

VARIATIONS SUR LA DÉRIVATION

Luc TROUCHE,
IREM de Montpellier

Il était vraiment possible que le déplacement de la banquise ne fût qu'apparent, et qu'au contraire, l'île Victoria, entraînée par le champ de glace, dérivât vers le détroit. Mais cette dérive, si elle existait, on ne pouvait la constater, on ne pouvait l'estimer, on ne pouvait la relever ni en longitude, ni en latitude.

Jules Verne, *le Pays des fourrures.*

Dans ce numéro de Repères, Jean-Luc Gasser évoque les problèmes posés par la formule de calcul approché de la dérivée utilisée par les physiciens. Surprenant (en apparence !) : c'est la même formule qui est implantée sur la plupart des calculatrices graphiques pour ce calcul. Coïncidence fortuite, ou manifestation d'une convergence plus profonde ?

Introduction : l'instrument boîte à outils/boîte à problèmes

Sur la plupart des calculatrices graphiques, on trouve aujourd'hui une commande (*nDeriv* pour Texas Instruments par exemple) qui permet d'obtenir une valeur approchée du nombre dérivée d'une fonction en un point (où elle est dérivable !). Cette commande coexiste d'ailleurs, sur les

calculatrices symboliques, avec une commande de calcul formel des dérivées et donc de calcul exact du nombre dérivé.

Sur une calculatrice graphique "ordinaire" (fig. 1) l'écran d'une TI-82), on trouve la commande *nDeriv* dans le Menu Mathématique :

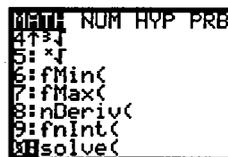


Figure 1

– *nDeriv* (*sinx*, *x*, 0) fournit une valeur approchée de la dérivée de la fonction *sinx*, de variable *x*, en 0 (sans indication de précision) ; cette commande fournit le résultat de 0, 999.999.833.3 (fig. 2)

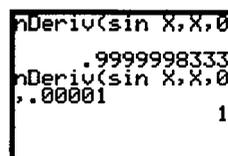


Figure 2

VARIATIONS SUR LA DÉRIVATION

– $nDeriv(\sin x, x, 0, h)$ donne une valeur approchée du même nombre pour un pas de calcul de h ; ci-contre un pas de 10^{-5} donne comme valeur approchée 1.

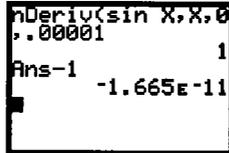


Figure 3

Un peu moins que 1 d'ailleurs : l'examen du troisième écran (fig. 3) indique que, si l'on soustrait 1 au nombre dérivé donné, on trouve une différence de $-1,665 \cdot 10^{-11}$.

L'utilisateur ne connaît pas nécessairement le mécanisme qu'il met en œuvre quand il actionne la commande $nDeriv$ – sauf s'il regarde le mode d'emploi avec beaucoup d'attention.

L'explicitation de cette commande est par contre possible avec une calculatrice symbolique.

Voici par exemple l'écran d'une TI-92 qui fait apparaître, dans le Menu Calcul de l'application Derive :

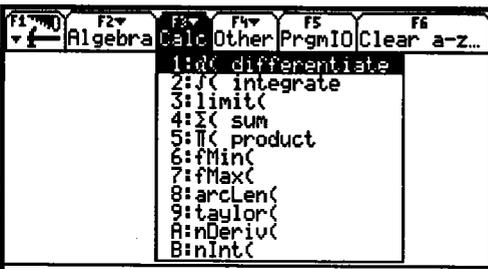


Figure 4

- une commande de calcul exact des dérivées (1 : differentiate) ;
- une commande de calcul approché (A : $nDeriv$).

L'appel de cette dernière commande fait apparaître son contenu :

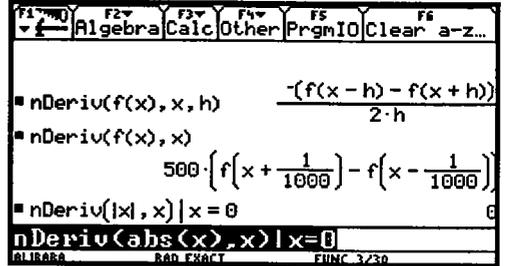


Figure 5

$nDeriv(f(x), x, h)$ donne la pente de la corde entre les points d'abscisse $(x-h)$ et $(x+h)$:

- si la valeur de h n'est pas mentionnée, la valeur $\frac{1}{1000}$ est choisie par défaut ;
- la particularisation en un point ($|x=0$ veut dire sachant que $x=0$) permet d'obtenir une valeur approchée de la dérivée en un point.

Cette commande de calcul approché correspond donc à la formule qu'utilisent les physiciens pour approcher les vitesses en un point.

On peut se poser à propos de cette commande (au moins) trois questions :

- une première question, théorique : le processus symbolisé par $nDeriv$ est-il bien convergent, autrement dit a-t-on $\lim_{h \rightarrow 0} nDeriv(f, x, h) = f'(x)$?

- une deuxième question, théorico-pratique : pourquoi avoir choisi une approche de la dérivée, symétrique, par $\frac{f(x+h) - f(x-h)}{2h}$ et pas par la formule, plus classique (dans le cours de mathématique), $\frac{f(x+h) - f(x)}{h}$?

– une troisième question, à nouveau théorique : pourrait-on imaginer une approche encore meilleure de la dérivée par d'autres formules ?

1. Un processus convergent sous conditions

A propos d'abord de nDeriv.

Soit donc :

$$nDeriv(f, x, h) = \frac{f(x+h) - f(x-h)}{2h}$$

a-t-on $\lim_{h \rightarrow 0} = f'(x)$?

Il est clair d'abord que l'existence de $\lim_{h \rightarrow 0} nDeriv$ n'implique pas la dérivabilité de la fonction f en x . L'exemple de la fonction $x \rightarrow |x|$ en 0 (cf. copie d'écran fig. 5) le montre bien. Plus généralement d'ailleurs, pour toute fonction f paire, nous disposerons de $f(h) = f(-h)$, donc $nDeriv(f, 0, h) = 0$, sans que cela n'implique nécessairement que f soit dérivable en 0 !

Réciproquement, si f est dérivable en x , nous pourrons écrire :

$$* f(x+h) = f(x) + hf'(x) + h\varepsilon(h)$$

avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$;

$$* f(x-h) = f(x) - hf'(x) - h\varepsilon'(h)$$

avec $\lim_{h \rightarrow 0} \varepsilon'(h) = 0$.

Par soustraction, nous obtenons :

$$nDeriv(f, x, h) = \frac{f(x+h) - f(x-h)}{2h} = f'(x) + h[\varepsilon(h) + \varepsilon'(h)].$$

Ceci prouve bien que $\lim_{h \rightarrow 0} nDeriv = f'(x)$.

Géométriquement, on peut d'ailleurs constater que $nDeriv(f, x, h)$ est l'isobary-

centre de $\frac{f(x+h) - f(x)}{h}$ et de $\frac{f(x-h) - f(x)}{-h}$, c'est-à-dire que $\lim_{h \rightarrow 0} nDeriv$ apparaît

comme la moyenne arithmétique des dérivées à gauche et à droite de f . Ceci explique que $\lim_{h \rightarrow 0} nDeriv$ existe (et égale 0)

lorsque la fonction n'est pas dérivable en x mais que les dérivées à gauche et à droite existent.

A propos de processus d'approches plus généraux.

Ce que nous venons de voir s'applique à toute formule du type $\frac{f(x+g(h)) - f(x-g(h))}{2g(h)}$,

où g est une fonction de h qui a pour limite 0 en 0. Nous en reparlerons d'ailleurs dans la troisième partie de cet article. Peut-on théoriquement envisager un résultat encore plus général, en envisageant un processus d'approche non symétrique à gauche et à droite de x , c'est-à-dire peut-on affirmer que, si f est dérivable en x , $\lim_{(a,b) \rightarrow (x,x)} \frac{f(a) - f(b)}{a - b} = f'(x)$?

(on dira dans ce cas que la fonction f est strictement dérivable en x , cf. l'article de Henri Lombardi dans ce numéro de *Repères*).

Pour cause d'unicité, on peut affirmer que, si cette limite existe, elle est égale nécessairement à $f'(x)$. Mais est-on sûr que cette limite existe ? Examinons cela de plus près : la forme même de l'expression invite à utiliser le théorème des accroissements finis.

$\frac{f(a) - f(b)}{a - b} = f'(c)$ avec c appartenant à $]a, b[$. Si a et b tendent simultanément vers x , c va tendre aussi vers x . Si la fonction f' est continue en x , on pourra alors

VARIATIONS SUR LA DÉRIVATION

affirmer que $f'(c)$ tendra vers $f'(x)$. Conclusion : dans ce cas, $\lim_{(a,b) \rightarrow (x,x)} \frac{f(a) - f(b)}{a - b} = f'(x)$.

Mais que dire si f' n'est pas continue en x ? On peut examiner un exemple pour se faire une petite idée. Trouver une fonction dérivée non continue n'est pas simple : une fonction dérivée vérifie la propriété des valeurs intermédiaires (d'après le théorème de Darboux), on sait aussi que si une fonction dérivable est convexe (ou concave) sur un intervalle, alors sa dérivée est nécessairement continue. Nous devons donc trouver un exemple de fonction que ne soit ni convexe ni concave au voisinage d'un point (c'est-à-dire qui oscille assez frénétiquement). On peut penser aux fonctions trigonométriques...

Soit la fonction f définie par $f(0) = 0$ et $f(x) = x^2 \sin \frac{1}{x}$ pour $x \neq 0$. La présence du facteur x^2 assure la continuité et la dérivabilité de f en 0 :

$$f'(0) = \lim_{h \rightarrow 0} \frac{f(h) - f(0)}{h} = \lim_{h \rightarrow 0} h \sin \frac{1}{h} = 0.$$

Cette contrainte imposée par le facteur x^2 se retrouve aussi graphiquement (cf. fig. 6) : le graphe de la fonction f est

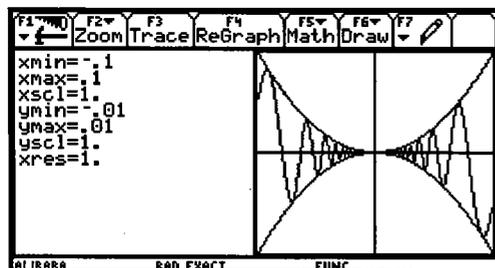


Figure 6

“coincée” entre les graphes des fonctions $x \rightarrow x^2$ et $x \rightarrow -x^2$.

Par ailleurs, pour $x \neq 0$, $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$. Le premier terme a pour limite 0 en 0 . Le deuxième terme n'a pas de limite en ce point. La fonction f' n'est donc pas continue en 0 . C'est bien le type de fonction que nous cherchions.

Examinons donc maintenant

$\lim_{(a,b) \rightarrow (0,0)} \frac{f(a) - f(b)}{a - b}$. On peut commencer notre étude par une observation graphique en considérant la fonction g de deux variables :

$$g \rightarrow g(x,y) = \frac{f(x) - f(y)}{x - y}$$

Ci-contre (fig. 7) un point de vue sur la représentation graphique de la fonction g . On a choisi pour x et y le même intervalle de variation $[-0,1 ; 0,1]$ que pour l'observation de la fonction f à une variable.

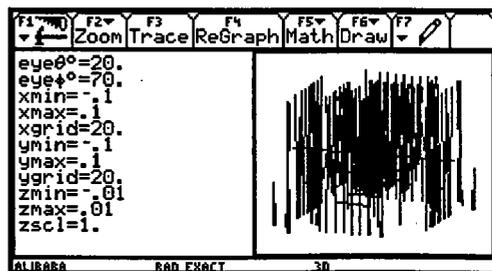


Figure 7

La relative régularité de la fonction f laisse ici la place à un relatif désordre (apparent bien sûr).

Nous nous orientons donc vers la recherche d'un contre-exemple, c'est-à-dire vers la recherche d'un itinéraire de \mathbb{R}^2 où

(a, b) tend vers $(0, 0)$ avec $\frac{f(a) - f(b)}{a - b}$ ne tendant pas vers $f'(0) = 0$.

$\frac{f(a) - f(b)}{a - b}$ est la pente d'une corde passant par les points d'abscisse a et b . Nous allons construire deux suites (a_n) et (b_n) qui convergent vers 0 et examiner alors $\lim_{(a_n, b_n) \rightarrow (0, 0)} \frac{f(a_n) - f(b_n)}{a_n - b_n}$. Pour construire ces deux suites, on ne part pas tout à fait au hasard. A nouveau des considérations sur les oscillations de la fonction f peuvent donner d'utiles indications.

Ci-dessous (fig. 8) un aspect de la représentation graphique de f sur l'intervalle $[0; 0,1]$, encadrée par les fonctions x^2 et $-x^2$.

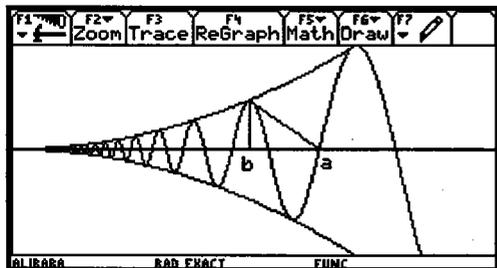


Figure 8

Si l'on considère la suite (b_n) des points de contact entre la courbe de f et la parabole supérieure et (a_n) la suite des points où la fonction, sur ses intervalles de croissance, s'annule, on peut imaginer que la pente de ces cordes ne tend pas vers 0 quand (a_n) et (b_n) tendent vers 0.

La vérification est aisée :

* $f(x) = x^2$ si et seulement si $\sin \frac{1}{x} = 1$,

c'est-à-dire $x = \frac{1}{\sqrt{2} + 2n\pi}$; ainsi $b_n = \frac{1}{\sqrt{2} + 2n\pi}$

* $f(x) = 0$ si et seulement si $\sin \frac{1}{x} = 0$, c'est-à-dire $x = \frac{1}{n\pi}$; si l'on veut en plus être sur les intervalles de croissance de f , cela impose $x = \frac{1}{2n\pi}$; ainsi $a_n = \frac{1}{2n\pi}$.

$$\begin{aligned} \text{Ainsi } \frac{f(a_n) - f(b_n)}{a_n - b_n} &= \frac{\frac{1}{(\sqrt{2} + 2n\pi)^2}}{\frac{1}{\sqrt{2} + 2n\pi} - \frac{1}{2n\pi}} \\ &= -\frac{2}{\pi} \cdot \frac{2n\pi(\sqrt{2} + 2n\pi)}{(\sqrt{2} + 2n\pi)^2} \end{aligned}$$

Pour des raisons de "rapport des termes de plus haut degré", ce rapport tend vers $-\frac{2}{\pi}$ quand n tend vers $+\infty$.

Pour conclure :

* $\lim_{(a_n, b_n) \rightarrow (0, 0)} \frac{f(a_n) - f(b_n)}{a_n - b_n} = -\frac{2}{\pi}$

* $\lim_{h \rightarrow 0} \frac{f(h) - f(0)}{h} = 0$ (pour cause de dérivée).

Pour cause d'unicité de la limite, on peut donc affirmer que $\lim_{(a, b) \rightarrow (0, 0)} \frac{f(a) - f(b)}{a - b}$ n'existe pas (pour plus d'information sur cette fonction, voir [1]).

Ainsi, si f est dérivable en x et f' non continue en ce point, l'existence de

$\lim_{(a, b) \rightarrow (0, 0)} \frac{f(a) - f(b)}{a - b}$ n'est pas garantie

Pour résumer cette première partie :

VARIATIONS SUR LA DÉRIVATION

* si f est dérivable en x,

$$\lim_{h \rightarrow 0} \frac{f(x+kh) - f(x-kh)}{2kh} = f'(x) ;$$

* si f est dérivable en x et f' continue en x,

$$\lim_{(a,b) \rightarrow (0,0)} \frac{f(a) - f(b)}{a - b} = f'(x).$$

2. Efficacité de l'approximation et contraintes de l'instrument

Comparons d'abord l'approche "classique" de f'(x) et l'approche par nDeriv.

On utilise pour cela la formule de Taylor, généralisation du théorème des accroissements finis. Nous supposons, pour simplifier les choses, que f est dérivable autant que nécessaire.

$$f(x+h) = f(x) + h.f'(x) + \frac{h^2}{2}f''(c)$$

où c ∈]x, x+h[

Ainsi $\frac{f(x+h) - f(x)}{h} - f'(x) = \frac{h}{2}f''(c)$, ce

que nous pouvons écrire :

$$\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| \leq \frac{h}{2}M_2$$

où $M_2 = \text{Max}_{[x, x+h]} |f''|$.

Pour l'approche de nDeriv, la soustraction des deux expressions nous autorise à aller un rang plus loin dans l'écriture de la formule de Taylor :

$$f(x+h) = f(x) + h.f'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(c)$$

où c ∈]x, x+h[

$$f(x-h) = f(x) - h.f'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f'''(c')$$

où c' ∈]x-h, x[

Par soustraction, on obtient

$$\frac{f(x+h) - f(x-h)}{2h} = \frac{h^3}{12} [f'''(c) + f'''(c')],$$

ce qui peut s'écrire :

$$\left| \frac{f(x+h) - f(x-h)}{2h} - f'(x) \right| \leq \frac{h^2}{6}M_3$$

où $M_3 = \text{Max}_{[x-h, x+h]} |f'''|$.

La comparaison des deux approches est éclairante :

- l'approche classique par $\frac{f(x+h) - f(x)}{h}$ est contrôlée par $\frac{h}{2}M_2$;

- l'approche par nDeriv = $\frac{f(x+h) - f(x-h)}{2h}$ est contrôlée par $\frac{h^2}{6}M_3$.

Localement, la deuxième approche est donc incontestablement meilleure. Les professeurs de physique, comme les concepteurs de calculatrices, n'ont pas fait des choix arbitraires !

Petite illustration de cette situation :

On se propose d'approcher ci-dessous (fig. 9) la dérivée de la fonction exponentielle en 0 par :

- l'approche classique "à droite" ;
- l'approche classique "à gauche"
- l'approche par nDeriv.

Le tableau de valeurs ci-dessous (fig. 10) permet de comparer les trois approches.

La qualité supérieure de l'approxi-

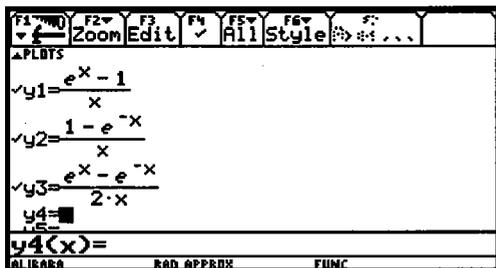


Figure 9

mation par nDeriv apparaît clairement, ainsi que son caractère isobarycentrique par rapport aux deux approches précédentes.

x	y1	y2	y3
-0.5	.78693868	1.2974425	1.0421906
-0.4	.82419988	1.2295617	1.0268808
-0.3	.86393926	1.166196	1.0150676
-0.2	.90634623	1.1070138	1.00668
-0.1	.95162582	1.0517092	1.0016675
0.	undef	undef	undef
0.1	1.0517092	0.95162582	1.0016675
0.2	1.1070138	0.90634623	1.00668

Figure 10

Précisons cependant la signification de la proposition "l'approche symétrique de f' par $\frac{f(x+h) - f(x-h)}{2h}$ est meilleure localement que l'approche par $\frac{f(x+h) - f(x)}{h}$ ".

Cette signification n'est pas la même que pour la proposition "la tangente est la droite qui approche le mieux une courbe au voisinage d'un point" (cf. l'article de Marie-Jeanne Perrin-Glorian) dans ce numéro de *Reperes* :

- dans ce dernier cas, la proposition

signifie qu'il existe effectivement un voisinage de ce point où l'approche de la courbe par la tangente est meilleure que l'approche par n'importe quelle autre droite ;

- il n'en est pas de même pour l'approche de f' par les deux formules de dérivation. On ne peut pas affirmer qu'il existera nécessairement un voisinage du point x sur lequel l'approche symétrique de f'(x) par $\frac{f(x+h) - f(x-h)}{2h}$ sera de meilleure

qualité que l'approche par $\frac{f(x+h) - f(x)}{h}$. Le caractère isobarycentrique de la première formule peut nous aider à trouver un contre exemple adéquat : il suffit de trouver une fonction pour laquelle l'approche du point à droite et à gauche diffèrent sensiblement. Prenons par exemple la fonction qui vaut x^4 à gauche de 0 et x^{20} à droite de 0 :

- elle est bien dérivable en 0. Elle est même dérivable 4 fois en ce point, ce qui fait que les formules que nous avons établies grâce aux formules de Taylor pour le contrôle de l'écart entre les dérivées et ses différentes approches s'appliquent bien ;

- cependant, vu "l'aplatissement" des fonctions puissances au voisinage de 0 implique que l'approche de la dérivée en 0 à droite sera incomparablement meilleure que l'approche à gauche : pour h positif $\frac{f(h) - f(0)}{h} = h^{19}$ et $\frac{f(h) - f(-h)}{2h} = \frac{h^{19} - h^3}{2}$.

On aura alors $0 < h^{19} < \frac{h^{19} - h^3}{2}$ | dès que $0 < 3 h^{19} < h^3$, c'est-à-dire $\forall h \in [0, \sqrt[19]{\frac{1}{3}}]$ (cf.

VARIATIONS SUR LA DÉRIVATION

ci-dessous (fig. 11) en y2 l'approche symétrique de $f'(0)$, et y3 l'approche non symétrique, bien meilleure à droite de 0.

F1	F2	F3	F4	F5	F6
← f	Setup	Cell	Header	Del	Row
x	y2			y3	
.5			-.0624990463		-.125
.1			-.0005		-.001
.01			-.0000005		-.000001
0.			undef		undef
.001			-.0000000005		1.E-57
.01			-.0000005		1.E-38
.1			-.0005		1.E-19
1.			0.		1.

x = -.5
ALIBABA RAD EXACT FINC

Figure 11

Que signifie alors l'affirmation que l'approche symétrique est "localement meilleure" ?

Reprenons les deux formules de majoration que nous avons établies :

$$\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| \leq \frac{h}{2} M_2$$

où $M_2 = \text{Max}_{[x, x+h]} |f''|$;

$$\left| \frac{f(x+h) - f(x-h)}{2h} - f'(x) \right| \leq \frac{h^2}{6} M_3$$

où $M_3 = \text{Max}_{[x-h, x+h]} |f^{(3)}|$.

Supposons que nous voulions bénéficier d'une précision de ϵ pour le calcul de $f'(x)$. Il suffira pour cela de prendre :

- avec la première formule $h \leq \frac{2\epsilon}{M_2}$;

- avec la deuxième formule $h \leq \sqrt{\frac{6\epsilon}{M_3}}$.

On aura alors $\frac{2\epsilon}{M_2} \leq \sqrt{\frac{6\epsilon}{M_3}}$ dès que $\epsilon \leq \frac{3M_2^2}{2M_3}$.

Autrement dit, à partir d'une certaine précision voulue pour le calcul de f' , le

choix de la formule d'approche symétrique nous permettra de choisir un pas de calcul moins petit. Ce gain sur h est tout à fait utile car nous travaillons avec des instruments qui disposent d'une précision limitée !

On le voit, cette idée de meilleure approche locale n'est pas très simple à préciser...

Revenons maintenant à l'interprétation des formules d'approximation de la dérivée.

Ces approches différentes en un point ont aussi un rapport avec l'interpolation polynomiale des fonctions.

Les polynômes de Lagrange permettent de construire l'unique polynôme de degré au plus n dont le graphe passe par $(n + 1)$ points d'abscisses distinctes.

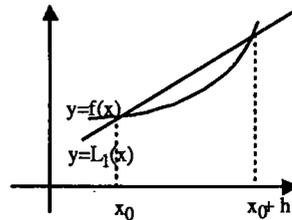


Figure 12

Si on considère l'approximation polynomiale (Lagrange) à deux points x_0 et $x_1 = x_0 + h$ ($h \neq 0$) d'une fonction f , ce polynôme peut s'exprimer par :

$$L_1(x) = \frac{x - x_1}{x_0 - x_1} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1)$$

C'est l'interpolation affine de f entre x_0 et x_1 . La dérivée de la fonction affine L_1 en x_0 est alors la pente de la corde elle-même : $\frac{f(x_0+h) - f(x_0)}{h}$.

Si on considère maintenant (fig. 13) l'approximation polynomiale à trois points : $x_{-1} = x_0 - h$; x_0 et $x_1 = x_0 + h$ ($h \neq 0$) de f , ce polynôme pourra s'exprimer par :

$$L_2(x) = \frac{(x-x_0)(x-x_1)}{(x_1-x_0)(x_1-x_1)}f(x_{-1}) + \frac{(x-x_{-1})(x-x_1)}{(x_0-x_{-1})(x_0-x_1)}f(x_0) + \frac{(x-x_{-1})(x-x_0)}{(x_1-x_{-1})(x_1-x_0)}f(x_1).$$

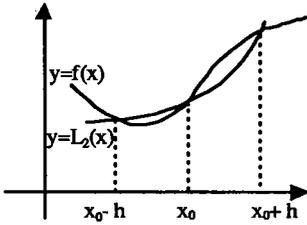


Figure 13

La dérivée de L_2 en x_0 est alors : $\frac{f(x_0+h) - f(x_0-h)}{2h}$.

* $\frac{f(x+h) - f(x)}{h}$ consiste ainsi à approcher $f'(x)$ par la dérivée en ce point de la fonction affine qui coïncide avec f en deux points $(x, x+h)$;

* $nDeriv(f, x, h)$ consiste à approcher $f'(x)$ par la dérivée en ce point du polynôme de degré au plus deux qui coïncide avec f en trois points $(x, x-h, x+h)$.

Ceci peut aussi faire comprendre la meilleure qualité de l'approximation par cette deuxième formule (cf. sur ce point [2]).

Remarque : il est clair que, pour les polynômes de degré au plus deux, $nDeriv(f,x,h)$ donne la valeur exacte de la dérivée $f'(x)$ (puisque dans ce cas l'inter-

polation à trois points de f donne... f). On a ainsi :

$$\forall x \in \mathbb{R}, \forall h \in \mathbb{R}^*, \frac{f(x+h) - f(x-h)}{2h} = f'(x) \text{ (i) .}$$

Graphiquement, cela se traduit, pour toute parabole, par le parallélisme entre une tangente au point d'abscisse a et toute corde reliant deux points d'abscisses symétriques par rapport à x (cf. illustration de la fig. 14).

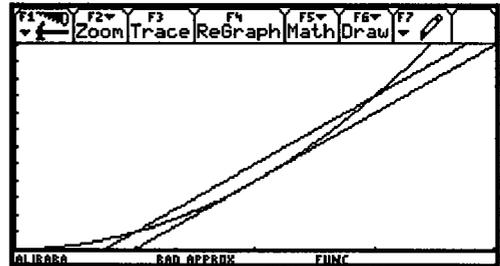


Figure 14

On peut à ce sujet se poser la question : cette propriété est-elle caractéristique des paraboles ? Elle l'est, et la démonstration est assez simple. L'égalité (i) ci-dessus peut s'écrire : $f(x+h) - f(x-h) = 2h.f'(x)$. La dérivation par rapport à la variable h donne :

$$f'(x+h) + f'(x-h) = 2f'(x), \text{ c'est-à-dire } f'(x) = \frac{f'(x+h) + f'(x-h)}{2}.$$

Ainsi f' possède la "propriété du point milieu" (l'image du milieu de deux points est le milieu des images de ces deux points). Comme f' est continue (ainsi que l'atteste l'égalité (i)), on peut en déduire que f' est à la fois convexe et concave, c'est-à-dire affine. D'où la propriété voulue : f est un polynôme de degré au plus deux.

VARIATIONS SUR LA DÉRIVATION

Revenons au calcul pratique. Peut-on, en mathématique comme en physique, disposer de valeurs approchées de la dérivée aussi précises que l'on veut ?

Reprenons un calcul simple (même simpliste, puisque la dérivée est ici assez simple à calculer...) : la dérivée de la fonction exponentielle en 0. Si on veut une précision égale à ϵ , il suffit (cf. *supra*) de choisir h tel que $\frac{h^2}{6} M_3 \leq \epsilon$, où représente le maximum de la dérivée troisième de la fonction exponentielle sur $]-h, h[$. Il suffit donc de prendre $h \leq \sqrt{\frac{6\epsilon}{M_3}}$. Si on convient de

prendre h dans $[-0,1 ; 0,1]$, on aura $M_3 = e^{0,1}$. La contrainte sur h devient alors plus simplement : $h \leq \sqrt{6\epsilon} e^{-0,1}$.

Observons les résultats que donne une TI-82 (tableau ci-dessous).

La question posée au début de ce paragraphe était bien sûr faussement naïve. Bien entendu la précision disponible à partir de tout outil de calcul n'est pas infinie. Sur une calculatrice, la précision est de l'ordre de 10^{-12} ; toute exigence supérieure n'a pas de sens. Ainsi, le calcul de nDeriv pour un pas de $h = 2.10^{-15}$, débouche sur un calcul "équivalent" à $\frac{0}{0}$.

Précision voulue ϵ	h théoriquement suffisant $h \leq \sqrt{6\epsilon} e^{-0,1}$	Résultats TI-82	Commentaires
10^{-3}	0,07	<pre>nDeriv(e^X,X,0,.07) 1.000816867</pre>	Le résultat pratique correspond au résultat théorique attendu.
10^{-6}	0,002	<pre>nDeriv(e^X,X,0,.002) 1.000000667</pre>	<i>Idem</i>
10^{-9}	0,0007	<pre>nDeriv(e^X,X,0,.00007) 1.000000001</pre>	<i>Idem</i>
10^{-15}	7.10^{-8}	<pre>nDeriv(e^X,X,0,7*10^-8) Ans-1</pre>	La calculatrice ne détecte plus ici de différence entre valeur approchée et valeur exacte.
10^{-30}	2.10^{-15}	<pre>nDeriv(e^X,X,0,2*10^-15) 0</pre>	Pour un pas h trop petit, le résultat affiché est "délirant".

qui peut donner à peu près n'importe quel résultat.

On peut l'observer aussi sur le plan graphique : fig. 15 et 16, la représentation graphique de

$$f(h) = n\text{Deriv}(e^x, 0, h) = \frac{e^h - e^{-h}}{2h}$$

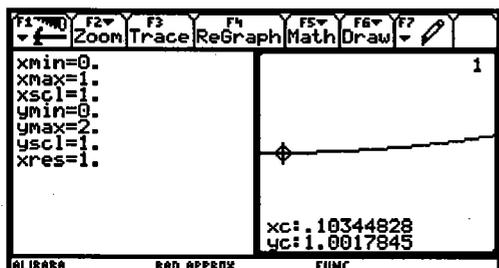


Figure 15

Sur des intervalles "raisonnables", du point de vue de la précision de l'outil, la convergence vers 1 est manifeste.

Dès que l'on approche des limites (de la calculatrice), la recherche de la limite (de la fonction) n'a plus de sens : des oscillations étranges apparaissent.

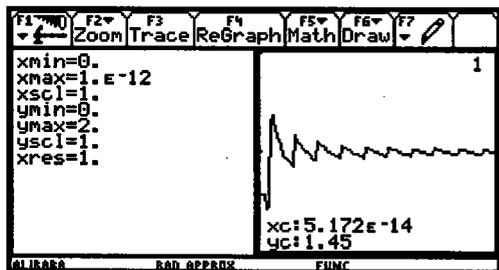


Figure 16

Ces phénomènes sont bien connus ([3], [4]). Ils ont nécessairement des conséquences sur la construction des connaissances ([5]).

Avec un outil de calcul travaillant avec une précision de 10^{-n} , on doit donc se limiter à des pas "h", pour le calcul approché de la dérivée, supérieurs à 10^{-n} .

Revenons au choix de la formule de "dérivée symétrique", concrétisée par les calculatrices, $\frac{f(x+h) - f(x-h)}{2h}$, pour ce calcul

approché. Nous avons vu qu'il était pertinent puisque la précision sur le calcul de dérivée est de l'ordre de h^2 . Mais ne pourrait-on pas envisager d'autres algorithmes qui permettent une précision encore plus grande ? C'est ce que nous allons examiner dans la dernière partie de cet article.

3. Recherche d'un équilibre entre efficacité, précision et validité

Une triple approche de la dérivation numérique (cf. sur ce point [6])

On peut tenter de généraliser ce qui a été fait pour l'étude de la formule de dérivation symétrique. Cette formule reposait sur l'utilisation des points $(x-h)$ et $(x+h)$. On peut utiliser de la même façon $(x-2h)$ et $(x+2h)$. Soit une fonction que l'on supposera dérivable cinq fois. On obtient (formule de Taylor) :

On dispose alors de la formule de Taylor :

$$f(x+h) = f(x) + h.f'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f^{(3)}(x) + \frac{h^4}{24}f^{(4)}(x) + \frac{h^5}{5!}f^{(5)}(c) \text{ où } c \in]x, x+h[$$

$$f(x+h) = f(x) - h.f'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f^{(3)}(x) + \frac{h^4}{24}f^{(4)}(x) - \frac{h^5}{5!}f^{(5)}(c') \text{ où } c' \in]x-h, x[$$

La soustraction de ces deux égalités donne :

VARIATIONS SUR LA DÉRIVATION

$$f(x+h) - f(x-h) = 2hf'(x) + 2 \frac{h^3}{6} f^{(3)}(x) + \frac{h^5}{5!} [f^{(5)}(c) + f^{(5)}(c')] \text{ (égalité 1).}$$

En remplaçant h par 2h, on obtient :

$$f(x+2h) = f(x) + 2h.f'(x) + 4 \frac{h^2}{2} f''(x) + 8 \frac{h^3}{6} f^{(3)}(x) + 16 \frac{h^4}{24} f^{(4)}(x) + 32 \frac{h^5}{5!} f^{(5)}(d)$$

où d ∈]x, x+2h[

$$f(x-2h) = f(x) - 2h.f'(x) + 4 \frac{h^2}{2} f''(x) - 8 \frac{h^3}{6} f^{(3)}(x) + 16 \frac{h^4}{24} f^{(4)}(x) - 32 \frac{h^5}{5!} f^{(5)}(d')$$

où d' ∈]x-2h, x[

La soustraction de ces deux égalités donne :

$$f(x+2h) - f(x-2h) = 4h.f'(x) + 16 \frac{h^3}{6} f^{(3)}(x) + 32 \frac{h^5}{5!} [f^{(5)}(d) + f^{(5)}(d')] \text{ (égalité 2).}$$

Il est clair alors qu'une combinaison linéaire des égalités 1 et 2 permet d'exprimer f'(x) en fonction des différences symétriques :

$$8 [f(x+h) - f(x-h)] - [f(x+2h) - f(x-2h)] = 12h.f'(x) + \frac{h^5}{5!} [12f^{(5)}(c) + 12f^{(5)}(c') - 32f^{(5)}(d) - 32f^{(5)}(d')].$$

D'où :

$$\frac{8[f(x+h) - f(x-h)] - [f(x+2h) - f(x-2h)]}{12h} = f'(x) + \frac{h^4}{5!} \cdot \frac{40M}{12}$$

où |M| ≤ $\frac{\text{Max}}{[x-2h, x+2h]} f^{(5)}$ (égalité 3).

On dispose ainsi d'une approximation qualitativement supérieure à l'approche

symétrique en (x-h), (x+h) précédente. Elle est d'ordre h⁴ désormais.

On peut observer par exemple les résultats comparés de l'approche "symétrique à deux points" et de l'approche "symétrique à quatre points" pour le calcul de la dérivée de la fonction exponentielle en 0 :

Dans la colonne x (fig. 17) se trouvent les différentes valeurs attribuées à h, dans la colonne y1 l'approche de la dérivée par la formule symétrique à deux points, en y2 l'approche par la formule à quatre points.

F1	F2	F3	F4	F5	F6
Setup	Cell	Header	Del	Row	Ins Row
x	y1	y2			
-5	1.0421906	.99785375			
-1	1.0016675	.99999666			
-.01	1.0000167	1.			
0.	undef	undef			
.001	1.0000002	1.			
.01	1.0000167	1.			
.1	1.0016675	.99999666			
1.	1.1752012	.96245819			
x = -.5					
CALC		RAD APPROX		FUNC	

Figure 17

Les deux fonctions semblent bien converger vers 1, mais pas à la même allure !

Cette approche algébrique par la formule de Taylor correspond bien sûr à une stratégie d'interpolation de la fonction f par un polynôme. L'unique polynôme P de degré 4 qui coïncide avec f aux points x, x+h, x-h, x+2h et x-2h vérifiera P'(x) = f'(x) : il suffit de réécrire l'égalité (3) ci-dessus pour le polynôme P et de considérer que bien sûr pour un polynôme du quatrième degré la dérivée cinquième est nulle.

Ainsi la formule d'approche symétrique à quatre points revient à interpoler la

fonction f par l'unique polynôme P du quatrième degré passant par $x, x-h, x+h, x-2h, x+2h$, et à approcher $f'(x)$ par $P'(x)$.

Peut-on imaginer aussi une interprétation barycentrique ? Elle est moins directe que pour la formule de dérivation symétrique à deux points, où on prenait l'isobarycentre des sécantes à gauche et à droite. Mais cette interprétation est aussi possible. Pour simplifier les écritures, on posera :

$$t_1 = \frac{f(x+h) - f(x)}{h}, \quad t_{-1} = \frac{f(x) - f(x-h)}{h},$$

$$t_2 = \frac{f(x+2h) - f(x)}{2h}, \quad t_{-2} = \frac{f(x) - f(x-2h)}{2h}.$$

Le lecteur pourra alors vérifier que l'on a :

$$\frac{8[f(x+h) - f(x-h)] - [f(x+2h) - f(x-2h)]}{12h}$$

$$= \frac{1}{12} [8 t_1 + 8 t_{-1} - 2t_2 - 2t_{-2}]$$

La dérivée symétrique à quatre points est ainsi le barycentre des pentes des cordes $(x, x+h), (x, x-h), (x, x+2h), (x, x-2h)$, affectées des coefficients 8, 8, -2 et -2. Ce résultat fournit au passage une propriété caractéristique des polynômes de degré au plus quatre.

Petite illustration (fig. 18 et 19) avec un polynôme de degré 4, la tangente (horizontale) au point d'abscisse -1 et les quatre cordes passant par ce point et les points d'abscisses respectives -3, -2, 0 et 1 (les deux dernières sont confondues).

Les coefficients de l'écriture barycentrique ne ressortent pas de l'évidence... On peut cependant remarquer que, dans la formule, les cordes les plus proches sont affectées du plus grand coefficient (en

valeur absolue), ce qui semble bien naturel.

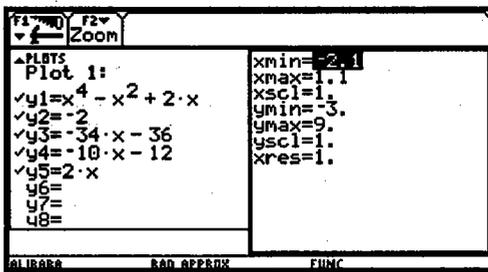


Figure 18

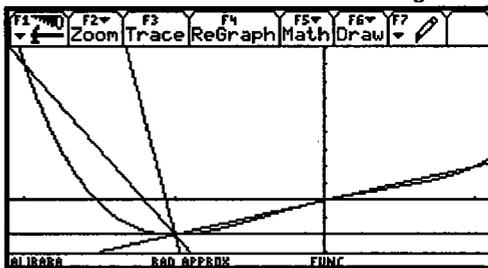


Figure 19

On comprend bien que ce que nous avons fait avec 4 points pourrait être étendu à 2n points. Il conviendrait alors d'écrire la formule de Taylor à l'ordre 2n+1 pour les points $(x+kh)$ et $(x-kh)$, k entier compris entre 1 et n . Une stratégie d'élimination des termes intermédiaires permettrait alors d'obtenir la formule générale :

$$f'(x) = \sum_{k=1}^n a_k \frac{f(x+kh) - f(x-kh)}{h} + o(h^{2n-1})$$

Pour $n = 5$, on trouverait par exemple :

$$a_1 = \frac{5}{6}, \quad a_2 = -\frac{5}{21}, \quad a_3 = \frac{5}{84},$$

$$a_4 = -\frac{5}{504}, \quad a_5 = \frac{1}{1260} \text{ (cf. sur ce point [7]).}$$

Une dernière comparaison (fig. 20) des approches symétriques à 2 points, à 4

VARIATIONS SUR LA DÉRIVATION

points (déjà faites) et à 10 points (à partir des coefficients ci-dessus) est instructive : le gain en rapidité de convergence est considérable.

F1	F2	F3	F4	F5	F6
▼	Setup	Cell	Header	Del Row	Ins Row
x	y1	y2	y3		
-.5	1.0421906	.99785375	1.0000004		
-.1	1.0016675	.99999666	1.		
-.01	1.0000167	1.	1.		
0.	undef	undef	undef		
.001	1.0000002	1.	1.		
.01	1.0000167	1.	1.		
.1	1.0016675	.99999666	1.		
1.	1.1752012	.96245819	1.0005127		
x = -.5					
ALIBABA RAD EXACT FUNC					

Figure 20

Ce qui est excessif est insignifiant

On pourrait alors se poser (encore naïvement) la question : pourquoi ne pas avoir choisi une formule de calcul approché des dérivées encore plus efficace que la formule symétrique à deux points ? On pourrait répondre de façon simpliste que l'échelle de la précision est infinie et qu'il faut bien se situer quelque part... Mais, plus profondément, il faut évoquer ici trois raisons :

- une raison liée à l'instrument : la précision, avec la dérivée symétrique à deux points, est de l'ordre, nous l'avons vu de h^2 . Si la précision de l'instrument est de 10^{-n} , on ne pourra pas choisir un pas de h inférieur à 10^{-n} . On disposera alors, pour le calcul de la valeur approchée de la dérivée via nDeriv, d'une précision théorique de l'ordre de $(10^{-n})^2$, plus grande que ce que l'instrument lui-même peut représenter. Il est donc vain de vouloir aller plus loin. D'autant que le passage à une formule à quatre points se traduit par un alourdissement des calculs (donc une occupation supplémentaire de la mémoire de la calculatrice et un temps de calcul plus long) ;

- une raison mathématique : si on veut contrôler la qualité de l'approximation symétrique avec $2n$ points, cela nous contraint à contrôler la dérivée d'ordre $(2n+1)$. Il y a évidemment un certain paradoxe à devoir contrôler la dérivée $(2n+1)^{ième}$ pour disposer d'une valeur approchée de... la dérivée première ! Ce qui est simple dans le cas de fonctions élémentaires (mais dans ce cas, on pourra disposer souvent des valeurs exactes de la dérivée !) devient beaucoup plus complexe dans le cas général ;

- une raison physico-mathématique : le physicien a souvent des contraintes plus fortes que le mathématicien. Le mathématicien peut faire varier "h" comme il le veut (dans la limite cependant de la précision des outils de calcul), alors que le physicien est soumis en plus à la précision des instruments de mesure... et aux phénomènes eux-mêmes. Si bien qu'il disposera souvent d'une série de mesures données (à intervalle de temps constant par exemple) : il ne pourra plus alors faire varier le pas h, qui sera dès lors *une donnée* du système de mesures. Il s'agira alors plus d'une approche *globale* que *locale* de la fonction.

Si on considère par exemple ces 11 mesures (fig. 21), on peut décider de modéliser le phénomène à partir d'un polynôme qui passe exactement par tous ces points.

On sait qu'il est unique, qu'il est de degré au plus 10, on sait aussi l'exprimer à partir par exemple de la formule de Lagrange.

On obtient alors le polynôme dont le graphe est ci-dessous (fig. 22) : ses variations fortes ne sont pas tout à fait une surprise, vu le degré du polynôme. Il faut imaginer alors que la formule de dérivation numérique à 10 points revient justement à prendre comme dérivée au point central la

dérivée de ce polynôme. On comprend alors que les variations fortes d'un tel polynôme peuvent donner des résultats parfois surprenants...

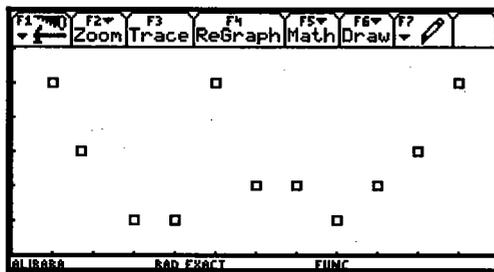


Figure 21

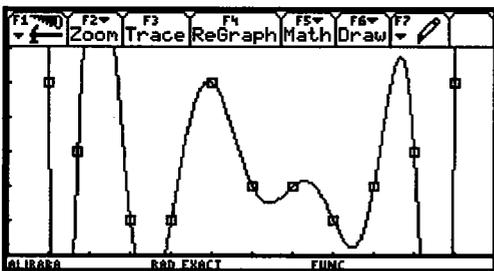


Figure 22

Au point central, la fonction apparaît fortement décroissante

On peut remarquer aussi des "effets de bord important" (phénomènes de Runge, cf. [8]) : l'interpolation fait apparaître aux deux extrémités de l'intervalle des variations encore plus brutales qu'au centre. Le choix de points de part et d'autre du point à approcher limite cet effet : on est pour le calcul de la dérivée au centre de l'intervalle d'interpolation, donc le plus loin des bords possible !

A la place d'une stratégie d'interpolation, on peut choisir une stratégie d'approximation, en recherchant un polynôme de degré donné (ici 4) approchant au mieux les points donnés (par exemple

par la méthode des moindres carrés) (fig. 23). Cela donne une autre vision du phénomène, donnant au cinquième point de la série un caractère d'exception (cela peut aussi révéler une erreur de mesure...).

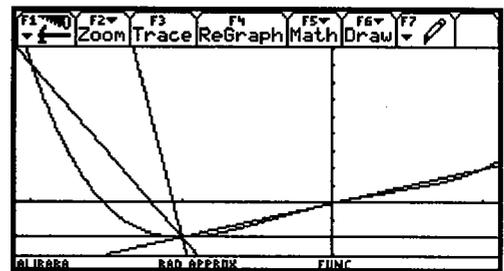


Figure 23

Au point central, la fonction apparaît quasi-constante

Les dérivées qui en découlent sont aussi plus sages.

On peut enfin choisir, entre l'interpolation ou l'approximation globales, une méthode intermédiaire : analyser les mesures par petits paquets, pour en gérer à chaque fois un petit nombre. Si on en prend 3 par exemple, on peut interpoler par un polynôme de degré au plus deux, et on débouche, pour le calcul de la dérivée, sur la formule des physiciens et des calculatrices :

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$

Retour au point de départ ! On peut aussi constater que les paraboles, qui sont derrière cette formule, sont aussi derrière la formule de calcul approché des intégrales par la méthode de Simpson (où l'on retrouve aussi un caractère barycentrique).

Le caractère édifiant de la parabole est "incontournable"...

 VARIATIONS SUR LA DÉRIVATION

La morale de la parabole

La généralisation de calculatrices graphiques faisant du calcul approché avait, curieusement, relégué au second plan le calcul numérique dans le cours de mathématiques. Au professeur la tâche de calculer "exactement", à l'élève la charge de comprendre les réponses approximatives des étranges lucarnes.

Les choses changent, avec une évolution lente des programmes. Elles changent aussi du fait d'une évolution des matériels eux-mêmes, en particulier avec l'irruption des calculatrices symboliques :

– la coexistence, sur ces mêmes calculatrices, du mode "calcul exact" et du mode "calcul approché" oblige les élèves à considérer les différents résultats, à passer d'un mode à l'autre ;

– les écritures symboliques peuvent dévoiler, on l'a vu, le contenu de certaines commandes comme nDeriv. La boîte noire s'éclaircit un peu.

Il y a là, pour le professeur de mathématiques, un nouveau champ d'activité qui s'ouvre. Dans ces activités, il posera nécessairement le problème du contrôle des calculs approchés, de l'efficacité, de la validité, du sens des algorithmes de calcul mis en place. Il devra préciser le sens des approches locales et globales d'une fonction. Il pourra faire interagir des cadres géométriques (les barycentres) et analytiques.

Il trouvera aussi matière à discussion – et pourquoi pas à travail interdisciplinaire ? avec le professeur de physique !

BIBLIOGRAPHIE

Merci à Henri Lombardi qui m'a transmis des notes de cours fort utiles pour la mise en forme de cet article.

- [1] RAMIS, DESCHAMPS, ODOUX, 1985 : *Analyse 2*, pp 145-146, Paris : Masson.
- [2] Bernard RENÉ, Christian FAURE, Maryse NOGUÈS, Yvon NOUAZÉ et Luc TROUCHE, 1995 : *Des fonctions et des graphes*, Montpellier : Irem, Université de Montpellier II.
- [3] KUNTZ Gérard, 1993 : "L'outil informatique ne peut donner que ce qu'il a", *Repères-Irem* n°11, pp. 39-55, Pont-à-Mousson : Topiques Éditions.
- [4] TROUCHE Luc, 1994 : "Calculatrices graphiques, la grande illusion", *Repères-Irem* n°14, pp. 5-32. Pont-à-Mousson : Topiques Éditions.
- [5] TROUCHE Luc, 1996 : *A propos de l'apprentissage de la notion de limites dans un environnement calculatrice, étude des rapports entre processus d'instrumentation et processus de conceptualisation*, thèse de doctorat, Montpellier : Irem, Université de Montpellier II.
- [6] SIBONY Moïse et Jean-Claude MARDON, 1988 : *Analyse numérique II, Approximations et équations différentielles*, Paris : Hermann.
- [7] LYGEROS Nyk, Olivier MARGUIN et Michel MIZONY, 1999 : "Réflexions méthodologiques en calcul formel", *Repères-Irem*, n° 34, Pont-à-Mousson : Topiques Éditions.
- [8] OVAERT Jean-Louis et Jean-Luc VERLEY, 1981 : *Algèbre 1*, pp. 101-102. Paris : Cedic/Fernand Nathan.