À propos du théorème des accroissements finis

Introduction

Sans doute avez-vous un cousin qui ne fait pas de mathématiques mais qui conduit une voiture. La prochaine fois que vous le voyez demandez lui à brûle pourpoint :

> «Suppose que tu roules pendant une heure entre 30 et 40 km /h, combien penses-tu que tu auras fait de kilomètres ? ». (1)

La réponse sera sans doute

> «Eh, tu me prends pour un imbécile ?»

ou bien

> «Je ne vois pas ce qu'il y a de drole.»

Dans le premier cas, faites vous tout petit (il serait absurde de se fâcher avec son cousin pour si peu).
Dans le deuxième cas, vous pouvez enchaîner

> «Ce qu'il y a de drole, c'est qu'on le démontre en mathématiques, qu'on a fait entre 30 et 40 kilomètres».

Et vous aurez peut-être fait faire un grand progrès au développement de l'esprit mathématique dans toutes les couches de la société.
Si vous indiquez en outre qu'il s'agit d'un théorème assez difficile à démontrer, vous aurez sans doute créé un nouvel adepte des démonstrations rigoureuses. Si vous êtes enseignant en mathématiques([1]) en terminales vous aurez la joie d'avoir à démontrer l'évidence (1) ci-dessus au moyen de l'évidence (2) suivante:

> «Suppose que tu n'avances à aucun moment, est-ce que tu penses que tu as une chance d'avoir reculé ? » (2).

Bref, vous l'avez compris, il s'agit de faire quitter à vos élèves la planète terre pour les emmener dans la planète "mathématiques formelles".
Et si vos élèves n'y pigent que couic, et ne recherchent que des recettes, demandez-vous si par hasard, la nécessité de démonstrations convaincantes ne devrait pas être illustrée sur des exemples pour lesquels le résultat n'est pas a priori évident.
En fait la "démonstration" du théorème des accroissements finis, démontre, non pas la réponse à l'évidente question (1), mais l'adéquation du modèle "nombres réels, fonctions dérivables" pour la réalité "voiture qui roule avec un compteur de vitesse".
Cette adéquation, on le sait, n'est d'ailleurs que partielle et approximative (il est douteux que $\mathbb{R}$ modélise parfaitement l'écoulement du temps, ou l'espace sur une ligne).

Bref d'une part, il est absurde de vouloir démontrer le théorème des accroissements finis en terminales, d'autre part, ou ne devrait jamais l'enseigner (en DEUG par exemple) en tant que preuve d'une chose évidente, mais en tant que "vérification de l'adéquation d'un modèle mathématique à la réalité qu'il veut représenter".

---

1 c'était en l'année 1987 où a été écrit cet article

Première discussion

Plaçons nous au niveau du DEUG, et supposons que nous ayons déjà fait sentir suffisamment en quoi la structure $\mathbb{R}$ est un assez bon modèle pour certaines choses, et en quoi une fonction réelle uniformément continue sur un intervalle $[t_0, t_1]$ est un assez bon modèle pour la description du mouvement d'un mobile sur une ligne.

Nous en arrivons au chapitre "dérivée comme modélisation de la notion de vitesse" .

Nous parlons de la vitesse moyenne, et nous arrivons au point : qu'il serait agréable d'avoir une fonction "vitesse instantanée".

Si nous utilisons le réflexe de Pavlov, nous aurons tendance à débiter ce qu'on nous a enseigné quand nous étions jeunes, c'est-à-dire en fait ce qu'une certaine école mathématique nous a enseigné comme étant "la vérité" .

Dans ce cas nous définirons, sans plus nous poser de problème, la vitesse instantanée comme limite de la vitesse moyenne au sens de la définition usuelle.

Si nous prenons un tout petit peu de recul, et si nous nous débarrassons de cet autre réflexe de vouloir toujours rechercher des hypothèses minimales (même quand, dans la pratique mathématique, les problèmes ne se présentent jamais avec ces seules hypothèses minimales), nous serons amenés à nous demander si la définition minimale est, après tout, bien "réaliste". En classe de physique (seconde ou après), la vitesse à l'instant t_0 est mesurée expérimentalement comme étant:

$$\big(f(t_0 + \Delta t) - f(t_0 - \Delta t) \big) / 2\,\Delta t$$

avec Δt petit. Ceci suggérerait une aussi bonne définition minimale, de la "dérivée de la fonction f au point t_0" qui serait:

$$f'(t_0) = \lim_{h \to 0} \frac{f(t_0 + h) - f(t_0 - h)}{2h} \qquad (3)$$

Avec cette nouvelle définition, certaines fonctions qui étaient non dérivables en t_0 , deviennent dérivables, comme par exemple la fonction $f(t) = |\,t\,|$ (qui devient dérivable en 0 avec $f'(0) = 0$) .

Évidemment, vous sentez bien que cette définition est peut-être "un peu trop" minimale(2). Mais pourquoi alors la définition usuellement donnée en classe de mathématiques serait-elle "juste assez minimale mais pas trop" ? Si on ne veut pas utiliser des arguments purement esthétiques ("c'est plus élégant", "ça donne de plus jolis théorèmes", "ça me plaît mieux", "si vous remettez tout en cause, vous risquez de recevoir votre poing dans la figure par vos élèves à la sortie, près du sous-bois(3)" etc...) on est bien obligé de se fatiguer un peu et d'ouvrir une discussion qui, ordinairement, est renvoyée dans les "notes historiques", mais qui, du point de vue de la formation mathématique, devrait être partie intégrante du cours.

Si la 2ème définition est "trop" minimale, la 1ère l'est aussi, nécessairement, et pour la même raison: le fait de fixer, arbitrairement, un accroissement égal de chaque côté, est déplaisant pour une raison de fond (et pas seulement esthétique). De même, le fait de fixer, arbitrairement, un

---

2 Avec cette définition le "théorème de Rolle" devient faux, et cela a de quoi chagriner bien des gens

3 cf. une déclaration de Jacques Médecin sur "les instituteurs marxistes " qui les menaçait en ces termes, ... depuis, il a d'autres chats à fouetter

accroissement égal à h d'un côté et à 0 de l'autre, est encore plus déplaisant. Nous en arriverons donc à la définition un peu moins minimale suivante (f étant supposée continue):

la dérivée de f au point t_0 est la limite suivante, si elle existe

$$\lim_{h,k \longrightarrow 0,\ h \neq k} \frac{f(t_0 + h) - f(t_0 + k)}{h - k} \qquad (4)$$ [4]

Une démarche plus systématique

Jusqu'à présent, nous avons raisonné, par jeu, en digression à partir de la définition "standard".

Mais nous avons en quelque sorte continué à jouer le même jeu, en ce sens que nous n'avons pas procédé systématiquement, ce qui nécessite de commencer par préciser les propriétés que nous voulons voir vérifier par le modèle mathématique.

Personnellement, il me semble que <u>la</u> propriété que nous voulons voir vérifier à la fonction dérivée est la propriété (1) . Celle justement que votre cousin, qui ne fait pas de mathématique, trouverait absolument aberrant qu'elle ne soit pas vérifiée.
En langage mathématique formalisé, nous obtenons la traduction (5) de la propriété (1)

La fonction $g : [\, t_0, t_1\,] \longrightarrow \mathbb{R}$, uniformément continue, sera dite fonction dérivée de la fonction $f : [\, t_0, t_1\,] \longrightarrow \mathbb{R}$ si on a :

$$\forall\, a, b \in [\, t_0, t_1\,] \quad \forall\, m, M \in \mathbb{R} \quad \text{avec } a < b \text{ et } m < M$$

$$g([\, a\, ,\, b\,]) \subset [m, M] \quad \Rightarrow \quad \frac{f(b) - f(a)}{b - a} \in [m, M]$$

Insistons sur le fait que (5) est **exactement la traduction de** (1), avec l'a priori de modéliser au moyen de fonctions uniformément continues de $[\, t_0, t_1\,]$ vers $\mathbb{R}$. [5]

Notons que la définition (5) implique que g est la dérivée au sens "usuel" (preuve donnée un peu plus loin).

Expliquons ici la suite de l'article:

- tout d'abord nous montrons que la définition (5) implique une propriété (6) "locale au sens uniforme", qui est la définition de la dérivabilité en mathématiques constructives;
- ensuite nous montrons que (6) implique (5);
- ensuite nous vérifions que (6) est une version très agréable, pour laquelle les "théorèmes classiques" se démontrent très facilement;
- une conclusion est que les fonctions usuelles vérifient bien "le théorème des accroissements finis" c'est-à-dire (5), et que ceci a été démontré par une preuve simple, et constructive (contrairement à la preuve que subissent les étudiants);

---

4 l'appellation officielle pour cette propriété est : f est strictement dérivable en t_0

5 Le "réflexe minimaliste" inciterait à ne pas demander la continuité uniforme, au moins pour la fonction g. Sans vouloir discuter cette question en tant que telle, je signalerai que, dans la mesure où on modélise au moyen de fonctions, il est à peu près inévitable de s'en tenir aux fonctions uniformément continues (pour un segment borné) ceci parce que:

- (a) la continuité est "dans la nature de la modélisation" (les mesures en physique sont toujours "approchées", sauf lorsque la variable est dans un ensemble discret),
- (b) l'uniforme continuité est la forme manipulable de la continuité.

– nous comparerons enfin avec la démonstration usuelle de (5) donnée en général avec des hypothèses minimales.

Équivalence de la définition "globale" avec une définition "locale uniforme"

La définition (5) peut être qualifiée de globale. Elle implique la propriété (6) suivante, "locale au sens uniforme", qui est la définition de la dérivabilité donnée en mathématiques constructives.

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x, y \in [t_0, t_1] :$$
$$|y - x| \leq \delta \Rightarrow |f(y) - f(x) - g(x)(y - x)| \leq \varepsilon |x - y| \qquad (6)$$

En effet si on a $|x - y| \leq \delta \Rightarrow |g(y) - g(x)| \leq \varepsilon$
alors "en fixant x" on obtient :

$$|x - y| \leq \delta \Rightarrow g(y) \in [g(x) - \varepsilon, \ g(x) + \varepsilon]$$

et en appliquant (5) pour l'intervalle $[x - \delta, x + \delta] \cap [t_0, t_1]$ on obtient (6).
On notera que (6) implique que $g(x)$ est la dérivée de f au point x "au sens usuel" mais aussi au sens de la stricte dérivabilité (définition (4)). [6]

Avant de démontrer (6) $\Rightarrow$ (5) notons que (6) admet la formulation équivalente suivante :

la fonction $z \longmapsto g(z)$ est limite uniforme, lorsque y et x distincts tendent vers z, de la fonction "taux d'accroissement de f"[7]

$$h(x,y) := (f(y) - f(x)) / (y - x) \qquad (7)$$

Or cette définition peut être lue comme suit :

la fonction "dérivée" est la limite de la "fonction vitesse moyenne" lorsque l'intervalle de temps tend vers 0.

Autrement dit, du point de vue strictement mathématique, la phrase "la vitesse instantanée est la limite de la vitesse moyenne" recèle un flou suffisamment important pour laisser ouvertes une foultitude d'interprétations très diverses, aussi bien celle correspondant à la définition usuelle que celle correspondant à la définition constructive.
Passons à la preuve que (6) $\Rightarrow$ (5).
Supposons (6), et, avec $a, b \in [t_0,t_1]$, supposons $g([a,b]) \subset [m,M]$. On veut montrer que le taux d'accroissement de la fonction f sur l'intervalle $[a,b]$ est compris entre m et M. Soit $\varepsilon > 0$, il suffira de montrer que le taux d'accroissement est inférieur à $M + \varepsilon$ (la même démonstration prouverait qu'il est supérieur à $m - \varepsilon$).
Considérons un δ convenant à ε dans (6), et découpons l'intervalle $[a,b]$ en intervalles $[x_i,x_{i+i}]$ de longueur $\leq \delta$.
On a alors

$$|f(x_{i+i}) - f(x_i) - g(x_i)(x_{i+i} - x_i)| \leq \varepsilon |x_{i+i} - x_i|$$

d'où

$$(f(x_{i+i}) - f(x_i)) / (x_{i+i} - x_i) \in [m - \varepsilon, M + \varepsilon]$$

---

6 On peut montrer, qu'une fonction qui admet une dérivée "au sens usuel" uniformément continue, est dérivable au sens (6). On peut aussi montrer (au moins en mathématiques classiques) qu'une fonction qui est partout strictement dérivable est dérivable au sens (6)

7 On peut donc écrire (modulo une petite modification en t^1) que g est limite uniforme de la suite de fonctions : h",,(x) : = n(f(x+~)- f(x)) . Cela montre que l'hypothèse "g uniformément continue", nécessaire dans (5), est superflue dans (6)

en particulier

$$f(x_{i+i}) - f(x_i) \leq (M + \varepsilon)\ (x_{i+i} - x_i)$$

et enfin

$$f(b) - f(a) \leq (M+\varepsilon)\ (b - a)$$

Il est bon de signaler que cette preuve est essentiellement celle que Cauchy donne pour le théorème des accroissements finis. C'est en fait une preuve simple et naturelle, celle qui va de soi lorsqu'on a formalisé la notion de limite et qu'on cherche à en vérifier l'efficacité. Que cette preuve de Cauchy soit réputée fautive est à vrai dire un grand sujet d'étonnement, car la définition qu'il donne de la fonction dérivée (il ne donne pas celle de la dérivée en un point) est trop imprécise pour qu'on puisse trancher entre l'interprétation maintenant usuelle et l'interprétation (6).

Efficacité de la définition "locale uniforme"

La définition (6) "locale uniforme" est celle donnée en mathématiques constructives (cf. [BB]). C'est une définition "très efficace" parce qu'elle permet de donner des démonstrations constructives des théorèmes essentiels sur les fonctions dérivées.

Nous laissons à la lectrice le soin de vérifier que les preuves classiques concernant la dérivabilité d'une somme, d'un produit ou d'un quotient, ou celle qui donne l'existence d'une primitive pour une fonction uniformément continue, fonctionnent pratiquement sans aucun changement avec la définition (6), et sont des démonstrations entièrement constructives.

Ici il faut que nous expliquions un peu en quoi une preuve peut être déclarée constructive. Par exemple avec le théorème des accroissements finis lui-même.

Nous le prenons sous la forme équivalente suivante :

(TAF1) Si la fonction f est dérivable au sens (6) et si $f' \geq 0$ sur $[a, b]$ alors $f(b) \geq f(a)$

Si nous examinons en détail la démonstration que nous avons donnée, elle donne en fait un résultat qui a une signification algorithmique plus forte.

(TAF2) Si f est dérivable au sens (6) et si $f(b) < f(a)$, alors on peut expliciter en un temps fini un x dans $[a, b]$ tels que $f'(x) < 0$

Plus précisément cette démonstration fournit le plan d'un algorithme qui traite les données f, a, b, (la fonction f étant dérivable au sens (6)) et fournit en sortie un rationnel $r \in [0,1]$ tel que $f'(a+r(b-a)) < 0$. ([8])

À partir de la définition "usuelle" de dérivée, il est également possible de fournir une preuve algorithmique de (TAF1), mais nettement moins intuitive et dont le caractère algorithmique est

---

8 Un problème qui se pose est "comment un algorithme traite-t-il une fonction uniformément continue en tant que donnée ?". C'est-à-dire encore: quelle est la forme "manipulable" (algorithmique) prise par une fonction continue $[a, b] \longrightarrow \mathbb{R}$? Une possibilité est la suivante.
Nous avons besoin de 2 choses : d'une part, le module de continuité uniforme de f, qui est une fonction $mc : \mathbb{N} \longrightarrow \mathbb{N}$ vérifiant : $(|x - y| \leq mc(n)) \Rightarrow (|f(x) - f(y) \leq 1/2^n)$. Et d'autre part, pour chaque $n \in \mathbb{N}$ et chaque nombre rationnel $t \in [a, b]$, nous devons connaître un rationnel z vérifiant $|f(t) - z| \leq 1/2^n$.
Ainsi, un algorithme qui traite "une donnée qui est une fonction uniformément continue f " peut être conçu comme un algorithme utilisant 2 "oracles" :
le premier répond $mc(n)$ lorsqu'on lui pose la question "module de continuité pour l'écart $1/2^n$? ",
le deuxième répond z (rationnel) lorsqu'on lui pose la question "valeur de f au rationnel t avec la précision $1/2^n$? " ($t \in \mathbb{Q} \cap [a,b]$, $n \in \mathbb{N}$).

moins convaincant. En effet, cette preuve ne fournit pas (TAF2) mais une version une peu plus faible (TAF3), qui implique aussi (TAF1), mais dans laquelle le réel x cherché n'est pas explicité en un temps fini; c'est seulement le n-ème terme d'une suite de rationnels convergeant vers x (de manière géométrique) qui est explicité en temps fini :

(TAF3) Si f est dérivable au sens usuel sur l'intervalle $[a, b]$ et si $f(b) < f(a)$, alors on peut expliciter un $x \in [a, b]$ tel que $f'(x) < 0$.

Voici cette preuve, qui fonctionne par dichotomie (quelques détails techniques supplémentaires devraient être rajoutés, mais l'essence de la preuve est donnée).
On note $t = (f(b) - f(a))/(b - a)$. On considère $\varepsilon > 0$, $\varepsilon_n = \varepsilon/2^n$. A l'étape numéro n on détermine un intervalle $[a_n,b_n]$ de longueur $(b - a)/2^n$ sur lequel le taux d'accroissement de la fonction f est $\geq t - \varepsilon + \varepsilon_n$, les intervalles successifs étant emboités. L'initialisation avec $n = 0$ et l'intervalle $[a,b]$ est claire. Ensuite on remarque que lorsqu'on divise un intervalle en deux intervalles égaux, le taux d'accroissement global T est la moyenne des deux taux d'accroissements T' et T'' sur les intervalles de longueur moitié. Il suffit alors de calculer chacun des deux réels T' et T'' avec une précision suffisante, pour être sûr que l'un des deux au moins est supérieur $T - \delta$ où δ est un rationnel strictement positif arbitraire. Enfin, la limite commune des deux suites (a_n) et (b_n) est un réel x de $[a,b]$ en lequel le nombre dérivé (supposé exister) est $\geq t - \varepsilon$.

Si nous comparons le contenu algorithmique des deux preuves, nous voyons que la première fournit un rationnel r convenable en un nombre fini d'étapes, tandis que la seconde fournit un réel x comme limite d'une suite de Cauchy (explicite et explicitement convergente).
Mais surtout, bien que nous venions de faire fonctionner la définition usuelle à l'intérieur d'une preuve de nature algorithmique, la définition usuelle de la dérivabilité en tout point n'a pas de signification algorithmique connue. Plus précisément, personne n'a la moindre idée de comment on pourrait rendre la définition usuelle de la dérivabilité d'une fonction en tout point d'un intervalle, entièrement explicite. En effet, on ne sait pas expliciter le fait que "le nombre dérivé existe en tout point de l'intervalle" autrement qu'en exprimant ce nombre dérivé comme fonction uniformément continue de la variable x et en donnant un "module de convergence", uniforme par rapport à x, qui explicite la manière dont le taux d'accroissement tend vers le nombre dérivé en x. Mais alors on tombe sur la définition (6) et non sur la définition usuelle.

Signalons en outre que, puisque les théorèmes essentiels sur les fonctions dérivées sont démontrables constructivement avec la définition (6), nous obtenons une preuve complètement explicite du théorème (TAF2) lorsque la fonction f est une fonction polynome, ou une fraction rationnelle dont le dénominateur reste $\geq m > 0$ sur un intervalle $[a, b]$.
Ceci nous amène à poser la question suivante : le théorème des accroissements finis (sous la forme (TAF2)) est-il un théorème d'algèbre ? (f est un polynome ou une fraction rationnelle dont le dénominateur est supposé $\geq m > 0$ sur l'intervalle considéré, et le corps des réels est remplacé par un corps ordonné arbitraire).
La réponse est sûrement oui lorsque l'on considère un corps ordonné $K \subset \mathbb{R}$, puisque la preuve "complètement explicite" de (TAF2) n'utilise aucun passage à la limite[9] mais seulement des minorations et des majorations. Il existe également une preuve "purement algébrique" dans le cas général d'un corps ordonné, mais seulement avec une fonction polynome, nous la donnons en annexe.

---

9 Contrairement à celle donnée pour (TAF3)

Mais revenons pour terminer à un autre aspect de la discussion générale sur le caractère plus ou moins constructif, algorithmique, des preuves en analyse. Et ceci, à partir de l'exemple du "théorème de Rolle".

Le théorème de Rolle : un théorème non algorithmique

Donnons tout d'abord la version constructive du théorème de Rolle:

(R1) Si la fonction f est dérivable au sens (6) et si $f(a) = f(b)$, alors $\inf \{ | f'(t) | ;\ t \in [a,b] \} = 0$

Preuve algorithmique : soit un rationnel $\varepsilon > 0$ et découpons $[a,b]$ en intervalles $[x_i, x_{i+i}]$ de longueur $\leq \delta$ sur chacun desquels la fonction f' varie de moins que $\varepsilon/4$. Pour chaque i, on calcule $f'(x_i)$ avec une précision meilleure que $\varepsilon/4$, ce qui permet d'affirmer au moins une des 2 éventualités:

$$| f'(x_i) | \geq \varepsilon/2 \quad \text{ou} \quad | f'(x_i) | \leq \varepsilon$$

Si on avait $| f'(x_i) | \geq \varepsilon/2$ pour chaque i, alors $f'(x_i)$ resterait de signe constant (car $| f'(x_{i+i}) - f'(x_i) | \leq \varepsilon/4$), par exemple > 0, et on aurait, par (5)

$$f(b) \geq f(a) + \varepsilon/4\ (b - a)$$

Cette preuve algorithmique nous donne le moyen de trouver, pour chaque entier n, en un nombre fini d'étapes, majoré à priori, un rationnel x_i tel que

$$| f'(x_i) | \leq 1/2^n$$

Elle ne nous donne cependant pas le moyen de construire un nombre réel x (c'est-à-dire : une suite de Cauchy de rationnels) pour lequel $f'(x) = 0$.
En outre, il semble difficile de modifier convenablement cette preuve pour la rendre explicite dans le cas où la fonction f est seulement supposée dérivable au sens usuel.

Quant à la preuve "usuelle" du théorème de Rolle, elle est basée sur le fait que la fonction f atteint son maximum sur l'intervalle $[a, b]$. Mais ce "théorème" d'analyse est fortement non algorithmique : en effet, on peut construire une fonction f uniformément continue sur un intervalle $[a,b]$, de manière complètement explicite, et telle que cependant f n'atteigne son maximum en aucun point "mécaniquement calculable" de l'intervalle $[a,b]$. En conséquence, aucun algorithme ne peut fournir, à partir d'une fonction uniformément continue arbitraire explicite f, un point x où cette fonction atteindrait son maximum. Même dans l'hypothèse où f est supposée dérivable au sens (6), il est impossible de réaliser le théorème de Rolle par un algorithme général, parce que cet algorithme général serait capable de réaliser le théorème de la valeur intermédiaire pour la fonction dérivée([10]).

Conclusion

Une préoccupation de réalisme nous a tout d'abord incité à changer le statut du théorème des accroissements finis : il démontre, non pas "une vérité", mais l'adéquation d'un modèle à une réalité qu'il voulait représenter.
Cette même préoccupation de réalisme nous a ensuite conduit à proposer une définition alternative pour la notion de "fonction dérivable".
Sans machiavélisme prémédité, nous avons abouti à une formulation équivalente à celle donnée par les mathématiques constructives. Cette définition alternative de fonction dérivable s'avère en

---

[10] Voir le chapitre « Hier et Demain » pour une discussion sur le théorème de la valeur intermédiaire.

fin de compte très efficace du point de vue algorithmique, et fournit par exemple une preuve "purement algébrique" du théorème des accroissements finis dans le cas d'une fraction rationnelle à dénominateur $\geq m > 0$ et à coefficients rationnels.

Pour passer de la formulation usuelle "vitesse instantanée = limite de la vitesse moyenne" à la formulation constructive alternative, il suffit de lire le mot "limite" au sens uniforme.

Ceci n'est pas sans rappeler une mésaventure qui arriva à Cauchy, célèbre pour avoir introduit la rigueur en analyse. Cauchy, (qui passait parfois pour un dangereux faiseur de contre-exemples irréalistes), proposa une définition de la continuité (c'était la continuité en tout point, formulée dans son langage) et, comme il ne prétendait pas "démontrer des vérités", mais "vérifier l'adéquation d'un modèle", il s'attaqua à un "théorème" qui, à l'époque, avait quasiment le statut d'une évidence: "si une suite de fonctions continues admet une limite, sa limite est une fonction continue".

Ce théorème n'est ni vrai, ni faux. Tout dépend du contexte, du sens précis qu'on attribue au mot limite. Cauchy donna en 1821 une démonstration "fausse". On connut rapidement des contre-exemples. Mais l'erreur dans la démonstration ne fut découverte que 26 ans plus tard. C'est Seidel qui mit clairement à jour le point faible dans le modèle: la notion de limite d'une suite de fonctions devait être prise "au sens uniforme" et non au sens "point par point", le premier qui venait à l'esprit. Pour plus de détails sur ce sujet, on lira "Preuves et réfutations" de Imre Lakatos, p. 165-182.

Annexe : le théorème algébrique des accroissements finis

Il affirme que le taux d'accroissement d'un polynome de degré fixé sur un intervalle est une moyenne, pondérée par des coefficients positifs, de valeurs de la dérivée en des points connus de l'intervalle, indépendants du polynome.

Exemple : Pour tout polynôme de degré ≤ 4 on a l'identité:

$$P(a) - P(b) = (a - b)\ (P'(a/6+5b/6)/3+P'(a/3+2b/3)/6+P'(2a/3+b/3)/6+P'(5a/6+b/6)/3)$$

Plus généralement on a les résultats suivants :

Lemme : Il existe deux suites $(\lambda_{i,j})_{1\leq i\leq j;j=1,2,\ldots,n,\ldots}$ et $(r_{i,j})_{1\leq i\leq j;j=1,2,\ldots,n,\ldots}$ de rationnels $\in\ ]\,0\,,\,1\,[$ telles que, pour tout polynôme $P \in \mathbb{Q}[X]$ de degré $\leq n$, on ait :

$$P(a) - P(b) = (a - b) \times \sum_{i=1}^{n} r_{i,n} \,.\, P'(a+\lambda_{i,n}(b - a))$$

Théorème 1 : (théorème algébrique des accroissements finis pour les polynomes)

Il existe deux suites $(\lambda_{i,j})_{1\leq i\leq j;j=1,2,\ldots,n,\ldots}$ et $(r_{i,j})_{1\leq i\leq j;j=1,2,\ldots,n,\ldots}$ de rationnels $\in\ ]\,0\,,\,1\,[$ telles que, pour tout corps ordonné $\mathbb{K}$ et tout polynôme $P \in \mathbb{K}[X]$ de degré $\leq n$, on ait :

$$P(a) - P(b) = (a - b) \times \sum_{i=1}^{n} r_{i,n} \,.\, P'(a+\lambda_{i,n}(b - a))$$

En particulier, si P' est de signe positif sur un intervalle, la fonction polynôme est croissante sur tout l'intervalle, et si le taux d'accroissement de P est <0 entre a et b la dérivée de P est <0 en l'un au moins des points $a+\lambda_{i,n}(b - a)$

preuve> Le théorème est une conséquence immédiate du lemme: ce dernier fournit en effet des identités algébriques concernant les variables " a, b, et les coefficients du polynôme " qui s'appliquent alors dans tout anneau commutatif qui est une $\mathbb{Q}$-algèbre, et en particulier dans les corps ordonnés.

Démontrons le lemme.

Par changement de variable affine, on se ramène au cas où $a = -1$ et $b = 1$. Considérons le degré n fixé. L'application $P \longmapsto P(1) - P(-1)$ est une forme linéaire ne faisant pas intervenir le coefficient constant. Les formes linéaires ne faisant pas intervenir le coefficient constant forment un espace de dimension n. Pour tout choix de n rationnels $\lambda_{i,n}$, les formes linéaires $P \longmapsto P'(\lambda_{i,n})$ sont indépendantes et ne font pas intervenir le coefficient constant. Il correspond donc à ce choix des rationnels $r_{i,n}$ qui rendent la formule vraie. La difficulté consiste à déterminer des $\lambda_{i,n} \in\ ]\,0\,,\,1\,[$ tels que les $r_{i,n}$ correspondants soient également dans $]\,0\,,\,1\,[$. Les formules de quadrature de Gauss correspondent à un tel choix, mais avec des réels alors que nous voulons des rationnels. Il suffit alors de choisir des $\lambda_{i,n}$ rationnels suffisamment voisins des $\lambda_{i,n}$ de Gauss (zéros des polynômes de Legendre) pour que les $r_{i,n}$ correspondants restent positifs. ❑