|f(x+y) - f(x) - f(y)| \leqq \varepsilon. \tag{73}$$

Il existe alors une solution $g: \mathbb{R} \to \mathbb{R}$ de l'équation de CAUCHY (52) et

$$|f(x) - g(x)| \leqq \varepsilon, \quad (x \in \mathbb{R}). \tag{74}$$

La stabilité des équations fonctionnelles fait l'objet de nombreux travaux contemporains, et on comprendra que nous ne puissions préciser ceux-ci sans quitter le domaine historique et épistémologique et entrer dans celui de la recherche mathématique.

Conclusion

Nous avons dû négliger bien des aspects importants du développement du concept de fonction, notamment en n'insistant pas sur les équations différentielles, les équations aux différences ou les équations aux dérivées partielles. Loin d'être exhaustif, à partir du thème restreint des équations fonctionnelles, notre but était de baliser la pratique des mathématiciens quant aux fonctions en l'opposant éventuellement à des définitions contraignantes qu'eux-mêmes proposèrent. Il existe des définitions claires, facilement admises, pascaliennes en un mot, mais il figure aussi des définitions rhétoriques, suggestives au-delà des mots, et en ce sens, de nature expérimentale. L'aventure mathématique s'en colore d'autant et l'étude épistémologique de cette science doit s'enrichir de tels essais. Seules la pratique et l'analyse minutieuse de la manipulation des définitions dans les démonstrations, avec l'irruption possible de limitations par rapport aux définitions de départ, permettent de distinguer. L'historien des mathématiques est donc contraint à une lecture très attentive des textes et ne peut en rester aux résultats les plus marquants. Ce faisant, une image contrastée de l'évolution des idées fonctionnelles nous semble obtenue et digne d'intérêt. Comme dans toute histoire lue à rebours et en focalisant sur quelques pratiques, on peut distinguer des contributions fortes

---

[151] E. HILLE et R. S. PHILLIPS, *Functional analysis and semi-groups*, A.M.S. Coll. Publ., Vol. XXXI, 1957, 808 p.

Deux ouvrages fournissent quelques notes historiques au fur et à mesure de l'exposé: T. KATO, *Perturbation theory for linear operators*, Springer-Verlag, 2nd ed. 1980.

M. REED & B. SIMON, *Methods of modern mathematical physics*, Academic Press, Vols. I, II, III, IV, 1972.

[152] D. H. HYERS, On the stability of the linear functional equation, Proc. Nat. Acad. Sc. U.S.A., **27**, (1941), p. 222–224.

opposées à des résultats plus brillants que novateurs, souvent parce que leurs inventeurs pressés d'autres applications ne prirent pas soin d'en dégager la signification, et même des apports conduisant à des erreurs.

Avec ORESME et l'équivalence graphique d'une propriété fonctionnelle analytique, GALILÉE et la propriété fonctionnelle de la chute des corps, D'ALEMBERT sans doute avec la séparation fonctionnelle des variables, CAUCHY et la continuité fonctionnelle, LEBESGUE et la liaison fonctionnelle entre intégration et mesure, BANACH enfin sur les opérations linéaires et le lancement de l'analyse fonctionnelle, nous tenons les moments les plus éclatants.

Il apparaît même un caractère presque teigneux chez les auteurs mentionnés, dans la mesure où creusant autour d'un cercle d'idées, la méthode fonctionnelle, ils ne l'isolent pas et l'installent volontiers dans l'architecture même des mathématiques. Le plus notable dans l'histoire du concept de fonction est la permanence d'une tension entre ce cercle *a priori* restreint d'idées, volontairement restreint pour que les définitions et méthodes soient identifiables avec précision, et l'élargissement dynamique potentiel. L'histoire du concept fonctionnel n'est donc pas linéaire mais présente des heurts dialectiques dont nous avons rendu compte.

Jusqu'à nos jours, nous avons aussi pu constater que les équations fonctionnelles sont un lieu où des techniques nouvelles sont expérimentées: ainsi la théorie de la mesure de LEBESGUE, l'algèbre des corps, l'axiome du choix, la topologie de BAIRE et plus près de nous la géométrie des espaces de BANACH ou la représentation des opérateurs linéaires. Régularités aussi bien que pathologies en résultent. Telle est aussi l'utilité – modeste mais attachante – des équations fonctionnelles.

Institut de Mathématiques
Chemin de la Houssinière
Nantes, France

(Manuscrit reçu le 28 décembre 1985)