

SIGNE DE f' ET VARIATIONS DE f : LA FABRIQUE D'UNE CHAÎNE DÉDUCTIVE LONGUE

Renaud CHORLAY
ESPE de l'académie de Paris
SPHERE (UMR 7219 CNRS – Paris Diderot)
IREM de Paris

Résumé. Le théorème relatif au lien entre le signe de f' et les variations de f joue un rôle particulier à la transition secondaire-supérieur : familier mais admis dans le secondaire, il est démontré dans le supérieur dans le cadre du cours sur les fondements de l'analyse. Nous proposons un dossier formé de textes historiques permettant d'illustrer deux aspects de ce théorème : d'une part la multiplicité des idées de preuve (nous en étudierons trois) ; d'autre part, son insertion progressive dans une longue chaîne d'énoncés formant, depuis le début du 20^{ème} siècle, l'entrée standard dans l'analyse réelle du supérieur. Après l'étude de ces textes, nous proposerons deux pistes d'utilisation : pour l'enseignement de l'analyse au niveau L1, et pour la formation des enseignants à l'épistémologie des mathématiques.

Mots clés. Analyse mathématique, dérivation, épistémologie, histoire des mathématiques, démonstration.

Abstract. In France, the theorem which links the sign of the derivative f' to the variations of the function f plays a specific part in the transition from secondary to tertiary education: stated without proof in high-school, the proof is given in tertiary education, as part of the course on the foundations of mathematical analysis. This paper presents a selection of historical texts which shed a twofold light on this theorem: firstly, on the multiplicity of historical proofs (we shall study three of them); secondly, on its progressive insertion into a long deductive chain of theorems; a chain which has been standard in the courses on the foundations of analysis since the beginning of the 20th century. After studying this collection of texts, we shall suggest two potential contexts of use: one in the teaching of mathematical analysis in the first year of higher education, and one in teacher-training, as a means to trigger reflection on the nature of mathematics.

Key-words. Mathematical analysis, derivative, epistemology, history of mathematics, proof.

1. Les théorèmes de l'Analyse en classe post-bac

Le passage du secondaire au supérieur s'accompagne souvent d'une nette intensification des tâches relatives à la démonstration : on attend que l'étudiant comprenne le grand nombre de preuves présentées en cours, et qu'il puisse en produire dans le cadre d'exercices. Un théorème d'analyse semble pouvoir jouer un rôle charnière dans cette transition secondaire-supérieur : soit f une fonction numérique définie et dérivable sur un intervalle, la fonction dérivée f' est positive si et seulement si la fonction primitive f est croissante. Ce théorème est énoncé au Lycée, et justifie la tâche routinière de construction du tableau de « signes de f' et variations de f » ; cependant, seul l'un des deux sens de l'implication est démontrable dans le secondaire (variation \Rightarrow signe), la démonstration de la réciproque relevant typiquement d'un enseignement d'analyse du supérieur. Outre cette place spécifique dans les cursus, le travail sur ce théorème est

d'autant plus intéressant que l'on dispose déjà de deux éclairages complémentaires : l'un sur son histoire (Dugac 1979), l'autre sur les questions didactiques liées à son enseignement, dans le contexte de la transition secondaire-supérieur (Praslon 1994).

Au niveau L1/MPSI¹, ce théorème s'insère dans un ensemble déductif cohérent relatif aux fondements de l'analyse réelle². Précisons que l'affirmation ne porte pas ici sur les pratiques d'enseignement en post-bac dans les filières générales (cursus scientifique des universités, classe préparatoires scientifiques), mais sur les objectifs d'enseignement et architectures démonstratives retenues dans les programmes ou les manuels les plus courants. On pourra par exemple consulter le programme de MPSI (BO spécial n°3 du 30 mai 2013), programme qui constitue également la base du programme de l'épreuve écrite du CAPES de mathématiques.

La démonstration est alors une application directe de l'égalité des accroissements finis³ ; cette égalité est elle-même déduite du « théorème de Rolle »⁴. La démonstration de ce dernier utilise de manière fondamentale le théorème établissant qu'une fonction numérique continue sur un segment (i.e. un intervalle fermé et borné) admet un maximum et un minimum. Quoique bien intuitif, ce dernier résultat peut lui-même se démontrer, en s'appuyant sur des propriétés non-triviales de la droite numérique⁵. Le strict enchaînement de la longue suite de résultats permettant d'établir le théorème visé rend difficile d'imaginer qu'il ait pu être démontré autrement, et qu'une telle *chaîne déductive longue* ait pu être mise au point progressivement. Ce sont pourtant d'autres preuves du théorème final, et une mise en place progressive des éléments de la chaîne déductive que nous allons illustrer par des textes historiques.

Du point de vue historique, si l'on trouve des énoncés de notre théorème dès le début du 19^{ème} siècle, la chaîne déductive longue que nous venons de rappeler se forme à la fin de ce siècle : on la trouve, par exemple, dans les cours de Jordan (Jordan, 1893), Stolz (Stolz, 1893) ou W.F. Osgood (Osgood, 1912) ; elle se transmet de manière *stable* depuis. Au moment de sa formation, elle est explicitement perçue comme illustrant la nouvelle manière d'exposer systématiquement l'analyse, en s'appuyant non seulement sur la notion de limite (déjà fondamentale dans l'enseignement de Cauchy dans les années 1820), mais, en dernier ressort, sur les propriétés des ensembles de points de la droite réelle. Ainsi, dans le chapitre sur le calcul différentiel et intégral de la grande

1 L1 : Licence première année ; MPSI : classe préparatoire Mathématiques, Physique et Sciences de l'Ingénieur.

2 Précisons que l'affirmation ne porte pas ici sur les pratiques d'enseignement en post-bac dans les filières générales (cursus scientifique des universités, classe préparatoires scientifiques), mais sur les objectifs d'enseignement et architectures démonstratives retenues dans les programmes ou les manuels les plus courants. On pourra par exemple consulter le programme de MPSI (BO spécial n°3 du 30 mai 2013), programme qui constitue également la base du programme de l'épreuve écrite du CAPES de mathématiques.

3 Soient $a < b$ deux nombres réels. Si f est une fonction numérique continue sur $[a, b]$ et dérivable sur $]a, b[$, il existe c dans $]a, b[$ tel que $f'(c) = \frac{f(b) - f(a)}{b - a}$. Géométriquement parlant : il existe un point de l'arc de courbe de f où la tangente est parallèle à la corde reliant les points de la courbe d'abscisses respectives a et b .

4 Soient $a < b$ deux nombres réels. Si f est une fonction numérique continue sur $[a, b]$ et dérivable sur $]a, b[$ prenant la même valeur en a et en b , il existe c dans $]a, b[$ tel que $f'(c) = 0$.

5 Par exemple: de toute suite bornée on peut extraire une suite convergente. ; ou : toute suite d'intervalles emboîtés non vide est d'intersection non vide.

Encyclopédie des sciences mathématiques, Voss souligne clairement la dépendance de l'égalité des accroissements finis envers le théorème démontré par Weierstrass assurant l'existence d'extrema pour les fonctions continues sur les segments (Voss, 1899, p.65-66) (Chorlay, 2007b).

Nous proposons ici un dossier documentaire permettant d'explorer la période séparant les premiers énoncés du théorème (début 19^{ème}) de son insertion dans la chaîne stable (fin 19^{ème}). Pour bien les distinguer de notre commentaire, les textes-sources sont présentés dans des encadrés. Les documents historiques sélectionnés permettent d'illustrer deux points différents. Premièrement, les textes de Lagrange, Cauchy et Serret montrent trois justifications différentes du même théorème ; les différences ne sont pas ici de simple forme, les trois justifications reposent sur des idées-de-preuve⁶ réellement différentes : la justification de Lagrange repose sur l'approximation de la fonction f par des fonctions affines par morceaux issues des approximations affines locales fournies par f' ; l'argument de Cauchy repose sur la simple définition du nombre dérivé comme limite ; l'argument de Serret repose sur l'égalité des accroissement finis, elle-même établie de manière semi-intuitive.

Nous aurions pu nous arrêter au texte de Serret, puisqu'il propose peu ou prou la preuve actuelle. Nous avons cependant choisi, dans la cinquième partie, de compléter le dossier historique par des textes illustrant des épisodes de la formation de la chaîne déductive longue dans laquelle le théorème se trouve inséré de manière stable depuis le début du 20^{ème} siècle. Le théorème étant valide, aucun contre-exemple ne peut lui être opposé ; en revanche, les preuves peuvent en être critiquées de plusieurs manières : telle preuve peut reposer sur un Lemme erroné, susceptible d'être infirmé par ce que Lakatos appelle un contre-exemple *local* (Lakatos 1976) ; on peut déceler dans une preuve la présence de Lemmes cachés, demandant à être explicités et établis ; on peut aussi critiquer une démonstration sans remettre en question sa validité, en considérant qu'elle n'est pas « éclairante », qu'elle manque de « généralité », de « simplicité » etc. Deux aspects de ce travail critique d'*analyse de preuve*, mené par les mathématiciens eux-mêmes, seront illustrés dans cette 5^{ème} partie. Quoi qu'ayant longtemps été reprise dans les manuels du 19^{ème} siècle, la preuve de Lagrange sera finalement abandonnée, et cet abandon n'est qu'un épisode d'un mouvement plus général et bien connu de l'histoire de l'analyse : celui de la distinction progressive entre propriétés ponctuelles et propriétés uniformes, de la critique des hypothèses *implicites* d'uniformité (ici, d'uniforme dérivabilité) dans les démonstrations de la première moitié du 19^{ème} siècle. Nous illustrerons ce point par la critique que fait Peano d'une preuve de l'égalité des accroissements finis donnée par Jordan. Le second aspect est d'une nature épistémologique différente. La chaîne déductive longue et stable fait reposer la démonstration de l'égalité des accroissements finis sur le théorème relatif aux extrema des fonctions continues sur les segments ; ce théorème reposant lui-même sur la définition de la continuité en un point, et sur les propriétés topologiques de la droite numérique. C'est en particulier dans les cours de Weierstrass que cette partie de la chaîne déductive se met en place, nous en donnerons donc des extraits, en traduction inédite.

Que ce soit par ses objectifs, sa problématisation ou son mode de constitution du corpus, cet article ne constitue pas un travail de recherche en histoire des

6 Nous reprenons ici la notion de *proof-idea* ou *proof-germ* proposée dans Downs et Mamona-Downs (2010), à distinguer de la preuve comme produit final et rédigé.

mathématiques. Il s'agit d'un travail *s'appuyant* sur l'histoire des mathématiques, dont nous espérons qu'il pourra permettre à chacun – enseignant du secondaire, du supérieur, formateur d'enseignant – d'enrichir sa vision de ce théorème classique, tout en éprouvant le plaisir (parfois, la difficulté) de lire des textes historiques. La lecture de ce dossier de textes nous semble fournir des pistes de réflexion et de prolongement, de plusieurs natures. Nous en indiquerons deux dans la sixième et dernière partie : la première piste concerne l'enseignement des mathématiques au niveau L1, et suggère des modes d'utilisations de parties de ce dossier documentaire ; la seconde piste concerne la formations en épistémologie des mathématiques des futurs enseignants. La formulation de ces deux pistes constitue, en outre, un appel à collaboration.

2. L'énoncé et la preuve de Lagrange (1806)

Le neuvième chapitre des *Leçons sur la théorie des fonctions* (Lagrange, 1884) de Joseph-Louis Lagrange (1736-1813) s'intitule : *De la manière d'avoir les limites du développement d'une fonction, lorsqu'on n'a égard qu'à un nombre fini de termes*. Il s'agit donc d'établir ce que nous nommons l'inégalité de Taylor-Lagrange, qui majore l'écart entre une fonction et son développement polynomial à l'ordre $n-1$ au moyen d'une majoration (en valeur absolue) de la dérivée n -ième.

Citons le début du chapitre, dans lequel le théorème qui nous intéresse est d'abord motivé, puis énoncé et démontré.

Toute fonction $f(x+i)$ se développe, ainsi qu'on l'a vu, dans la série

$$f(x) + i f'(x) + \frac{i^2}{2!} f''(x) + \frac{i^3}{3!} f'''(x) + \dots$$

laquelle va naturellement à l'infini, à moins que les fonctions dérivées de $f(x)$ ne deviennent nulles, ce qui a lieu lorsque $f(x)$ est une fonction rationnelle et entière⁷ de x .

Tant que ce développement ne sert qu'à la génération des fonctions dérivées, il est indifférent que la série aille à l'infini ou non; il l'est aussi lorsqu'on ne considère le développement que comme une simple transformation analytique de la fonction; mais, si on veut l'employer pour avoir la valeur de la fonction dans les cas particuliers, comme offrant une expression d'une forme plus simple à raison de la quantité i qui se trouve dégagée de dessous la fonction, alors, ne pouvant tenir compte que d'un certain nombre plus ou moins grand de termes, il est important d'avoir un moyen d'évaluer le reste de la série qu'on néglige, ou du moins de trouver des limites de l'erreur qu'on commet en négligeant ce reste.

La détermination de ces limites est surtout d'une grande importance dans l'application de la Théorie des fonctions à l'Analyse des courbes et à la Mécanique, pour pouvoir donner à cette application la rigueur de l'ancienne Géométrie, comme on le voit dans la seconde Partie de la *Théorie des fonctions analytiques*.

Dans la solution que j'ai donnée de ce problème dans l'Ouvrage cité, j'ai commencé par chercher l'expression exacte du reste de la série, ensuite j'ai déterminé les limites de cette expression. Mais on peut trouver immédiatement ces limites d'une manière plus élémentaire, et également rigoureuse.

Nous allons, pour cela, établir ce principe général, qui peut être utile dans plusieurs occasions :

⁷ Autrement dit : une fonction polynôme.

Une fonction qui est nulle lorsque la variable est nulle aura nécessairement, pendant que la variable croîtra positivement, des valeurs finies et de même signe que celles de sa fonction dérivée, ou de signe opposé si la variable croît négativement, tant que les valeurs de la fonction dérivée conserveront le même signe et ne deviendront pas infinies.

Ce principe est très important dans la théorie des fonctions, parce qu'il établit une relation générale entre l'état des fonctions primitives et celui des fonctions dérivées, et qu'il sert à déterminer les limites des fonctions dont on ne connaît que les dérivées.

Nous allons le démontrer d'une manière rigoureuse.

Considérons la fonction $f(x+i)$, dont le développement général est

$$f(x) + i f'(x) + \frac{i^2}{2!} f''(x) + \dots$$

Nous avons vu, dans la Leçon précédente, que la forme du développement peut être différente pour des valeurs particulières de x ; mais que, tant que $f''(x)$ ne sera pas infinie, les deux premiers termes de ce développement seront exacts, et que les autres contiendront par conséquent des puissances de i plus hautes que la première, de manière qu'on aura

$$f(x+i) = f(x) + i(f'(x) + V),$$

V étant une fonction de x et i , telle qu'elle devienne nulle lorsque $i = 0$.

Donc, puisque V devient nul lorsque i devient nul, il est clair que, en faisant croître i par degrés insensibles depuis zéro, la valeur de V croîtra aussi insensiblement depuis zéro, soit en plus ou en moins, jusqu'à un certain point, après quoi elle pourra diminuer; que par conséquent on pourra toujours donner à i une valeur telle que la valeur correspondante de V , abstraction faite du signe, soit moindre qu'une quantité donnée, et que pour les valeurs moindres de i la valeur de V soit aussi moindre.

Soit D une quantité donnée qu'on pourra prendre aussi petite qu'on voudra; on pourra donc toujours donner à i une valeur assez petite pour que la valeur de V soit renfermée entre les limites D et $-D$;

donc, puisqu'on a $f(x+i) = f(x) + i(f'(x) + V)$,

il s'ensuit que la quantité $f(x+i) - f(x)$ sera renfermée entre ces deux-ci

$$i(f'(x) \pm D).$$

Comme cette conclusion a lieu quelle que soit la valeur de x , pourvu que $f''(x)$ ne soit pas infinie, elle subsistera aussi en mettant successivement

$$x+i, x+2i, x+3i, \dots, x+(n-1)i$$

à la place de x ; de sorte qu'on pourra toujours prendre i positif et assez petit pour que les valeurs des quantités

$$f(x+i) - f(x),$$

$$f(x+2i) - f(x+i),$$

$$f(x+3i) - f(x+2i),$$

...

$$f(x+ni) - f(x+(n-1)i),$$

soient renfermées respectivement entre les limites

$$i(f'(x) \pm D),$$

$$i(f'(x+i) \pm D),$$

$$i(f'(x+2i) \pm D),$$

...

$$i(f'(x+(n-1)i) \pm D),$$

en prenant pour D la même quantité dans chacune de ces limites, ce qui est permis, pourvu qu'aucune des quantités $f'(x), f'(x+i), f'(x+2i), \dots, f'(x+(n-1)i)$

ne soit infinie.

Donc, si toutes ces dernières quantités sont de même signe, c'est-à-dire, toutes positives ou toutes négatives, il est facile d'en conclure que la somme des quantités précédentes, laquelle se réduit à $f(x+ni) - f(x)$

aura pour limite la somme des limites, c'est-à-dire les quantités

$$if'(x) + if'(x+i) + if'(x+2i) + \dots + if'[x+(n-1)i] \pm niD.$$

Si donc on prend la quantité arbitraire D moindre que la somme

$$f'(x) + f'(x+i) + f'(x+2i) + \dots + f'[x+(n-1)i]$$

divisée par n , abstraction faite du signe de cette somme, la quantité $f(x+ni) - f(x)$ sera nécessairement renfermée entre zéro et la somme

$$2i(f'(x) + f'(x+i) + f'(x+2i) + \dots + f'[x+(n-1)i]).$$

Donc, si P est la plus grande valeur positive ou négative des quantités

$$f'(x), f'(x+i), f'(x+2i), \dots, f'(x+(n-1)i),$$

la quantité $f(x+ni) - f(x)$ sera, à plus forte raison, renfermée entre zéro et $2niP$.

Or, comme, en prenant i aussi petit qu'on voudra, on peut en même temps prendre n aussi grand qu'on voudra, on pourra supposer in égale à une quantité quelconque z , positive ou négative, puisque la quantité i peut être prise positivement ou négativement.

La quantité $f(x+ni) - f(x)$ deviendra ainsi $f(x+z) - f(x)$, et pourra représenter une fonction quelconque de z , qui s'évanouit lorsque $z = 0$, la quantité x pouvant maintenant être regardée comme une constante arbitraire. De même, la quantité $f'(x+ni)$ deviendra $f'(x+z)$, et représentera la fonction dérivée de la même fonction de z , puisque $f'(x+z)$ est également la fonction dérivée de $f(x+z)$, soit par rapport à x , soit par rapport à z .

On peut donc conclure en général que, si $f'(x+z)$ a constamment des valeurs finies et de même signe, depuis $z = 0$, et que P soit la plus grande de ces valeurs, abstraction faite du signe, la fonction primitive dont il s'agit sera renfermée entre 0 et $2zP$; par conséquent elle aura toujours aussi des valeurs finies, et de même signe que la fonction dérivée si z est positive, ou de signe différent si z est négative.

(Lagrange, 1884, p.86-89)

On voit dans ce passage qu'on aurait tort de voir en Lagrange le promoteur d'une analyse reposant sur du pur calcul formel. Certes, dans la préface de la Théorie des fonctions analytiques (dans le sous-titre de l'ouvrage, même), il rejetait la notion de limite comme fondement et point de départ d'un exposé systématique de l'analyse mathématique. On le sait, il définit la dérivée f' d'une fonction f comme le coefficient du facteur i dans le développement en série

$$f(x+i) = f(x) + i f'(x) + \frac{i^2}{2!} f''(x) + \frac{i^3}{3!} f'''(x) + \dots$$

On voit cependant qu'il reconnaît une place légitime aux aspects numériques, aspects dans lesquels les questions de convergence et de majoration d'erreur sont centrales. Le théorème ici démontré servira à établir un résultat de majoration, et est lui-même établi par un raisonnement numérique. Plus encore, on voit dans la démonstration que quoiqu'il ne souhaite pas donner de la dérivée une définition reposant sur la notion de limite, il utilise librement cet aspect de la dérivée. Bien que la notion de limite n'ait pas explicitement été définie, on voit Lagrange expliciter la notion de limite nulle en zéro au moyen d'une implication entre inégalités : V étant fonction de i , l'hypothèse « V devient nulle lorsque i devient nul » est reformulé en « on pourra toujours donner à i une valeur telle que la valeur correspondante de V , abstraction faite du signe, soit moindre qu'une quantité donnée, et que pour les valeurs moindres de i la valeur de V soit aussi moindre » ; la phrase qui suit est encore plus explicite. Ces remarques ne visent pas à défendre l'idée selon laquelle, à quelque variantes de formulation près, le système de l'analyse selon Lagrange est le même que le nôtre ; ne serait-ce que dans le paragraphe que nous commentons, on voit que l'explicitation est amenée par un raisonnement qualitatif qui évoque plutôt la continuité en zéro. Au contraire, la tâche de comprendre en quoi l'analyse selon Lagrange diffère de la nôtre nous semble un chantier de recherche historique d'autant plus intéressant qu'il invite à travailler, au-delà de l'histoire des concepts proprement mathématiques (comme celui de fonction), sur l'historicité des catégories épistémologiques de démonstration, de définition ou de cas général (Ferraro et Panza, 2012) (Chorlay, à paraître).

Tel n'est toutefois pas le propos du présent dossier. Vraisemblablement, le lecteur du 21ème siècle éprouve une certaine gêne devant cette démonstration ; une gêne qui n'est pas réductible à la difficulté qu'il y a à lire une preuve assez longue et rédigée dans un style qui n'est plus le nôtre. Certes le théorème énoncé est juste, et l'idée d'utiliser les nombres dérivés pour obtenir des approximations affines locales semble excellente. Sans doute quelque chose dans l'argument de sommation puis de passage à la limite éveille la suspicion : trop de variables dépendent les unes des autres pour qu'on puisse faire tendre les unes vers zéro sans toucher aux autres.

Si avant même toute analyse plus précise le lecteur du 21ème siècle ressent un sentiment d'incrédulité devant ce passage, c'est qu'il a été formé à se défier de ce genre de raisonnements ; nous verrons au §6 une partie des origines historiques de cette éducation à la défiance.

3. L'énoncé et la preuve de Cauchy (1823)

Problème. La fonction $y = f(x)$ étant supposée continue par rapport à x dans le voisinage de la valeur particulière $x = x_0$, on demande si, à partir de cette valeur, la fonction croît ou diminue, tandis que l'on fait croître ou diminuer la variable elle-même.

Solution. Soient Δx , Δy les accroissements infiniment petits et simultanés des variables x et y . Le rapport $\Delta y/\Delta x$ aura pour limite $dy/dx = y'$. On doit en conclure que, pour de très petites valeurs numériques de Δx et pour une valeur particulière x_0 de la variable x , le rapport $\Delta y/\Delta x$ sera positif si la valeur correspondante de y' est une quantité positive et finie, négatif si cette valeur de y' est une quantité finie, mais négative. Dans le premier cas, les différences infiniment petites Δx , Δy étant de même signe, la fonction croîtra ou diminuera, à partir de $x = x_0$, en même temps que la variable x . Dans le second cas, les différences infiniment petites étant de signe contraire, la fonction croîtra si la variable x diminue, et décroîtra si la variable augmente.

Ces principes étant admis, concevons que la fonction $y = f(x)$ demeure continue entre deux limites données $x = x_0$, $x = X$. Si l'on fait croître la variable x par degrés insensibles depuis la première limite jusqu'à la seconde la fonction y ira en croissant toutes les fois que sa dérivée étant finie aura une valeur positive (...).

(Cauchy, 1823, p.37)

A la différence de Lagrange, Augustin-Louis Cauchy (1789-1857) avait choisi dans ses cours de *définir* la notion de dérivé au moyen de celle de limite. Comme Lagrange, il tire des conséquences correctes de cet aspect ; en l'occurrence : si une fonction de x (ici le taux de variation $\Delta y/\Delta x$) a en $x = x_0$ une limite strictement positive, alors elle est positive sur un voisinage de x_0 . Il obtient ainsi un résultat local en un chaque point.

La stratégie de démonstration diffère ensuite totalement de celle de Lagrange. Ce dernier passait du local au global en mettant bout à bout entre 0 et une valeur donnée z des approximations affines locales, puis en cherchant à contrôler l'erreur pour montrer que le signe de la valeur approchée obtenue en z était le même que celui de la fonction en z . L'approximation affine locale ne joue aucun rôle dans le raisonnement de Cauchy : il interprète le résultat sur le signe de $\Delta y/\Delta x$ au voisinage d'un point comme un résultat de locale croissance (ou décroissance)⁸, et passe au global par un argument assez qualitatif mais plutôt convaincant. On pourrait le rendre explicite et rigoureux en utilisant des arguments d'ailleurs utilisés par Cauchy en d'autres endroits, par exemple dans la démonstration du théorème des valeurs intermédiaires donnée en annexe de l'*Analyse algébrique* (Cauchy, 1821) : supposons la fonction de dérivée strictement positive sur un segment (a,b) , mais telle que $f(b)$ soit inférieur à $f(a)$. On localise la situation par intervalles emboîtés : partageons le segment (a,b) en 10 parties égales, il existe deux valeurs successives telles que l'image de la plus petite est supérieure à celle

8 Signalons qu'au sens strict cette interprétation est erronée. La stricte positivité de $\Delta y/\Delta x$ en un unique point x_0 ne garantit pas la croissance au voisinage de x_0 . Un exemple construit à partir d'une fonction de la famille $x^n \sin \frac{1}{x}$ permet de l'illustrer : la fonction définie au voisinage de 0 (privé de 0) par la formule $x + 10x^2 \sin \frac{1}{x}$, prolongée par 0 en 0, est dérivable en 0, de dérivée strictement positive en zéro (égale à 1), mais croissante sur aucun voisinage de 0 (on le voit en calculant f' sur \mathbf{R}^*).

de la plus grande ; on repartage ce nouveau segment en 10 parties égales etc. Au point commun à tous les segments emboîtés on obtient une contradiction avec la stricte positivité du nombre dérivé. Bien entendu cet argument suppose la complétude de \mathbf{R} , toujours implicite chez Cauchy.

Signalons que le type de démonstration par aboutage d'approximations affines locales puis étude de la limite lorsque le pas de discrétisation tend vers zéro est aussi utilisée parfois par Cauchy. On le voit dans le troisième volume de son *Cours d'analyse à l'école polytechnique*, dans lequel il entreprend d'établir que pour une équation différentielle ordinaire du premier ordre $y' = f(x, y)$, et pour des conditions initiales convenables (conditions dite de Cauchy), les fonctions affines par morceaux données par la méthode d'Euler convergent vers une fonction limite, solution de l'équation (Cauchy, 1981). Pour le lecteur du 21^{ème} siècle, cette preuve présente le même type de défaut que celle de Lagrange donnée plus haut : Cauchy y suppose des indépendances entre majorations de grandeurs qui reviennent à supposer que toute continuité est uniforme, que toute fonction dérivable est uniformément dérivable, que toute convergence de suite de fonctions est uniforme.

4. La preuve d'Ossian Bonnet dans le manuel de J.-A. Serret (1868)

Un élément de la chaîne déductive finale apparaît très explicitement dans le *Cours de calcul différentiel et intégral* de Joseph Serret (1819-1885), à savoir l'égalité des accroissements finis.

THÉORÈME I. - Soit $f(x)$ une fonction de x qui reste continue pour les valeurs de x comprises entre des limites données, et qui, pour ces valeurs, ait une dérivée $f'(x)$ déterminée. Si x_0 et X désignent deux valeurs de x comprises entre les mêmes limites on aura

$$\frac{f(X) - f(x_0)}{X - x_0} = f'(x_1),$$

x_1 étant une valeur comprise entre x_0 et X .

En effet, le rapport $\frac{f(X) - f(x_0)}{X - x_0}$

a, par hypothèse, une valeur finie; et, si l'on nomme A cette valeur, on aura

$$(1) (f(X) - AX) - (f(x_0) - Ax_0) = 0.$$

Désignons par $\varphi(x)$ la fonction de x définie par la formule (2)

$$\varphi(x) = (f(x) - Ax) - (f(x_0) - Ax_0)$$

on aura, à cause de l'égalité (1),

$$\varphi(x_0) = 0, \quad \varphi(X) = 0,$$

en sorte que $\varphi(x)$ s'annule pour $x = x_0$ et pour $x = X$.

Supposons, pour fixer les idées, $X > x_0$ et faisons croître x de x_0 à X ; la fonction $\varphi(x)$ est d'abord nulle. Si l'on admet qu'elle ne soit pas constamment nulle, pour les valeurs de x comprises entre x_0 et X , il faudra qu'elle commence à croître en prenant des valeurs positives, ou à décroître en prenant des valeurs négatives, soit à partir de $x = x_0$, soit à partir d'une

valeur de x comprise entre x_0 et X . Si les valeurs dont il s'agit sont positives, comme $\varphi(x)$ est continue et qu'elle doit s'annuler pour $x = X$, il est évident qu'il y aura une valeur x_1 entre x_0 et X telle, que $\varphi(x_1)$ sera supérieure ou au moins égale aux valeurs voisines

$$\varphi(x_1 - h) \quad , \quad \varphi(x_1 + h) \quad ,$$

h étant une quantité aussi petite que l'on voudra. Si la fonction $\varphi(x)$, en cessant d'être nulle, prend des valeurs négatives, le même raisonnement prouve qu'il existe une valeur x_1 entre x_0 et X telle, que $\varphi(x_1 - h)$, $\varphi(x_1 + h)$,

sera inférieure ou au plus égale aux valeurs voisines.

Ainsi, dans l'un et l'autre cas, la valeur de x_1 sera telle, que les différences

$$\varphi(x_1 - h) - \varphi(x_1) \quad , \quad \varphi(x_1 + h) - \varphi(x_1)$$

seront de même signe, et, par suite, les rapports

$$\frac{\varphi(x_1 - h) - \varphi(x_1)}{-h} \quad \frac{\varphi(x_1 + h) - \varphi(x_1)}{h}$$

seront de signes contraires.

Il faut remarquer que nous n'excluons pas l'hypothèse dans laquelle l'un des rapports, précédents se réduit à zéro, ce qui exige que la fonction $\varphi(x)$ conserve la même valeur pour les valeurs de x comprises dans un intervalle fini. En particulier, si la fonction $\varphi(x)$ est constamment nulle pour les valeurs de x comprises entre x_0 et X , les rapports (3) sont nuls l'un et l'autre.

Les rapports (3) tendent vers la même limite quand h tend vers zéro, car nous admettons que la fonction $f(x)$ a une dérivée déterminée, et la même chose a lieu, en conséquence, à l'égard de $\varphi(x)$; d'ailleurs ces rapports sont de signes contraires: donc leur limite est zéro.

Ainsi l'on a $\lim \frac{\varphi(x_1 + h) - \varphi(x_1)}{h} = 0$,

ou, à cause de l'équation (2),

$$\lim \left(\frac{f(x_1 + h) - f(x_1)}{h} - A \right) = 0 \quad ,$$

C'est-à-dire

$$A = \lim \frac{f(x_1 + h) - f(x_1)}{h} = f'(x_1) \quad .$$

On a donc

$$\frac{f(X) - f(x_0)}{X - x_0} = f'(x_1) \quad ,$$

Ou $f(X) - f(x_0) = (X - x_0) f'(x_1)$,

comme on l'avait annoncé.

Nous avons supposé $X > x_0$; mais comme la formule précédente ne change pas par la permutation des lettres x_0 , X , elle est évidemment indépendante de cette hypothèse.

(...) REMARQUE. - La démonstration qui précède est due à M. Ossian Bonnet. Il faut remarquer qu'elle ne suppose en aucune façon la continuité de la dérivée $f'(x)$; elle exige seulement que cette dérivée existe et ait une valeur déterminée.

(Serret, 1900, p.17-19)

On reconnaît l'argument utilisé aujourd'hui : une transformation affine ramène à une fonction prenant les mêmes valeurs au bord ; l'égalité est obtenue par l'annulation du nombre dérivé en un extremum (local). La rédaction diffère toutefois sur un point important de celle que l'on donnerait aujourd'hui. Ici, l'existence de l'extremum local est moins démontrée que rendue intuitive au lecteur, en ayant recours à ce que nous avons appelé ailleurs le « style narratif » : on suit temporellement l'évolution des grandeurs indépendantes et dépendantes de manière à faire sentir l'« évidence » de tel ou tel aspect qualitatif. L'argument est très parlant ; il n'est pas faux, au sens où l'on pourrait lui opposer un contre-exemple. Il est cependant d'un type qui sera progressivement considéré comme illégitime, au sens où on ne l'accepte pas comme élément de démonstration.

L'histoire de cette perte de légitimité n'a pas encore été écrite, contentons-nous de deux remarques.

Premièrement, le poids de cet argument repose sur notre incapacité à imaginer certaines situations : étant donné les hypothèses, la fonction *ne peut pas ne pas* passer par un extremum (au sens où : on ne peut l'imaginer) ; de même, une fonction continue changeant de signe ne peut pas ne pas passer par zéro etc. On imagine bien que l'exhibition de fonctions *pathologiques* va saper la légitimité de ce recours à l'intuition des courbes et des évolutions temporelles de grandeurs, quand bien même aucun contre-exemple ne viendrait invalider les énoncés à l'appui desquels ces arguments sont avancés.

Une deuxième remarque invite à nuancer la première. Du point de vue historique, il n'est pas évident que la défiance envers ce type d'argument soit issue de la rencontre avec de plus en plus de fonctions pathologiques. Comme souvent Cauchy se trouve dans l'entre-deux : certaines fonctions pathologiques importantes jouent un rôle dans ses cours des années 1820 (par exemple $x \mapsto e^{-x^2}$, infiniment dérivable en zéro mais ne coïncidant avec son développement de Taylor sur aucun voisinage de zéro), mais la distinction entre ponctuel et uniforme lui échappe. Il donne deux démonstrations du théorème des valeurs intermédiaires, l'une de style narratif, l'autre reposant sur des intervalles emboîtés (comme souvent, il fait d'un procédé d'approximation bien connu le rouage principal d'une démonstration d'existence). Son raisonnement sur le lien entre signe de f' et variations de f semble encore partiellement narratif. En remontant dans le temps, on peut faire remarquer que les fonctions pathologiques ne sont pas dans l'horizon des préoccupations de Lagrange ; on a pourtant vu ce dernier donner une démonstration numérique très technique d'un énoncé parfaitement évident du point de vue intuitif (du type : si on part de 0 et qu'à chaque instant on augmente, on a toujours 0 ou plus).

Ce « Théorème I » de Serret – d'après Bonnet (1819-1892) – est suivi de deux autres :

THÉORÈME II. - *Si la fonction $f(x)$ est constante pour toutes les valeurs de x comprises entre deux limites données, la dérivée $f'(x)$ est nulle pour les mêmes valeurs de x . Réciproquement, si la dérivée $f'(x)$ est nulle pour les valeurs de x comprises entre deux limites, la fonction $f(x)$ a une valeur constante pour les valeurs de x comprises entre les mêmes limites.*

(...)

COROLLAIRE. - *Lorsque deux fonctions $f(x)$, $F(x)$ de la variable x ne diffèrent que par une constante pour les valeurs de x comprises entre deux limites données, les dérivées de ces fonctions sont égales entre elles pour les mêmes valeurs de x . Réciproquement, si les dérivées $f'(x)$, $F'(x)$ de deux fonctions $f(x)$, $F(x)$ sont égales entre elles pour toutes les valeurs de x comprises entre des limites données, les fonctions ne diffèrent que par une constante, pour ces ,mêmes valeurs de x . (Serret, 1900, p.20)*

Puis :

THÉORÈME III. - *Si la dérivée $f'(x)$ de la fonction $f(x)$ reste finie pour toutes les valeurs de x comprises entre les limites x_0 et $X > x_0$, et que l'on fasse croître x de x_0 à X , la fonction $f(x)$ croîtra tant que la dérivée $f'(x)$ ne sera pas négative, et elle décroîtra tant que $f'(x)$ ne sera pas positive.* (Serret, 1900, p.21)

Ce qui surprend, c'est qu'alors que le théorème II est déduit de l'égalité des accroissements finis, la démonstration du théorème III ne l'utilise pas. Serret reprend ici l'argument de Cauchy :

En effet, x étant comprise entre x_0 et X , le rapport

$$\frac{f(x \pm h) - f(x)}{\pm h}$$

a pour limite $f'(x)$, qui est une quantité finie; il aura donc le signe de cette limite pour toutes les valeurs de h comprises entre zéro et une quantité positive ε suffisamment petite.

Par conséquent, on aura, pour les mêmes valeurs de h ,

$$f(x-h) < f(x) < f(x+h)$$

si $f'(x)$ est > 0 , et

$$f(x-h) > f(x) > f(x+h)$$

si $f'(x)$ est < 0 .

Ainsi la fonction $f(x)$ ira en croissant à partir de chaque valeur de x pour laquelle $f'(x)$ est > 0 , tandis qu'elle ira en décroissant à partir de chaque valeur telle que $f'(x)$ soit < 0 .

(Serret, 1900, p.21-22)

Serret n'est pas le seul à ne pas faire le lien entre ce théorème et l'égalité des accroissements finis. Ainsi, les mêmes choix démonstratifs se retrouvent dans le très weierstrassien traité de Genocchi et Peano (Genocchi et Peano, 1889). On peut aussi noter que l'égalité des accroissements finis était établie par Cauchy dans ses cours.

5. Analyse de preuve : contre-exemple local et analyse régressive

5.1 Le rôle de la convergence uniforme

Des éléments relatifs à l'histoire de la différenciation conceptuelle des notions de continuité d'une fonction et de limite d'une suite de fonctions ont été rapportés par nombreux auteurs (Dugac, 2003) (Arsac 2013) ; Stokes, Seidel, Gudermann, Weierstrass distinguent pour ces deux notions des variantes ponctuelles (continuité en chaque point, convergence simple) et uniformes. Nous donnons ici un extrait relatif à la distinction entre dérivabilité *en chaque point* et dérivabilité uniforme. Cette distinction permet de souligner qu'une hypothèse implicite d'uniforme dérivabilité se cache dans les démonstrations par aboutement d'approximations affines.

Le texte permettent en outre d'illustrer la notion de contre-exemple local. Nous utilisons ici la terminologie proposée par Lakatos (Lakatos 1976) pour décrire l'activité d'analyse de preuve (*proof-analysis*) menée par les mathématiciens eux-mêmes : un contre-exemple est dit global lorsqu'il vient infirmer l'énoncé de la proposition ; local s'il vient infirmer une étape d'une démonstration, sans toutefois infirmer la proposition qu'on entend démontrer. C'est bien entendu cette seconde variété que nous rencontrons dans notre enquête, les théorèmes visés étant valides.

Lisons la critique que fait Giuseppe Peano (1858-1932) de la démonstration de l'égalité des accroissements finis que Camille Jordan (1838-1922) donne dans la première édition de son *Cours d'analyse*. L'échange de lettres fut publié en 1887 dans les *Nouvelles annales de mathématiques* :

Dans son *Cours d'Analyse de l'Ecole Polytechnique*, p.21, M. Jordan donne une démonstration peu rigoureuse du théorème suivant :

« Soit $y = f(x)$ une fonction de x dont la dérivée reste finie et déterminée lorsque x varie dans un certain intervalle.

Soient a et $a+h$ deux valeurs de x prises dans cet intervalle. On aura $f(a+h) - f(a) = \mu h$, μ désignant une quantité intermédiaire entre la plus grande et la plus petite valeur de $f'(x)$ dans l'intervalle de a à $a+h$. »

En effet, dit l'auteur, donnons à x une série de valeurs a_1, a_2, \dots, a_{n-1} , entre a et $a+h$; posons

$$f(a_r) - f(a_{r-1}) = (a_r - a_{r-1}) [f'(a_{r-1}) + \varepsilon_r]$$

.Supposons maintenant les valeurs intermédiaires a_1, a_2, \dots, a_{n-1} indéfiniment multipliées (et rapprochées). Les quantités $\varepsilon_1, \varepsilon_2, \dots$ tendront toutes vers zéro, car ε_r est la différence entre $\frac{f(a_r) - f(a_{r-1})}{a_r - a_{r-1}}$ et sa limite $f'(a_{r-1})$.

Cette affirmation n'est pas juste ; car $f'(a_{r-1}) = \lim_{a_r \rightarrow a_{r-1}} \frac{f(a_r) - f(a_{r-1})}{a_r - a_{r-1}}$

quand on suppose a_{r-1} fixe, et a_r variable et s'approchant indéfiniment de a_{r-1} ; mais on ne le peut pas affirmer quand variant en même temps a_r et a_{r-1} , si l'on ne suppose pas que la dérivée soit continue.

Ainsi, par exemple, posons $y = f(x) = x^2 \sin \frac{1}{x}$, avec $f(0) = 0$;

sa dérivée $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$ pour $x \neq 0$, et $f'(0) = 0$, reste toujours finie et déterminée, mais discontinue.

Soit $a = 0$, $h > 0$;

posons $a_1 = \frac{1}{2n\pi}$, $a_2 = \frac{1}{(2n+1)\pi}$, a_3, a_4, \dots quelconques.

On aura $\varepsilon_2 = \frac{f(a_2) - f(a_1)}{a_2 - a_1} - f'(a_1)$;

mais $f(a_1) = 0$, $f(a_2) = 0$, $f'(a_1) = -1$;

donc $\varepsilon_2 = 1$, et sa limite n'est pas zéro.

Presque la même faute a été commise par M. Hoüel (*Cours de Calcul infinitésimal*, t.I, p.145). J'ajouterai enfin que l'on démontre très facilement la formule

$$f(x_0+h) - f(x_0) = h f'(x_0+\theta h),$$

sans supposer la continuité de la dérivée.

(Peano, 1884)

Fair-play, Jordan répond « Je n'ai rien à redire à la critique de M. le Dr Peano, qui est parfaitement fondée. J'ai admis implicitement dans ma démonstration que $\frac{f(x+h) - f(x)}{h}$ tendait *uniformément* vers $f'(x)$ dans l'intervalle de a à b » (Peano, 1884, p.47). On sait que les éditions suivantes du traité de Jordan seront entièrement refondues et donneront en France le modèle de l'analyse nouveau style. Il y démontre en particulier que si la dérivée est continue, $\frac{f(x+h) - f(x)}{h}$ tend bien uniformément vers $f'(x)$ sur les segments (Jordan, 1893, p.68). Peano avait donc besoin d'une fonction à dérivée discontinue pour construire son contre-exemple. On aurait pu citer d'autres textes au contenu très proche. Mentionnons en particulier les critiques adressées par Darboux aux versions préparatoires du traité de Hoüel (Gispert, 1983).

5.2 Analyse régressive : existence d'un maximum

La preuve de Bonnet échappe bien sûr à ce type de reproche, puisqu'elle repose sur une idée entièrement différente. On peut cependant se montrer incrédule face à l'appel à l'« évidence » d'existence d'un extremum. Certes, si la fonction est monotone par intervalle, cette existence s'établit sans problème ; mais toute fonction, même de classe C^n , n'est malheureusement pas monotone par intervalle, comme le montrent les classiques $x \mapsto x^n \sin \frac{1}{x}$.

Ceci étant, l'existence d'un extremum ne dépend ici en rien de la locale monotonie, pas plus que de la dérivabilité, comme Weierstrass le démontre, par exemple dans ses leçons

de 1878 sur la théorie des fonctions (Weierstrass, 1988). Le passage que nous donnons en traduction n'est pas *a priori* lié aux questions de calcul différentiel. Il est placé après la construction des nombres réels (plus précisément, de la droite numérique achevée $\bar{R} = (-\infty, +\infty)$) à partir des nombres rationnels. Cette construction permet d'établir la validité d'une série de résultats utilisant l'argument des intervalles emboîtés, en particulier l'existence d'une borne supérieure (*obere Grenze*) pour toute partie non vide de la droite achevée. On passe insensiblement de résultats portant sur des parties de \bar{R} à des résultats portant sur les correspondances fonctionnelles :

Qu'à un point (x_1, x_2, \dots, x_n) d'un domaine corresponde toujours un point y ; y est alors aussi une grandeur variable et possède aussi une limite inférieure et supérieure ; soit g cette dernière. Il existe [dans le domaine des x]⁹ (il n'est pas nécessaire que le point appartienne au domaine défini) au moins un point ayant la propriété suivante : si je considère un voisinage quelconque de ce point, aussi petit que je veux, si je considère les valeurs de y correspondant aux points du domaine x se trouvant dans ce voisinage, la limite supérieure de ces dernières valeurs est égale à g . Semblablement pour la limite inférieure.

(Weierstrass, 1988, p.91, traduction libre)

Ce théorème un peu inhabituel n'a pas été intégré aux exposés standards des premiers éléments de l'analyse. Le résultat est ici très général : la fonction $y = f(x_1, \dots, x_n)$ n'est pas supposée continue ; le domaine dans lequel varie le point (x_1, \dots, x_n) n'est soumis à aucune hypothèse : il n'a pas à être borné (le fait d'être dans la droite achevée en dispense), ni à être fermé. La faiblesse des hypothèses n'empêche cependant pas d'établir (par intervalles emboîtés) ce résultat de localisation des bornes supérieures et inférieures d'une fonction.

C'est dans un second temps que Weierstrass en tire le corollaire relatif aux fonctions numériques continues étudiées sur un segment :

C'est une question courante que de se demander si, parmi les valeurs que peut prendre une grandeur, se trouve un maximum ou un minimum (maximum et minimum au sens absolu). Soit y une fonction continue de x , $y = f(x)$. x doit se trouver entre deux limites fixées a et b . Dans quelles circonstances y a-t-il pour y un maximum et un minimum ? Il existe pour y une borne supérieure. Il doit donc y avoir, d'après notre proposition, un point x_0 du domaine x tel que la borne supérieure de y entre $x_0 - \delta$ et $x_0 + \delta$ est aussi g . x_0 se trouve soit à l'intérieur de $a \dots b$ ou sur la frontière ($x_0 = a$ ou $x_0 = b$).

Dans le premier cas $f(x_0)$ est un maximum. En effet $f(x_0)$ doit être égal à g : on peut rendre $f(x) - f(x_0)$ aussi petit qu'on veut en choisissant $|x - x_0|$ suffisamment petit, et d'autre part, puisque x est dans l'intervalle $x_0 - \delta \dots x_0 + \delta$, $f(x)$ peut devenir aussi près qu'on veut de g , d'où $f(x_0) = g$. (Si on avait $f(x_0) = g + h$, alors $f(x) - f(x_0) = f(x) - g - h$, et $f(x)$ ne pourrait s'approcher d'aussi près qu'on veut de g si h n'était pas nul.)

Si x_0 coïncide avec a ou b , on ne peut plus considérer de a (resp. b) que $f(a)$ (resp. $f(b)$) est un maximum, que si $f(x)$ est varié aussi continûment en a (resp. b).

(Weierstrass, 1988, p.91-92, traduction libre)

⁹ C'est nous qui ajoutons la précision entre crochets.

Des preuves en tout point semblables avaient été publiées avant 1878, mais toutes dérivait des cours de Weierstrass : (Cantor 1871), (Heine 1872, p.186), (Darboux 1872)¹⁰. Quelques cinquante années plus tard, cette clarification des conditions d'existence d'extrema est toujours considérée comme l'un des succès obtenus par Weierstrass dans son travail de fondement de l'analyse, comme le montrent les premières lignes du texte de Hilbert sur l'infini¹¹ :

Par sa critique d'une magistrale pénétration, Weierstrass a établi un fondement solide pour l'analyse mathématique. En clarifiant – entre autres – les notions de minimum, de fonction, de quotient différentiel, il a libéré le calcul infinitésimal des imperfections qui y demeuraient ; il en a expurgé toute les considérations confuses relatives à l'infinitésimal, et a ainsi définitivement triomphé des difficultés découlant de cette notion.

(Hilbert, 1926, p.161, traduction libre)

Dans la chaîne déductive longue telle qu'elle se forme à la fin du 19^{ème} siècle, on observe dans la démonstration de l'égalité des accroissements finis le remplacement de l'argument narratif d'existence d'un extremum local par le renvoi à un théorème d'existence d'extremum global pour les fonctions continues sur un segment, lui-même démontré en utilisant explicitement l'une des propriétés équivalente à la complétude de \mathbf{R} . Cette substitution ne modifie pas l'architecture de la preuve ; elle n'est pas rendue nécessaire par l'exhibition d'un contre-exemple local invalidant une étape du raisonnement. On pourrait ici parler d'une analyse *régressive*, au sens où un élément qui était considéré comme premier, indémontrable ou ne nécessitant pas de démonstration (ici : l'existence d'un extremum local pour une fonction continue non constante s'annulant aux deux bornes d'un intervalle) se voit désormais déduit de quelque chose de plus primitif encore (les propriétés de la droite réelle).

6. Pistes

Nous espérons que le parcours de ce dossier documentaire a pu satisfaire – et aiguïser – la curiosité du lecteur concernant des théorèmes centraux au Lycée et en L1. Nous quittons à présent le dossier de textes historiques et sa mise en forme épistémologique pour nous demander ce que pourraient être des enseignements s'appuyant – entre autres – sur lui. Nous évoquerons deux pistes d'utilisation possible, une au niveau L1, puis au sein des Masters préparant aux métiers de l'enseignement. Il ne s'agit pas ici du bilan d'un travail empirique, mais de pistes de travail, aussi bien au niveau enseignement qu'au niveau recherche. L'évocation de ces pistes constitue en outre un appel à collaboration.

6.1 Accompagner l'entrée dans l'Analyse post-bac

De nombreux travaux ont permis de mieux comprendre les enjeux et difficultés de la transition entre le secondaire et le supérieur, soit pour le domaine spécifique de l'analyse, soit relativement aux tâches de compréhension et de rédaction de

¹⁰ On trouvera plus d'éléments, dont des traductions, dans (Chorlay 2007a, chapitre 4).

¹¹ Nous donnons cette citation dans le seul but d'illustrer ce point. Cet extrait, très rhétorique, ne saurait être regardé comme une analyse du rôle de Weierstrass dans l'histoire du fondement de l'analyse ; tel n'était pas le propos de Hilbert.

démonstrations¹² ; quoique peu pratiqué, le travail s'appuyant sur des textes historiques a régulièrement été envisagé (Robert et Schwartzberger, 1991, §2.3).

Pris dans son ensemble, le présent dossier documentaire nous semble bien trop copieux, et souvent trop ardu, pour servir d'appui à un travail en L1. Cependant, plusieurs éléments peuvent sans doute être repris, aussi bien pour travailler sur les notions mathématiques (limite, dérivation etc.) que pour travailler sur l'écriture des mathématiques. Donnons quelques exemples (non encore testés avec des étudiants) :

- La formulation du théorème par Lagrange (*Une fonction qui est nulle lorsque la variable est nulle aura nécessairement, pendant que la variable croîtra positivement, des valeurs finies et de même signe que celles de sa fonction dérivée, ou de signe opposé si la variable croît négativement, tant que les valeurs de la fonction dérivée conserveront le même signe et ne deviendront pas infinies*) est très éloignée de la formulation bien connue de tout élève sortant du Lycée ; ce serait sans doute une tâche non triviale à confier à des étudiants que de leur demander si l'énoncé de Lagrange leur évoque un théorème bien connu ; si oui, lui est-il équivalent ou en est-il une conséquence ? La comparaison des deux formulations permet aussi de revenir sur un point souvent oublié des étudiants : l'implication (signe de $f' \rightarrow$ sens de variations de f) n'est valide que si le domaine est un intervalle¹³ ; Lagrange l'exprime autrement, en limitant la validité du théorème aux domaines où ni la fonction ni sa dérivée ne devienne infinies. On pourrait demander aux étudiants d'illustrer le rôle de cette hypothèse de Lagrange, puis de chercher des exemples de domaines mettant en défaut la conclusion, sans toutefois que ni f ni f' ne devienne infinie¹⁴.
- La démonstration que Cauchy donne du théorème peut donner lieu à un débat scientifique partant d'une question très ouverte, du type « Ce texte vous semble-t-il démontrer le théorème cité ; en quoi ? ». Un traitement plus guidé peut amener à travailler sur les points suivants : reconnaissance de la notion de nombre dérivé derrière une formulation inhabituelle aux étudiants ; retour à la définition de la notion de croissance pour étudier les questions suivantes : la positivité du nombre dérivé en un point implique-t-elle la croissance sur un voisinage de ce point ? Si la fonction dérivée est positive sur l'intervalle $(0,1)$ l'argument de Cauchy permet-il d'affirmer que $f(0) \leq f(1)$? Ces questions conduisent à problématiser les notions de comportement local, et de passage du local au global
- La démonstration de Bonnet, donnée par Serret, est à peu près celle que nous donnons aujourd'hui. On peut demander aux étudiants de chercher si le texte de Serret leur semble totalement démonstratif ou si certains passages, quoique convaincants, semblent s'appuyer sur l'intuition ; on peut faire comparer le texte de Serret et une démonstration actuelle ; on peut demander de réécrire la preuve de Serret en isolant sous forme de lemmes les propositions intermédiaires que

12 Nous renvoyons, par exemple, aux bibliographies de (Praslon 2000) et (Barrier 2009).

13 Ainsi la fonction inverse est-elle dérivable et de dérivée strictement négative, elle n'est cependant pas décroissante sur \mathbf{R}^* .

14 L'exemple de la fonction inverse n'est plus ici utilisable, il faut bricoler un peu plus ; avec par exemple la fonction définie sur \mathbf{R}^* par $x \rightarrow x$ sur \mathbf{R}^{*-} et $x \rightarrow x-1$ sur \mathbf{R}^{*+} .

Serret démontre ou suppose (implicitement) en cours de démonstration : lemme de Rolle, théorème d'existence d'un maximum.

Ces quelques pistes montrent, nous l'espérons, qu'un travail sur des extraits de textes historiques peut permettre, dans un contexte familier des étudiants (théorème bien connu au Lycée), d'aider à l'entrée progressive dans le monde de l'Analyse post-bac, avec ses longues chaînes démonstratives, ses théorèmes aux très nombreuses hypothèses, son goût des contre-exemples bizarres, sa nécessité de retourner à la définition des concepts (ou de chercher une définition) quand bien même la notion semble intuitive¹⁵, ses distinctions conceptuelles byzantines, ses démonstrations contre-intuitives de faits qualitatifs intuitifs etc.

6.2 Nourrir la réflexion des (futurs) enseignants sur la nature des mathématiques

Un autre contexte d'enseignement nous semble fournir l'occasion d'utiliser le dossier dans son ensemble. Nous pensons aux masters préparant à l'enseignement des mathématiques dans le secondaire, dans la plupart desquels un enseignement d'histoire et/ou d'épistémologie des mathématiques a été intégré. Il semble peu douteux qu'il ne soit bon que les enseignants aient des connaissances en histoire des mathématiques ; c'est plutôt sur l'aspect *épistémologique* que nous voudrions avancer ici quelques éléments, plus précisément sur l'aspect relatif à la *nature* des mathématiques.

Sans chercher à définir ici le terme « épistémologie », on peut d'emblée noter que la liste des questions épistémologiques qui pourraient utilement être abordée en formation d'enseignants est vaste et diverse¹⁶.

A titre d'inventaire non-exhaustif et semi-structuré, on peut citer quelques questions :

- (1) Du côté de l'épistémologie de branches des mathématiques ou de concepts fondamentaux : qu'est-ce que la géométrie (est-elle une connaissance empirique ? en quel sens peut-il y avoir plusieurs géométries ?) ; épistémologie de notions problématiques : nombre négatifs ou imaginaires, probabilité et espérance, le continu, l'infini (les infinis) ... ;
- (2) Sur la nature des mathématiques : fonctions de la preuve, classification de types de preuve, caractère explicatif ou non des preuves ; en quel sens les mathématiques ont-elles une histoire (si, à la différence de la physique, la théorie suivante ne remplace pas la précédente, le développement historique des mathématiques n'est-il qu'une accumulation sans discontinuité ?) ; le formalisme mathématique (son rôle, son histoire, peut-on rendre les mathématiques entièrement formelles et que peut-on apprendre de cette question (ouverture sur la logique mathématique)) ; rôle des définitions ;
- (3) Sur l'utilité des mathématiques : qui utilise des mathématiques et lesquelles ? ... aujourd'hui, il y a 400 ans (très peu de technologie ou de physique reposant sur les mathématiques, pas de statistiques, pas d'usage des probabilités par les banques et assurances ...), il y a 4000 ans à Babylone (échanges commerciaux, comptabilité et impôts, astronomie, culture professionnelle d'enseignants ...).

Nous souhaitons ici aborder également un point qui nous semble peu travaillé dans la communauté des formateurs d'enseignants, celui de la représentation que les

15 Et la construction de cette intuition était un objectif fondamental du Lycée.

16 Un exemple d'enseignement centré sur des questions épistémologiques est donné par Lombardi (2011).

professeurs se font de leur propre discipline¹⁷. Il s'agit moins là d'une question que d'une série de questions, en particulier : comment décrire l'image qu'un praticien des mathématiques (par exemple un enseignant) se fait de sa discipline ? Comment cette image se forme-t-elle et comment évolue-t-elle au cours d'une carrière ? Quelles interactions entre cette image des mathématiques, la perception de son rôle d'enseignant et ses choix d'enseignement ? Certains aspects relatifs à la nature des mathématiques devraient-ils faire l'objet d'un enseignement ou de formations, si oui lesquels ? Un enseignement d'histoire des mathématiques est-il un cadre privilégié pour accueillir une telle formation ?

On doit souligner que ces questions forment depuis longtemps un chantier important et bien circonscrit en didactique des sciences expérimentales, où les travaux relatifs à la « *nature of science* » (acronyme : NOS) s'appuient sur la philosophie, l'histoire et la sociologie des sciences pour étudier, outre les questions dont nous avons fait la liste, celles relatives à l'image de la science induite par l'enseignement chez les élèves de tous niveaux scolaires. On peut se faire une idée de ce chantier, du moins dans sa version anglo-saxonne, en consultant l'article récent *Teaching With and About Nature of Science* (Abd-El-Khalick, 2012), ou (Abd-El-Khalick et Lederman, 2000) pour un article portant plus précisément sur la formation des enseignants et le lien avec l'histoire ; le chantier n'est pas exclusivement anglo-saxon, comme le montrent par exemple les travaux de Laurence Maurines (Beaufils et Maurines, 2013). Indiquons quelques grandes lignes de l'approche anglo-saxonne. S'appuyant beaucoup sur Popper et Kuhn, à l'occasion sur Bachelard ou Duhem, on y cherche à sensibiliser élèves et professeurs aux aspects suivants des sciences expérimentales : toute observation ou toute expérience est guidée par un cadre théorique (*theory-ladenness of observation*) ; la théorie est sous-déterminée par l'expérience, et les *expériences cruciales* censées départager deux théories concurrentes sont souvent reconstruites *a posteriori* à des fins didactiques ou polémiques ; le savoir scientifique est produit collectivement, par des collectifs dont les membres peuvent exercer des fonctions diverses (variété des *faiseurs* de sciences, chercheurs, artisans, ingénieurs, amateurs ...). On voit qu'on cherche ainsi à corriger ce qu'on estime être une vision dogmatique de la science, en invitant, par une réflexion sur l'*activité* scientifique, à se défaire d'une image du type : lois générales et éternelles obtenues par démarche purement inductive, par un génie isolé observant la nature en elle-même, d'un œil enfin dégagé de théories pré-scientifiques ou de naïvetés enfantines.

Nous ne pouvons pas ne pas voir des échos entre ces travaux et le présent dossier documentaire. Certes les différences sont nombreuses, ne serait-ce que parce que l'observation, l'expérience et l'induction ne jouent pas des rôles de même importance dans les sciences expérimentales et dans les mathématiques. Certes on peut se sentir perplexe devant ce qui peut sembler une liste de *connaissances* relatives à la nature de la science, et mal à l'aise à l'idée d'*enseigner* ce qu'est la nature des mathématiques, ne serait-ce que parce que nous serions bien en peine de le dire !

Il ne s'agit toutefois pas de cela, mais plutôt d'inviter les futurs enseignants à réfléchir sur les mathématiques en les regardant comme le fruit d'une activité humaine et non seulement comme un ensemble d'énoncés entretenant des liens déductifs. Le dossier

¹⁷ Nous sommes conscients qu'il est fort possible que cette affirmation traduise bien plus notre ignorance que la rareté des travaux. Mentionnons cependant le travail de Uffe Jankvist, par exemple (Jankvist et Kjeldsen, 2011).

documentaire présenté ici a été délibérément mis en forme de manière à faire ressortir les différents *moments* d'élaboration collective d'une chaîne déductive longue comme des moments d'action rationnelle ; des moments de natures différentes : analyse de preuve, recherche de contre-exemples locaux, analyse régressive, différenciation conceptuelle.

Quand bien même cette contribution au chantier semblerait modeste, nous faisons l'hypothèse que la mise à disposition des futurs enseignants d'un *lexique épistémologique* comprenant ces termes – et d'autres à identifier – fournit des outils pour décrire l'activité mathématique ; sans constituer directement des *connaissances* relatives à la nature des mathématiques, l'usage en situation d'un tel lexique nous semble permettre d'enrichir la compréhension des mathématiques, tout en renforçant les nécessaires connaissances mathématiques et historiques relatives à cette discipline.

Références

- ABD-EL-KHALICK F. et LEDERMAN G. (2000) The Influence of History Science Courses on Students' Views of Nature of Science. *Journal of Research in Science Teaching* **37**, 1057-1095.
- ABD-EL-KHALICK F. (2013) Teaching *With* and *About* Nature of Science, and Science Teacher Knowledge Domains. *Science & Education* **22**, 2087-2107.
- AMPERE A.-M. (1806) Recherche sur quelques points de la théorie des fonctions dérivées qui conduisent à une nouvelle démonstration de la série de Taylor, et à l'expression des termes qu'on néglige lorsqu'on arrête cette série à un terme quelconque. *Journal de l'Ecole Polytechnique* **13**, 148-181.
- ARSAC G. (2013) *Cauchy, Abel, Seidel et la convergence uniforme : de la difficulté historique des raisonnements sur les limites*. Paris : Hermann.
- BARRIER T. (2009) *Une perspective sémantique et dialogique sur l'activité de validation en mathématiques* (thèse dirigée par V. Durand-Guerrier). Lyon : Université Claude Bernard.
- BEAUFILS D. et MAURINES L. (2013) Teaching the Nature of Science in Physics Courses : The Contribution of Classroom Historical Inquiries. *Science & Education* **22**, 1443-1465.
- CANTOR G. (1871) Notiz zu dem Aufsatz: Beweis (...). *Journal für die reine und angewandte Mathematik* **73**, 84-86.
- CAUCHY A.-L. (1821) *Cours d'analyse de l'Ecole royale Polytechnique, 1^{ère} partie Analyse Algébrique*. Paris : chez Debure frères = Cauchy, A.-L., *Œuvres Complètes*, série 2, tome III. Paris : Gauthiers-Villars, 1882-1974.
- CAUCHY A.-L. (1823) *Résumé des Leçons données à l'Ecole royale Polytechnique sur le Calcul infinitésimal* = Cauchy, A.-L., *Œuvres Complètes*, série 2, tome IV. Paris : Gauthiers-Villars, 1882-1974.
- CAUCHY A.-L. (1981) *Equations différentielles ordinaires – cours inédits* (édition scientifique C. Gillain). Saint-Laurent: Etudes vivantes.
- CHORLAY R. (2007a) *L'émergence du couple local-global dans les théories géométriques, de Bernhard Riemann à la théorie des faisceaux (1851-1953)* (thèse dirigée par C. Houzel). Paris : Université Paris Diderot.

- CHORLAY R. (2007b) La multiplicité des points de vue en Analyse élémentaire comme construit historique, in *Histoire et enseignement des mathématiques : erreurs, rigueurs, raisonnements*, E. Barbin & D. Bénard (eds.). Lyon: INRP, 203-227.
- CHORLAY R. (à paraître) Questions of Generality as Probes into Nineteenth Century Mathematics, in K. Chemla, R. Chorlay et D. Rabouin (eds.) *Handbook on Generality in Mathematics and the Sciences*.
- DARBOUX G. (1872) Sur un théorème relatif à la continuité des fonctions. *Bulletin des sciences mathématiques et astronomiques* **3**, 307-313.
- DOWNS M. et MAMONA-DOWNS J. (2010) Necessary Realignment from Mental Argumentation to Proof Presentation, in *Proceedings of CERME 6* (Congress of the European Society for Research in Mathematics Education). Lyon: INRP édition électronique.
- DUGAC P. (1979) *Histoire du théorème des accroissements finis* (polycopié). Paris : Université Pierre et Marie Curie.
- DUGAC P. (2003) *Histoire de l'analyse*. Paris: Vuibert.
- FERRARO G. et PANZA M. (2012) Lagrange's theory of analytical functions and his ideal of purity of method. *Archive for History of Exact Sciences* **66**(2), 95-197.
- GENOCCHI A. et PEANO G. (1889) *Differentialrechnung und Grundzüge der Integralrechnung* (traduction allemande de G. Bohlmann et A. Schepp, depuis l'édition italienne de 1884)/ Leipzig: Teubner.
- GISPERT H. (1983) Sur les fondements de l'analyse en France. *Archive for History of Exact Sciences* **28**, 37-106.
- HEINE E. (1872) Die Elemente der Funktionenlehre. *Journal für die reine und angewandte Mathematik* **74**, 172-188.
- HILBERT D. (1926) Über das Unendliche. *Mathematische Annalen* **95**, 161-190.
- JANKVIST U. et KJELDSSEN T. (2011) New Avenues for History in Mathematics Education: Mathematical Competencies and Anchoring. *Science & Education* **20**, 831-862.
- JORDAN C. (1893) *Cours d'analyse de l'école Polytechnique*, tome premier (2^{ème} édition, entièrement refondue). Paris: Gauthier-Villars.
- LAGRANGE J.-L. (1884) *Leçon sur le calcul des fonctions* (réimpression de la seconde édition (1806)), in J. Serret (ed.) *Oeuvres Complètes de Lagrange* **10**. Paris: Gauthier-Villars.
- LAKATOS I. (1976) *Proofs and Refutations*. Cambridge : Cambridge University Press.
- LOMBARDI H. (2011) *Epistémologie mathématique*. Paris: Ellipses.
- OSGOOD W. (1912) *Lehrbuch der Funktionentheorie* (2^{te} Auflage). Leipzig: Teubner.
- OUVRIER-BUFFET C. (2011) A Mathematical Experience Involving Defining Processes: In-Action Definitions and Zero-Definitions. *Educational Studies in Mathematics* **76**, 165-182.
- PEANO G. (1884) Extrait d'une lettre de M. le Dr. J. Peano. *Nouvelles annales de mathématiques* **3** (3^{ème} série), 45-49.
- PRASLON F. (1994) *Analyse de l'aspect Méta dans un enseignement de Deug A concernant le concept de dérivée. Etude des effets sur l'apprentissage* (mémoire de D.E.A., dirigé par Marc Rogalski). Paris: Université Paris Diderot.

- PRASLON F. (2000) *Continuités et ruptures dans la transition terminale / Deug sciences en analyse* (thèse dirigée par M. Artigue). Paris : Université Paris Diderot.
- ROBERT A. et SCHWARTZENBERGER R. (1991) Research in Teaching and Learning Mathematics at an Advanced Level, in D. Tall (ed.), *Advanced Mathematical Thinking*, Boston: Kluwer Academic Press, 127-139.
- SERRET J.-A. (1900) *Cours de calcul différentiel et intégral* (5^{ème} édition). Paris: Gauthier-Villars.
- STOLZ O. (1893) *Grundzüge der Differential- und Integralrechnung* (vol.1). Leipzig: Teubner.
- VOSS A. (1899) “Differential- und Integralrechnung”, in *Enzyklopädie der mathematischen Wissenschaften II.1.1*. Leipzig: Teubner.
- WEIERSTRASS K. (1888) *Einleitung in die Theorie der analytischen Funktionen, Vorlesung, Berlin 1878* (P. Ulrich, éditeur scientifique). Braunschweig: D.M.-V. & Vieweg.