Mathématique et Pédagogie

Périodique bimestriel publié par la *Société Belge des Professeurs de Mathématique d'expression française*

Sommaire

N° 106 ***Mars – Avril 1996***

Éditorial

C. rédaction,

Motion

Le Conseil d'Administration de la SBPMef (Société Belge des Professeurs de Mathématique d'expression française), réuni ce jeudi 29 février 1996, a examiné la situation créée par la diminution du NTPP dans l'enseignement secondaire.

Il rappelle que la SBPMef situe ses activités exclusivement sur le terrain pédagogique. Ce sont donc l'efficacité de l'enseignement et la qualité de la formation dispensée aux élèves qui sont au centre de ses préoccupations.

Après les mesures précédentes qui ont réduit de manière importante les grilles-horaires des élèves, les dernières mesures proposées par le gouvernement de la Communauté Française vont avoir des effets néfastes sur l'offre des formations et les possibilités de remédiation.

Ainsi, la réduction brutale du nombre de périodes organisables conduira forcément à la disparition irrationnelle de sections techniques et professionnelles dans de vastes aires géographiques. Elle entraînera la disparition de sections réputées fortes aussi bien dans les cours de langues anciennes et modernes qu'en mathématique et en sciences.

A part certaines écoles élitistes qui concentreront les élèves les plus doués, les autres écoles se verront contraintes de ne proposer que des filières dont la fonction sera de regrouper le maximum d'élèves au détriment de véritables objectifs de formation.

Ces mesures qui modifient radicalement le paysage de l'enseignement vont accélérer la dualisation de la société.

Conscient de la responsabilité des enseignants vis-à-vis des jeunes, le Conseil d'Administration de la SBPMef refuse de cautionner ce processus. Il désire exprimer publiquement son inquiétude devant la mise en application de mesures dont les conséquences possibles sont très dommageables pour les générations futures.

Cette motion a été adoptée à l'unanimité des membres présents.

Rectificatif : Dans le n° 105 de la revue “Mathématique et Pédagogie” a été publié un texte intitulé “Suggestions concernant l'enseignement de la mathématique aux deuxième et troisième degrés de l'enseignement secondaire” (pp. 5–28). Contrairement à ce que l'on pourrait croire à la lecture de cet article, ce texte résulte d'un travail réellement collectif de la Commission pédagogique, travail dans lequel le rôle du Président n'a pas été le plus important (seule l'introduction est vraiment de lui). Ce texte a été amendé et approuvé par le C.A. de la SBPMef.

Plusieurs chemins mènent au produit scalaire

E. Jason,

Les élèves des écoles secondaires seraient désarçonnés s'ils voyaient introduire la géométrie euclidienne en "parachutant" un texte tel que celui-ci :

> "Axiome untel.
>
> L'espace vectoriel des vecteurs du plan est muni d'une forme bilinéaire, symétrique, définie positive, qui à chaque couple de vecteurs fait correspondre un réel appelé ≪produit scalaire de ces 2 vecteurs≫ (...)".
>
> (un tel espace vectoriel est appelé euclidien).

On remédie souvent à cette difficulté en construisant le produit scalaire de 2 vecteurs à l'aide de projections orthogonales et de mesures algébriques (ou de composantes), ce qui nécessite l'introduction préalable de la perpendicularité des droites, elle-même introduite à partir de la notion d'angle ou après des manipulations telles que des doubles pliages...

La commutativité du produit scalaire exige un nouvel axiome, parfois motivé par des manipulations telles que le pliage autour de "la bissectrice d'un angle", etc.

Les difficultés des élèves proviennent de leur manque de maîtrise des projections orthogonales, de notations propices à la confusion ($\overline{OA'}$ et $\overline{OA}$, etc), de la distinction pas toujours claire entre norme et mesure algébrique, du nombre élevé de notions à introduire ou à rappeler pour arriver au produit scalaire.

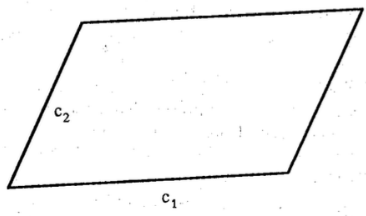
Le but de cet article est de présenter une approche plus directe.

Au terme de cette approche, la géométrie euclidienne est entièrement fondée sur la notion de distance, plus simple que celle d'angle ou de forme (classe de figures semblables).

1. Découverte

Les élèves reçoivent chacun un parallélogramme en carton, ou une feuille de papier sur laquelle sont dessinés des parallélogrammes.

Ils sont invités à vérifier, avec la précision de leurs instruments de dessin, qu'il s'agit bien de parallèlogrammes et à mesurer la longueur des côtés et des diagonales. Les parallélogrammes étant "différents", ils obtiennent des résultats tous différents qui sont centralisés au tableau.



Pour faciliter la collecte des résultats, lors de cette première étape, les longueurs à mesurer ont été choisies de sorte qu'elles comportent chacune un nombre entier de centimètres.

Les élèves reçoivent donc comme consigne d'arrondir éventuellement au centimètre près et de ne pas tenir compte des millimètres (ni d'éventuelles imperfections d'égalité de longueur des côtés opposés). Il obtiennent ainsi des résultats tels que ceux-ci (dans ce tableau, chaque ligne représente un parallélogramme) :

longueur des côtés (en cm)		longueur des diagonales (en cm)	
4	7	7	9
6	7	7	11
5	5	6	8
7	9	8	14
10	11	9	19
11	12	13	19
11	13	16	18
6	13	11	17
6	8	10	10
8	11	9	17
6	17	17	19
10	15	17	19
7	11	12	14
5	10	9	13
8	9	11	13
...	...	...	...
...	...	...	...

Ensuite, le professeur leur affirme que les nombres de chaque ligne ne sont pas indépendants : il existe une relation entre eux.

Bien sûr, les inégalités triangulaires sont vérifiées : $4+7>7$, $4+7>9$, etc. Mais en outre, il existe une relation qui s'exprime par une certaine *égalité.*

Les élèves sont invités à découvrir cette égalité. Ils disposent du temps nécessaire. Chaque élève est invité à travailler sur "son" parallélogramme.

Bien sûr, dans ce cas, il peut découvrir une relation qui n'est pas générale.

Par exemple, pour le dernier de la liste ci-dessus :

$$13-11=2\times(9-8)$$

ce qui, généralisé, donnerait

$$d_2-d_1=2\times(c_2-c_1),$$

où c_1 et c_2 sont les longueurs des côtés et où d_1 et d_2 sont les longueurs des diagonales ; mais cette égalité est mise en défaut sur tous les autres exemples cités.

D'autres proposeront $c_2=d_1, \quad c_2=\frac{1}{2}(d_1+d_2)$, etc.

Le professeur peut les mettre sur la bonne voie en leur confirmant qu'il faut chercher une égalité entre une fonction des côtés et une fonction des diagonales.

Comme on peut changer le nom des diagonales, il n'y a aucune raison de privilégier l'une des 2 diagonales ; de même pour les longueurs des côtés (certains verront peut-être ici l'ébauche d'une exigence de relativité...).

Donc, dans la formule cherchée, les 2 côtés jouent le même rôle et les 2 diagonales jouent le même rôle.

La relation cherchée, si elle est vraie, doit être indépendante de l'unité de longueur choisie. Ceci oriente les recherches vers une égalité entre polynômes de même degré.

Les polynômes en c_1 et c_2, qui sont du premier degré, et où c_1 et c_2 jouent le même rôle, sont tous des multiples de c_1+c_2. (On parlera d'opérations commutatives données par des polynômes du premier degré, s'ils n'ont pas vu les polynômes symétriques homogènes...).

Ainsi, on est amené à comparer $c_1 + c_2$ et $d_1 + d_2$:

c_1	c_2	d_1	d_2	$c_1 + c_2$	$d_1 + d_2$
4	7	7	9	11	16
6	7	7	11	13	18
5	5	6	8	10	14
7	9	8	14	16	22
6	8	10	10	14	20
8	11	9	17	19	26
5	10	9	13	15	22
8	9	11	13	17	23

A ce stade, on ne voit aucune relation apparaître entre les deux dernières colonnes.

A la question “Après les polynômes du premier degré, quelles sont les fonctions les plus simples ?” les élèves répondent d'habitude “ceux du second degré” (très peu répondent “on pourrait essayer la somme des inverses” ou la valeur absolue de la différence...).

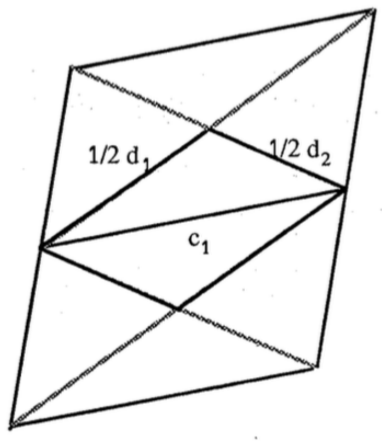
Les polynômes en c_1 et c_2, qui sont du second degré, et où c_1 et c_2 jouent le même rôle, sont c_1c_2, $c_1^2 + c_2^2$, et les sommes de leurs multiples.

On est donc amené à comparer les dernières colonnes de tableaux tels que :

c_1	c_2	d_1	d_2	$c_1.c_2$	$c_1^2 + c_2^2$	...	$d_1.d_2$
4	7	7	9	28	65		63
6	7	7	11	42	85		77
5	5	6	8	25	50		48
7	9	8	14	63	130		112
6	8	10	10	48	100		100
8	11	9	17	88	185		153
5	10	9	13	50	125		117
8	9	11	13	72	145		143

A ce stade, après avoir cherché suffisamment, une classe sur deux environ trouve la relation demandée (voir plus bas).

Dans le cas contraire, on peut faire la remarque suivante : si on “empile” deux exemplaires du même parallélogramme (en pratique, on trace sur le papier le contour du parallélogramme en carton, on translate celui-ci et on recommence) et si on trace leurs diagonales, par exemple dans une autre couleur :



on obtient un parallélogramme de côtés $\frac{1}{2}d_1$ et $\frac{1}{2}d_2$ et de diagonales c_1 et c_2 : les côtés peuvent jouer le rôle de diagonales, et à un facteur $\frac{1}{2}$ près, les diagonales peuvent jouer le rôle de côtés. Donc si on cherche une relation entre 2 fonctions $f(c_1, c_2)$ et $g(d_1, d_2)$, ce serait une bonne idée de chercher cette relation entre $f(c_1, c_2)$ et $f(d_1, d_2)$ pour la *même* fonction f. On compare donc c_1c_2 avec d_1d_2, puis séparément $c_1^2 + c_2^2$ avec $d_1^2 + d_2^2$:

c_1	c_2	d_1	d_2	$c_1.c_2$	$d_1.d_2$	$c_1^2 + c_2^2$	$d_1^2 + d_2^2$
4	7	7	9	28	63	65	130
6	7	7	11	42	77	85	170
5	5	6	8	25	48	50	100
7	9	8	14	63	112	130	260
6	8	10	10	48	100	100	200
8	11	9	17	88	153	185	370
5	10	9	13	50	117	125	250
8	9	11	13	72	143	145	290

A cette étape, toutes les classes remarquent la proportionnalité des deux dernières colonnes. A mon avis, le temps passé à ce type de recherche n'est jamais perdu. Il développe l'esprit d'observation, l'imagination, le sens géométrique, l'émulation, la réflexion sur la nature et la symétrie, et bien sûr le calcul mental.

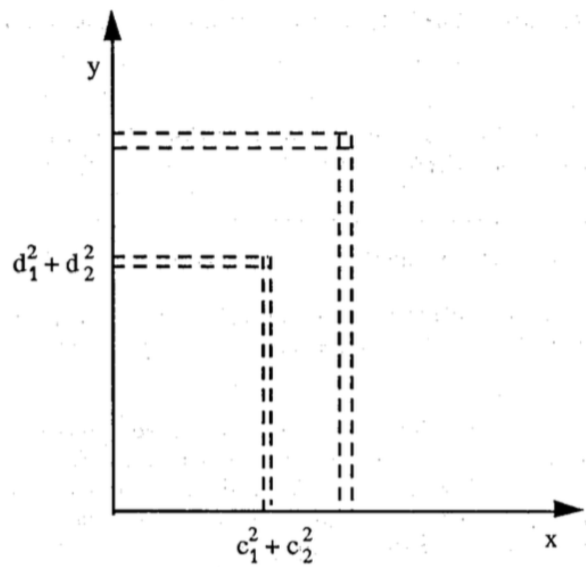
Finalement, ils constatent donc que

$$2(c_1^2 + c_2^2) = d_1^2 + d_2^2 :$$

pour chacun de leurs parallélogrammes, la somme des carrés de leurs 4 côtés égale la somme des carrés des 2 diagonales. (C'est l'occasion de leur faire exprimer leur trouvaille par une phrase complète.)

Mais ce n'est pas fini. Est-ce vrai pour *tous* les parallélogrammes ?

Que chaque élève trace arbitrairement un nouveau parallélogramme, en remontant à la définition première : deux droites parallèles distinctes qui coupent (sur la feuille) deux droites parallèles distinctes. Qu'il mesure les côtés et les diagonales, aux erreurs de mesure près (une occasion d'utiliser les calculatrices et d'appliquer le calcul des incertitudes).



La droite $y = 2x$ passe-t-elle dans toutes les “boîtes d'erreur” ? Sinon, vérifier le parallélisme des côtés, etc.

Après cet effort, ils seront sans doute bien disposés à admettre comme *axiome*, c'est-à-dire comme point de départ de déductions futures :

> *dans tout parallélogramme, la somme des carrés des diagonales vaut deux fois la somme des carrés de deux côtés adjacents.*

Cet axiome sera appelé par la suite *axiome des parallélogrammes.*

2. Extension à 2 vecteurs quelconques du plan

Avant de continuer, précisons quand cette approche peut avoir lieu.

Les élèves doivent bien connaître la géométrie affine plane, qui permet de définir la situation "parallélogramme" (voir ci-dessus) ; cette définition peut être vue après les deux premiers axiomes

"*par 2 points distincts du plan passe une et une seule droite*"

et

"*par un point extérieur à une droite, il passe dans le plan une et une seule parallèle à cette droite*"

et leurs conséquences directes.

Par la suite, il est utile d'avoir les conséquences de :

Si 2 bipoints sont liés à un même troisième par des parallélogrammes, ils sont liés entre eux par un parallélogramme, ou sont alignés

ce qui permet d'étudier les translations du plan (les vecteurs de la géométrie plane) ; il est aussi utile d'avoir introduit les nombres réels et l'axiome de bijection entre toute droite et l'ensemble $\mathbb{R}$ des nombres réels (une fois fixé un repère de cette droite) ; cet axiome confère à toute droite les propriétés de $\mathbb{R}$ (y compris la densité de l'ensemble des rationnels dans $\mathbb{R}$).

Il est utile d'avoir débuté la géométrie métrique, en ayant défini la norme $\|\vec{v}\|$ d'un vecteur $\vec{v}$, soit (ce qui est équivalent) en ayant postulé qu'il existe une fonction "distance" entre les points du plan (ou une fonction "longueur" sur les segments), qui vérifie les définitions d'une distance, qui donne la même longueur à deux côtés opposés d'un parallélogramme et qui vérifie "Thalès pour les longueurs" :

$$(\forall r \in \mathbb{R}) \;\; (\forall \vec{v} \text{ vecteur}) \qquad \|r.\vec{v}\| = |r| \,.\, \|\vec{v}\| \,.$$

Pour $\vec{u}, \vec{v}$ fixés, on supposera que la fonction $\mathbb{R}$ dans $\mathbb{R}$: $r \mapsto \|r\vec{u} + \vec{v}\|$ est continue. En 3ème et 4ème, il faudra se contenter d'une explication peu rigoureuse, quitte à y revenir plus tard.

Après avoir rappelé ces prérequis, nous pouvons maintenant exploiter comme axiome l'égalité qui a été découverte.

Autre formulation de l'axiome des parallélogrammes :

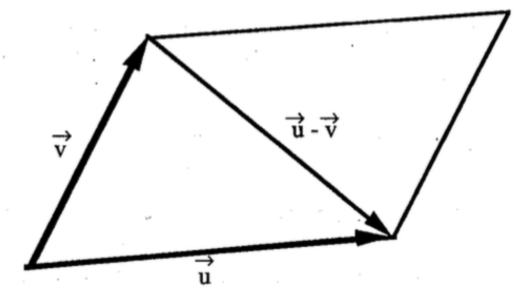
Pour tous vecteurs $\vec{u}$ et $\vec{v}$ du plan :

$$\|\vec{u}+\vec{v}\|^2 + \|\vec{u}-\vec{v}\|^2 = 2(\|\vec{u}\|^2 + \|\vec{v}\|^2)$$

Preuve : Si l'un des deux vecteurs est nul, c'est évident.

S'ils sont nuls et parallèles, il existe $r \times \mathbb{R}$ tel que $\vec{u} = r.\vec{v}$ et la formule annoncée se démontre à l'aide d'une propriété de la norme ("Thalès pour les longueurs").

Si les deux vecteurs $\vec{u}$ et $\vec{v}$ sont nuls et non parallèles, ils ont comme représentants des bipoints formant un parallélogramme, dont les longueurs des côtés sont $\|\vec{u}\|$ et $\|\vec{v}\|$ et les diagonales sont $\|\vec{u}+\vec{v}\|$ et $\|\vec{u}-\vec{v}\|$; l'axiome des parallélogrammes donne la formule annoncée.

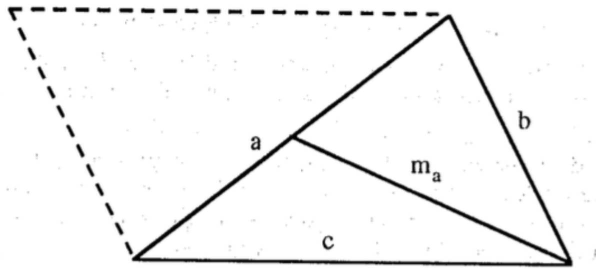


Encore une autre formulation de l'axiome des parallélogrammes : le *théorème de la médiane* :

Dans tout triangle, la somme des carrés des deux côtés vaut deux fois la somme du carré de la moitié du troisième côté et du carré de la médiane relative à ce côté.

Avec les notations de la figure ci-dessous,

$$b^2 + c^2 = 2[(\frac{1}{2}a)^2 + m_a^2]$$



Preuve : construire le symétrique de cette médiane par rapport au milieu de ce troisième côté, reconnaître un parallélogramme, et appliquer l'axiome des parallélogrammes : on obtient

$$2(b^2 + c^2) = a^2 + (2m_a)^2.$$

3. Le produit scalaire

A ce moment du cours, il est opportun de faire le bilan des opérations déjà définies et mettant en jeu les vecteurs.

On sait : additionner deux vecteurs
soustraire deux vecteurs
multiplier un vecteur par un nombre réel
diviser un vecteur par un nombre réel non nul
trouver la norme d'un vecteur
(ce n'est pas une opération binaire, ni interne...)

L'étape suivante serait logiquement de chercher à multiplier entre eux deux vecteurs. Mais, pour convenir, cette nouvelle opération doit être liée à la norme des vecteurs. Une relation naturelle à imposer est que si l'on multiplie un vecteur par lui-même, on retrouve soit sa norme, soit le carré de sa norme.

On préférera ici le carré de sa norme, pour les mêmes raisons d'invariance lors de changements d'unités qu'au paragraphe précédent.

On impose donc à la nouvelle multiplication : pour tout vecteur $\vec{v}$ du plan :

$$\vec{v}.\vec{v} = \|\vec{v}\|^2$$

Il reste à définir $\vec{u}.\vec{v}$.

C'est le moment de rappeler une petite expérience amusante de calcul rapide.

Les fanatiques de calcul mental remplacent parfois les multiplications entre nombres par des soustractions, plus rapides, à condition de connaître les carrés. Alors, ils utilisent

$$(\forall a, b \in \mathbb{R}) \qquad a.b = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2$$

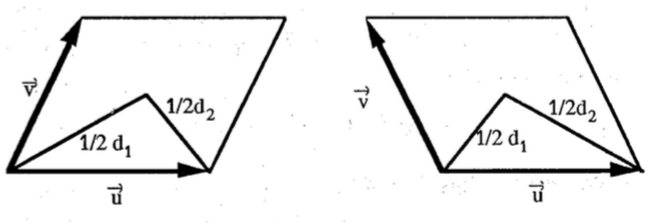
Exemple :

$$27 \times 23 = 25^2 - 2^2 = 625 - 4 = 621.$$

Transposons ceci aux vecteurs : la norme va permettre de définir le *produit scalaire* :

$$\vec{u}.\vec{v} = \frac{1}{4}\left(\|\vec{u}+\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2\right)$$

Cette formule, appliquée à tout vecteur $\vec{u}$ et à tout vecteur $\vec{v}$, définit une multiplication entre vecteurs, à valeurs dans $\mathbb{R}$.



$$\vec{u}.\vec{v} = \left(\frac{1}{2}\,d_1\right)^2 - \left(\frac{1}{2}\,d_2\right)^2 > 0 \qquad\qquad \vec{u}.\vec{v} = \left(\frac{1}{2}\,d_1\right)^2 - \left(\frac{1}{2}\,d_2\right)^2 < 0$$

4. Bifurcation

Ici, deux chemins sont possibles.

Soit on définit la perpendicularité de deux vecteurs non nuls par la nullité de leur produit scalaire et on constate que l'axiome des parallélogrammes devient, dans ce cas, le théorème de Pythagore, soit on étudie d'abord les propriétés du produit scalaire.

Le premier chemin conduit directement au théorème de Pythagore, mais ne permet pas d'aller plus loin dans les propriétés métriques des triangles rectangles. Il ne permet même pas de parler de *directions* perpendiculaires, en l'absence de la propriété :

$$\forall \vec{u}, \vec{v} \text{ vecteurs, } \forall r \in \mathbb{R}, \qquad (r\vec{u}) \cdot \vec{v} = r(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (r\vec{v}).$$

C'est pourquoi il semble préférable d'étudier d'abord les propriétés fondamentales du produit scalaire, ce que fait le théorème suivant.

5. Théorème de Jordan et von Neumann (références au §9)

La mutiplication définie par le produit scalaire ci-dessus est :

(1) *bilinéaire :* $\forall \vec{u}, \vec{v}, \vec{w}$ *vecteurs :*

$$(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$$

et

$$\vec{w} \cdot (\vec{u} + \vec{v}) = \vec{w} \cdot \vec{u} + \vec{w} \cdot \vec{v} \quad \textit{(additivité)}$$

$\forall \vec{u}, \vec{v}$ *vecteurs,* $\forall r \in \mathbb{R}$ *:*

$$(r\vec{u}) \cdot \vec{v} = r(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (r\vec{v})$$

(2) *symétrique :* $\forall \vec{u}, \vec{v}$ *vecteurs*

$$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$$

(3) *définie positive :* $\forall \vec{u}$ *vecteur :*

$$\vec{u} \cdot \vec{u} \geqslant 0 \textit{ et } \vec{u} \cdot \vec{u} \textit{ n'est nul que si } \vec{u} = \vec{0}.$$

(4) *liée à la norme par :* $\forall \vec{u}$ *vecteur :*

$$\vec{u} \cdot \vec{u} = \|\vec{u}\|^2$$

Preuve : (2), (4) et (3) sont évidents, vu les propriétés de la norme.

Montrons (1).

Soustrayons les lignes de rang pair des lignes de rang impair parmi les conséquences suivantes de l'axiome des parallélogrammes :

$$\begin{aligned}
\|(\vec{u}+\vec{w})+\vec{v}\|^2+\|\vec{u}+\vec{w}-\vec{v}\|^2 &= 2\cdot\|\vec{u}+\vec{w}\|^2+2\cdot\|\vec{v}\|^2 \\
\|(\vec{u}-\vec{w})+\vec{v}\|^2+\|\vec{u}-\vec{w}-\vec{v}\|^2 &= 2\cdot\|\vec{u}-\vec{w}\|^2+2\cdot\|\vec{v}\|^2 \\
\|\vec{u}+(\vec{v}+\vec{w})\|^2+\|\vec{u}-\vec{v}-\vec{w}\|^2 &= 2\cdot\|\vec{u}\|^2+2\cdot\|\vec{v}+\vec{w}\|^2 \\
\|\vec{u}+(\vec{v}-\vec{w})\|^2+\|\vec{u}-\vec{v}+\vec{w}\|^2 &= 2\cdot\|\vec{u}\|^2+2\cdot\|\vec{v}-\vec{w}\|^2
\end{aligned}$$

il reste :

$$8(\vec{u}+\vec{v})\cdot\vec{w}=8\vec{u}\cdot\vec{w}+8\vec{v}\cdot\vec{w}$$

d'où, en utilisant (2), l'additivité.

La dernière thèse $(\forall r\in\mathbb{R}),\quad (r\vec{u})\cdot\vec{v}=r(\vec{u}\cdot\vec{v})$ est plus "résistante".

Si r est un nombre naturel n, elle découle de l'additivité. En effet (si les deux sommes qui suivent et qui contiennent "…" ont n termes)É :

$$(n\vec{u})\cdot\vec{v}=(\vec{u}+\vec{u}+\cdots+\vec{u})\cdot\vec{v}=\vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{v}+\cdots+\vec{u}\cdot\vec{v}=n\cdot(\vec{u}\cdot\vec{v}).$$

Si $r=0$, elle revient à $\vec{0}\cdot\vec{v}=0$, qui est vérifié en appliquant la définition.

Si $r=-1$, elle revient à $(-\vec{u})\cdot\vec{v}=-\vec{u}\cdot\vec{v}$, donc à $(-\vec{u})\cdot\vec{v}+\vec{u}\cdot\vec{v}=0$, qui est vérifié par additivité.

Si $r\in\mathbb{Z}\setminus\mathbb{N}$,

$$(r\vec{u})\cdot\vec{v}=(-|r|\cdot\vec{u})\cdot\vec{v}=-(|r|\,\vec{u})\cdot\vec{v}=-|r|\,(\vec{u}\cdot\vec{v})=r(\vec{u}\cdot\vec{v}).$$

Si $r\in\mathbb{Q}$, $r=\frac{n}{d}$ avec $n\in\mathbb{Z}$, $d\in\mathbb{N}$. Par ce qui précède, $(n\vec{u})\cdot\vec{v}=n(\vec{u}\cdot\vec{v})$.

Divisons les deux membres par d :

$$\frac{1}{d}\,(n\vec{u})\cdot\vec{v}=\left(\frac{n}{d}\,\vec{u}\right)\cdot\vec{v}$$

car leur produit par d vaut chaque fois $n(\vec{u}\cdot\vec{v})$. Or,

$$\frac{1}{d}\,(n\vec{u})\cdot\vec{v}=\left(\frac{1}{d}\,n\right)(\vec{u}\cdot\vec{v})=r(\vec{u}\cdot\vec{v}).$$

Enfin, si $r \in \mathbb{R} \setminus \mathbb{Q}$, ayant admis que $(r\vec{u}) \cdot \vec{v}$ et $r(\vec{u} \cdot \vec{v})$ sont des fonctions continues de r, égales sur l'ensemble $\mathbb{Q}$ des nombres rationnels, on admettra qu'elles sont égales sur $\mathbb{R}$ (en 4ème, on devra promettre une explication en 5ème).

L'axiome des parallélogrammes implique donc la linéarité du produit scalaire qui, avec la positivité, implique l'inégalité de Cauchy-Schwarz et l'inégalité triangulaire ; on peut donc enlever celle-ci des axiomes de distance signalés au début de ce paragraphe, si on impose l'axiome des parallélogrammes.

6. Droites perpendiculaires

Deux vecteurs sont dits *orthogonaux* si leur produit scalaire est nul.

Deux vecteurs $\vec{u}$ et $\vec{v}$ non nuls sont dits *perpendiculaires* si leur produit scalaire est nul, ce que l'on note $\vec{u} \perp \vec{v}$.

La perpendicularité de deux vecteurs non nuls ne dépend que de leurs directions : en effet, si $\vec{u} \perp \vec{v}$, pour tout r non nul, $(r\vec{u}) \cdot \vec{v} = r(\vec{u} \cdot \vec{v}) = r \cdot 0 = 0$, donc $r\vec{u} \perp \vec{v}$.

On peut donc parler de directions perpendiculaires si ce sont celles de vecteurs non nuls perpendiculaires. A des directions perpendiculaires correspondent des droites perpendiculaires.

L'étude de la géométrie métrique plane peut alors se poursuivre comme dans les approches traditionnelles :

— par le théorème de Pythagore, s'il n'a pas encore été vu (cf. §4)

— par l'existence et l'unicité de la perpendiculaire menée par un point X à une droite a.
(Soit $\vec{v}$ un vecteur parallèle à la droite a et de norme 1.
Pour tout point $X \notin a$ fixé et tout point Y de a, $\overrightarrow{XY} - (\overrightarrow{XY} \cdot \vec{v})\vec{v}$ est perpendiculaire à a ; l'unicité découle du théorème de Pythagore ou du théorème de la médiane).

— par la propriété d'invariance : "un produit scalaire ne change pas si on ajoute à l'un des facteurs un vecteur orthogonal à l'autre", ce qui permet de démontrer très rapidement les propriétés métriques restantes des triangles rectangles.

— par l'étude des projections orthogonales, des bases orthogonales et orthonormées,...
— par les implications : ($\forall a, b, c$, droites du plan) :
$a \perp b$ et $b // c$ impliquent $a \perp c$
$a // b$ et $b \perp c$ impliquent $a \perp c$
$a \perp b$ et $b \perp c$ impliquent $a // c$.
(la dernière implication aurait été fausse en géométrie tri-dimensionnelle).
— par la définition et les propriétés du cosinus de deux vecteurs, y compris les formules des cosinus dans un triangle quelconque ABC :

$$a^2 = b^2 + c^2 - 2bc \cos A, \quad \dots$$

— par l'étude des isométries et des orientations du plan, des angles et des fonctions trigonométriques.

Au terme de cette approche, le produit scalaire n'est plus fondé sur les projections orthogonales, mais permet de les définir.

Ainsi, les élèves qui éprouvent des difficultés à manipuler les projections orthogonales (par exemple, s'ils ne maîtrisent pas le concept d'application du plan sur une droite) ne sont pas exclus de la compréhension du produit scalaire, ni des parties de la géométrie qui l'utilisent.

7. Observations lors de l'utilisation de cette méthode

Cette méthode a été utilisée comme première approche du produit scalaire en 4ème secondaire (horaire 4h/sem.) de 1983 à 1985 et dans l'enseignement supérieur pédagogique comme alternative aux méthodes traditionnelles, de 1985 à 1990.

Nous nous limiterons à des remarques sur les classes de 4ème secondaire.

La part importante consacrée au travail de découverte (voir §1) réserve cette approche à des classes imaginatives et "chercheuses".

Il est inutile de "vendre la mèche" et de servir "à froid" la réponse $2(c_1^2 + c_2^2) = d_1^2 + d_2^2$ à une classe qui n'a pas cherché suffisamment ou qui n'a pas bien compris ce qu'on lui demande. Les doubleurs doivent se taire.

Cette méthode permet d'arriver très rapidement au coeur de la géométrie métrique.

Il est inutile d'introduire des axiomes spécifiques pour l'existence de la direction perpendiculaire à une direction donnée, pour la commutativité du produit scalaire, etc. (Chaque axiome devant être motivé, devant être fondé sur une approche visant *primo* à lui donner un *sens* perceptible par les élèves, *secundo* à l'intégrer dans la rationalité de l'enseignement, les méthodes trop exigeantes en axiomes sont souvent coûteuses en temps ou déficientes au niveau du sens).

Elle bénéficie de l'attrait que présente toute méthode comportant des séquences de découverte, pendant lesquelles les élèves jouissent d'une assez grande liberté de recherche.

Comme dans les autres méthodes d'exposition du produit scalaire, il faut écrire explicitement les formules qui permettent de calculer le produit scalaire de deux vecteurs parallèles (de même sens ou de sens contraires) à partir de leurs normes.

La démonstration de la bilinéarité dans le théorème de Jordan-von Neumann n'est pas évidente. Il ne faut pas s'attendre à ce qu'un élève la retrouve par lui-même. (Pour preuve, le délai qui s'est écoulé entre la découverte du produit scalaire – 1843 – et la découverte de cette démonstration : 1935). Mais il n'est peut-être pas mauvais que les élèves rencontrent de temps à autre une démonstration assez "technique" : un autre exemple est la démonstration de la formule de Taylor-Maclaurin en 6ème secondaire. L'important est ici de bien déterminer l'objectif recherché : ce n'est pas de reproduire la démonstration (!) mais d'être convaincu de sa validité.

Ceci soulève à nouveau la question de l'utilité des démonstrations à l'école secondaire.

A mon avis, il ne faut présenter une démonstration que si les élèves en ressentent le besoin. Ils expriment ce besoin par des questions du genre "D'où vient cet énoncé ?" (Qu'est-ce qui me le prouve ?) ou bien "Est-ce toujours vrai ?" (Où interviennent les hypothèses ?) ou "A quoi sert (cette définition. . .) ?".

L'argument utilisant la continuité ne pose pas de problèmes aux élèves et fait partie d'un petit nombre d'arguments topologiques inévitables en géométrie. Pour rappel, la géométrie traditionnelle "justifie" l'existence de la perpendiculaire par un raisonnement cinématique ou d'ordre : si les deux angles formés par une demi-droite ou une droite sont inégaux, on fait tourner la demi-droite jusqu'à ce que l'inégalité change de sens. Implicitement, on applique le théorème des valeurs intermédiaires pour conclure qu'entre ces deux positions, il existe une demi-droite qui forme des angles égaux.

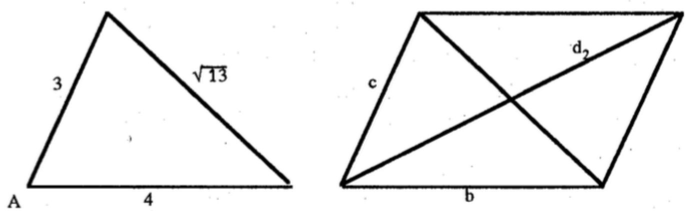
Parmi les (rares) effets secondaires de cette méthode, j'ai pu observer qu'au moment de calculer le cosinus d'un angle d'un triangle dont on donne les trois côtés, certains élèves remontent jusqu'à la définition du produit scalaire plutôt que d'utiliser la formule $a^2 = b^2 + c^2 - 2bc \cos A$.

C'est ainsi que pour calculer l'angle A ci-dessous, ces élèves calculent d'abord

$$d_2 = \sqrt{37} \quad \text{puis} \quad \vec{b}.\vec{c} = \frac{1}{4}(37 - 13) \quad \text{puis} \quad \cos A = \frac{\vec{b}.\vec{c}}{bc} = \frac{6}{3 \times 4} = \frac{1}{2}$$

au lieu d'utiliser

$$\cos A = \frac{b^2 + c^2 - a^2}{2bc}.$$



Les collègues professeurs de physique apprécient l'aspect expérimental de la méthode, sa rapidité pour arriver à la trigonométrie, et l'importance donnée aux parallélogrammes, figures qu'ils utilisent autant, sinon plus, que les triangles.

La question de la place de l'expérimentation dans l'apprentissage de la géométrie élémentaire est évoquée par M. Carral, dans son livre récent *Géométrie*, page 1, et par J. Dieudonné dans l'introduction de *Algèbre linéaire et Géométrie*, page 14.

8. Remarques complémentaires

a) Les quadruples de nombres présentés au §1 et vérifiant

$$2(c_1^2 + c_2^2) = d_1^2 + d_2^2, \quad c_1 + c_2 > d_1, \quad c_1 + c_2 > d_2$$

sont simples à manipuler parce qu'ils sont constitués d'entiers.

Pour les retrouver ou en trouver d'autres, on peut utiliser l'identité

$$2[(2a+b)^2+(a+3b)^2]=(3a+2b)^2+(a+4b)^2$$

pour a, b entiers > 0 et construire des quadruples

$$(c_1, c_2, d_1, d_2)=(2a+b,\ a+3b,\ 3a+2b,\ a+4b).$$

Une autre famille de tels quadruples est

$$(2a+9b,\ 8a+2b,\ 10a-b,\ 6a+13b) \qquad \text{pour } \ a>0,\ 0<b<2a.$$

On peut remarquer que si (c_1, c_2, d_1, d_2) vérifient la formule du parallélogramme, alors (C_1, C_2, D_1, D_2) tels que $C_1+iC_2=(a+bi)(c_1+ic_2)$, $D_1+iD_2=(a+bi)(d_1+id_2)$ la vérifient également (si $\arg(a+bi)$ est assez petit, on respecte encore les inégalités triangulaires).

Il y a une infinité de tels quadruples d'entiers, non proportionnels. Donc il y a une infinité de parallélogrammes non semblables à côtés et diagonales entiers copremiers.

b) Mieux : il existe des triangles à côtés entiers et médianes entières. (Notons a, b, c les longueurs des côtés, m_a, m_b, m_c les longueurs des médianes).

Par exemple : $a=136$, $b=170$, $c=174$, $m_a=158$, $m_b=131$, $m_c=127$ (ces nombres vérifient les trois égalités traduisant le théorème de la médiane : $a^2+b^2=2m_c^2+\frac{1}{2}c^2$, etc.).

De même, il existe des triangles à côtés entiers et hauteurs entières. Par exemple : $a=30$, $b=25$, $c=25$, $h_a=20$, $h_b=h_c=24$.

Mais on ignore encore s'il existe des triangles à côtés, médianes **et** hauteurs de longueurs entières (voir J.P. Delahaye, *Ignorance ou indécidabilité*, énigme n°5).

c) En géométrie euclidienne, on peut démontrer une "réciproque" de l'axiome des parallélogrammes : Si 4 points A, B, C, D d'un espace euclidien vérifient

$$AB^2+BC^2+CD^2+DA^2=AC^2+BD^2,$$

ils forment un parallélogramme (voir par exemple : Sortais, *Géométrie de l'espace et du plan*, p. 136).

9. Historique

Le théorème sur le carré d'un côté d'un triangle quelconque, duquel on déduit d'habitude le théorème de la médiane, est la proposition 13 du Livre II des *Eléments* d'Euclide. Le théorème de la médiane semble apparaître au Livre VII des *Collections mathématiques* de Pappus (voir T.L. Health, *Euclid : The Thirteen Books of the Elements*, Vol. 1, commentaires des propositions 9, 10 du Livre II, p. 401).

Grégoire de Saint-Vincent (1584–1667) et Viviani (1622–1703) en ont déduit l'énoncé relatif aux carrés des diagonales et des côtés de tout parallélogramme, cité plus haut sous le nom d'axiome des parallélogrammes.

En 1843, Grassmann montre l'existence du produit scalaire en géométrie euclidienne.

La démonstration de l'existence du produit scalaire, à partir de l'axiome des parallélogrammes, est attribuée à P. Jordan et J. von Neumann, *Ann. Math.* **36**, 719 (1935). La démonstration reproduite ici est adaptée de : A. Kolmogorov et S. Fomine, *Eléments de la théorie des fonctions et de l'analyse fonctionnelle*, (1974) p. 156–158.

[Un problème voisin, mais différent, la caractérisation des espaces euclidiens parmi les espaces métriques, par des conditions sur une seule distance (sans savoir que ce sont des espaces vectoriels, sans notion de droites parallèles ...) avait été résolu par K. Menger en 1928 (voir M. Berger, *Géométrie 1*, §9.7.4 p. 283).]

10. A quelle échelle et avec quelle précision notre espace physique est-il euclidien ?

A notre échelle humaine, nous avons pu vérifier, entre les longueurs des côtés (c_1 et c_2) et des diagonales (d_1 et d_2) des parallélogrammes, la relation

$$2(c_1^2 + c_2^2) = d_1^2 + d_2^2$$

avec une incertitue relative d'environ 1% (une incertitude de 1 mm sur 10 cm est est tout à fait à la portée des élèves).

Les techniques modernes de mesure pourraient faire beaucoup mieux, et sur une gamme de distances allant au moins du kilomètre au millimètre.

La vérification du parallélisme des côtés est un obstacle qui peut être contourné en remplaçant la relation du parallélogramme par le théorème de la médiane qui, dans ce contexte, lui est équivalent :

Si a, b, c sont les longueurs des côtés d'un triangle, et m_a la longueur de la médiane relative à a,

$$b^2 + c^2 = 2[(1/2\,a)^2 + m_a^2].$$

Nous avons vu que cette relation, jointe aux axiomes d'espace affine muni d'une fonction distance, suffit pour en faire un espace euclidien.

Réciproquement, tout espace euclidien vérifie le théorème de la médiane.

Dans deux domaines de la physique au moins, on est amené à étudier des espaces non euclidiens : en relativité générale, et en microphysique.

Le théorème de la médiane (ou l'axiome des parallélogrammes) n'étant plus d'application, par quoi doit-il être remplacé ?

Dans le cas d'un espace hyperbolique à définition globale (voir M. Berger, *Géométrie 2*, §19.2, p. 446), à fonction distance d, deux points distincts B et C possèdent un unique milieu M tel que

$$d(M, B) = d(M, C) = \frac{1}{2}\, d(B, C).$$

Pour tout point A, si l'on pose :

$$d(A, C) = b, d(A, B) = c, d(B, C) = a, d(M, A) = m_a,$$

alors

$$b^2 + c^2 \geqslant 2m_a^2 + \frac{1}{2}\, a^2 \qquad \text{(idem, §19.4.7 et 9.8.6.5)}$$

(cette propriété intervient dans la démonstration de la compacité du groupe des isométries conservant un ensemble compact).

L'inégalité devient stricte si A, B, C sont distincts 2 à 2.

Cette condition, cas particulier de la condition de Bruhat-Tits, généralise le théorème de la médiane. Dans le cas d'un espace elliptique (à définition globale), l'inégalité, appliquée localement, s'inverse (voir M. Berger, *Géométrie 1*, §9.8.6.4, p. 287).

En géométrie différentielle, on définit, sur l'espace vectoriel tangent en chaque point d'une variété riemannienne, un produit scalaire, un tenseur de torsion et un tenseur de courbure, qui peuvent varier avec le point considéré. Si ces deux tenseurs sont identiques à zéro, et si les groupes d'homotopie de la variété sont nuls, la variété riemannienne (supposée connexe et complète) est équivalente à un espace euclidien défini globalement.

On a effectué des mesures sur l'espace physique qui, dans le domaine des grandes longueurs, montrent qu'il ne peut pas être décrit avec précision comme un espace euclidien : on constate
a) la déviation des rayons électromagnétiques par la masse du Soleil.
b) l'avance du périhélie des planètes.
c) le décalage gravitationnel des raies spectrales vers le rouge.
d) l'effet de lentille gravitationnelle par les corps massifs.

Les écarts relatifs sont très faibles :

Au cours de l'éclipse totale de 1919, on a comparé l'arrière-plan d'étoiles au voisinage du Soleil avec ce même champ d'étoiles 6 mois plus tard : les rayons lumineux Etoile-Terre étaient déviés de $(1,98 \pm 0,12)$ secondes d'arc, alors que la relativité générale prévoit 1,75 secondes d'arc. La déviation par la Lune serait $1,6.10^7$ fois plus petite (pratiquement impossible à mesurer). (Voir, par exemple : G. Pascoli, *La gravitation*, 1989, p. 42–49).

Autre observation : lorsqu'un écho radar met environ 20 minutes pour l'aller-retour Terre-Mars, la durée de l'aller-retour augmente brusquement de quelques microsecondes quand Mars s'apprête à disparaître derrière le Soleil (effet Shapiro, 1964). En imaginant une sonde spatiale située hors de l'alignement Terre-Soleil-Mars, et capable de mesurer sa distance aux planètes, un tel effet prend en défaut le théorème de la médiane, donc l'"axiome des parallélogrammes".

Dans le domaine des petites longueurs, aucun écart au caractère euclidien de l'espace n'a pas encore été constaté pour des longueurs d'un ordre supérieur ou égal aux dimensions des atomes.

Donc, à notre échelle, pour les phénomènes usuels, et avec une très bonne précision, nous pouvons considérer l'espace comme euclidien.

Bibliographie

[1] Berger M. *Géométrie 1*, Cedic-Nathan, Paris (1990), 430 p. ; *Géométrie 2*, 541 p.

[2] Borceux F. *Invitation à la géométrie*, Ciaco, Louvain-La-Neuve (1986).

[3] Carral M. *Géométrie*, Editions Ellipses, Paris (1995), 244 p.

[4] Delahaye J.P. *Ignorance ou indécidabilité ?* Revue *Pour la Science*, juillet 1994, p. 94–99.

[5] Dieudonné J. *Algèbre linéaire et Géométrie*, Hermann, Paris (1969).

[6] Heath T.J. *Euclid - The Thirteen Books of the Elements. Vol. 1 (Books I and II)*, Dover, New York (1956), 432 p.

[7] Jordan P., von Neumann J. *Ann. Math.*, **36**, 719 (1935).

[8] Kolmogorov A., Fomin S. *Eléments de la théorie des fonctions et de l'analyse fonctionnelle* Mir, Moscou (1974) ; *Introductory Real Analysis*, Dover, New York (1973).

[9] Papy G. *Mathématique moderne 3*, Didier, Bruxelles-Montréal Paris (1972), 452 p.

[10] Pascoli G. *La gravitation*, Presses Universitaires de France, Paris (1989), 128 p.

[11] Sortais Y. et R. *Géométrie de l'espace et du plan*, Hermann, Paris (1988, 1993), 396 p.

[12] Tisseron C. *Géométrie affine, projective et euclidienne*, Hermann, Paris, Thème 5 (1994) 386 p.

Adresse de l'auteur :

E. JASON
Avenue Montgomery 13
1340 Ottignies

Variations autour de la définition des fonctions convexes (2ème partie)

J. Bair et G. Haesbroeck, *Université de Liège* (1)

8. Croissance de la dérivée pour des fonctions dérivables

Sachant que la convexité d'une fonction équivaut à la croissance de la pente des cordes par rapport à un point fixe (proposition 5.1), il convient de regarder si, pour une fonction dérivable, la convexité ne va pas de pair avec le caractère croissant de la dérivée. C'est effectivement ce que l'on peut démontrer dans le cas d'une fonction *aimable* sur I, c'est-à-dire d'une fonction continue sur I et dérivable sur $\overset{\circ}{I}$.

Proposition 8.1 *Soit f une fonction aimable sur I ; f est convexe sur I si et seulement si f' est croissante sur $\overset{\circ}{I}$.*

Preuve

Soient x_0, x_1 et x_2 trois points distincts de $\overset{\circ}{I}$, avec $x_1 < x_0 < x_2$. Par la proposition 7.1, $p(X_1X_0) \geqslant f'(x_1)$ et $p(X_2X_0) \leqslant f'(x_2)$. Or, d'après la proposition 5.1, $p(X_1X_0) \leqslant p(X_2X_0)$, d'où $f'(x_1) \leqslant f'(x_2)$ et f' est bien croissante sur $\overset{\circ}{I}$.

Soient x_1 et x_2 deux points quelconques de $\overset{\circ}{I}$ tels que $x_1 < x_2$. La formule des accroissements finis livre un réel $\theta \in]0,1[$ pour lequel

$$f(x_2) = f(x_1) + (x_2 - x_1)f'\left[x_1 + \theta(x_2 - x_1)\right].$$

Or, par hypothèse, f' est croissante sur $\overset{\circ}{I}$, d'où $f'(x_1) \leqslant f'\left[x_1 + \theta(x_2 - x_1)\right]$; dans ces conditions,

$$f(x_2) - f(x_1) \geqslant (x_2 - x_1)f'(x_1).$$

Comme x_1 et x_2 sont arbitraires dans $\overset{\circ}{I}$, on peut appliquer la proposition 7.1 qui garantit la convexité de f.

■

1. *La première partie de cet article est parue dans le n° 105 de Mathématique & Pédagogie (1996), pp. 57–70.*

9. Caractère non négatif de la dérivée seconde pour des fonctions deux fois dérivables

Dans la pratique, le moyen le plus commode pour reconnaître quand une fonction est convexe consiste, lorsque cela est possible, à étudier le signe de sa dérivée seconde qui caractérise effectivement la monotonie de la dérivée.

Proposition 9.1 *Soit f une fonction continue sur I et deux fois dérivable à l'intérieur de I ; f est convexe sur I si et seulement si $f''(x) \geqslant 0$ pour tout $x \in \mathring{I}$.*

Preuve

Ce résultat découle directement de la proposition précédente et de la caractérisation de la croissance d'une fonction à partir du signe de sa dérivée.

■

Cet énoncé permet de vérifier aisément la convexité des fonctions suivantes :

— pour tous réels a et b, $ax + b$ et e^{ax+b} sur $\mathbb{R}$;
— pour tout réel $a > 0$ (avec b, c quelconques), ax^2+bx+c, a^x et $(1+a)^x$ sur $\mathbb{R}$;
— pour $m \geqslant 1$, x^m sur $[0, +\infty[$, et même sur $\mathbb{R}$ quand, de surcroît, m est rationnel, à dénominateur impair et à numérateur pair ; pour $m \in]0, 1[$, $-x^m$ sur $[0, +\infty[$ et x^m sur $]-\infty, 0[$ quand, de plus, m est rationnel à numérateur et dénominateur impairs ; pour $m < 0$, x^m sur $]0, +\infty[$;
— $\ln \frac{1}{x} = -\ln x$ et $x \ln x$ sur $]0, +\infty[$; $\ln(1 + e^x)$ sur $\mathbb{R}$;
— $\sin x$ sur $[(2k + 1)\pi, 2(k + 1)\pi]$, $k \in \mathbb{Z}$;
— $\cos x$ sur $\left[\frac{\pi}{2} + 2k\pi, \frac{\pi}{2} + (2k + 1)\pi\right]$, $k \in \mathbb{Z}$;
— $\operatorname{tg} x$ sur $\left]k\pi, \frac{\pi}{2} + k\pi\right[$, $k \in \mathbb{Z}$;
— $\operatorname{Arcsin} x$ sur $[0, 1]$;
— $\operatorname{arctg} x$ sur $]-\infty, 0[$;
— $\sqrt{1 + x^2}$ sur $\mathbb{R}$;
— pour tout $a > 0$, $-\sqrt{a^2 - x^2}$ sur $]-a, a[$;
— e^{-x^2} sur $\left]-\infty, -\frac{\sqrt{2}}{2}\right[$ et sur $\left[+\frac{\sqrt{2}}{2}, +\infty\right[$; $\frac{-1}{1+e^{-x}}$ sur $[0, +\infty[$.

10. Inégalité de Jensen

Il est utile de rappeler qu'un sous-ensemble E de $\mathbb{R}^2$ est convexe si et seulement s'il contient toute *combinaison convexe* de ses éléments, étant entendu que $\sum_{i=1}^n \lambda_i X_i$ est une combinaison convexe de $X_1, \ldots, X_n$ lorsque $\lambda_i \geqslant 0$ pour tout $i = 1, \ldots, n$ et $\sum_{i=1}^n \lambda_i = 1$; de fait, pour E convexe, $X_i \in E$ et $\lambda_i \in [0,1]$ pour tout $i = 1, \ldots, n$, avec $\sum_{i=1}^n \lambda_i = 1$, on obtient de proche en proche :

$$\begin{aligned}
Y_1 &= (\frac{\lambda_1}{\lambda_1+\lambda_2})X_1 + (\frac{\lambda_2}{\lambda_1+\lambda_2})X_2 \in E,\\
Y_2 &= (\frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3})Y_1 + (\frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3})X_3 \in E,\\
&\vdots\\
Y_{n-1} &= (\frac{\lambda_1+\cdots+\lambda_{n-1}}{1})Y_{n-2} + \frac{\lambda_n}{1}X_n = \sum_{i=1}^n \lambda_i X_i \in E.
\end{aligned}$$

L'inégalité intervenant dans la définition analytique des fonctions convexes peut alors être généralisée au cas de combinaisons convexes, la formule obtenue étant attribuée à Jensen.

Proposition 10.1 *f est convexe sur I si et seulement si, pour tous $x_1, \ldots, x_n$ de I et tous $\lambda_1, \ldots, \lambda_n$ de $[0,1[$ tels que $\sum_{i=1}^n \lambda_i = 1$,*

$$f(\sum_{i=1}^n \lambda_i x_i) \leqslant \sum_{i=1}^n \lambda_i f(x_i).$$

Preuve

Grâce à la proposition 3.1, la convexité de f sur I et de $Epi(f)$ dans $\mathbb{R}^2$ va de pair et, compte tenu de la remarque précédente, équivaut au fait que toute combinaison convexe d'éléments de $Epi(f)$ est dans $Epi(f)$; la conclusion découle de l'appartenance des points $(x_i, f(x_i))$ dans $Epi(f)$.

■

De nombreuses inégalités, plus ou moins classiques et célèbres, entre nombres réels sont en réalité des conséquences de la formule de Jensen.

Ainsi, considérons n nombres réels positifs $x_1, \ldots, x_n$ et des réels positifs ou nuls $\lambda_1, \ldots, \lambda_n$ tels que $\sum_{i=1}^n \lambda_i = 1$.

On a $\prod_{i=1}^n x_i^{\lambda_i} \leqslant \sum_{i=1}^n \lambda_i x_i$ car la fonction $-\ln x$ est convexe sur $]0,+\infty[$, d'où $-\ln(\sum_{i=1}^n \lambda_i x_i) \leqslant -\sum_{i=1}^n \lambda_i \ln x_i = -\ln(\prod_{i=1}^n x_i \lambda_i)$; en particulier, en posant chaque λ_i égal à $\frac{1}{n}$, $\sqrt[n]{\Pi_{i=1}^n x_i} \leqslant \frac{1}{n}\sum_{i=1}^n x_i$, formule qui traduit la supériorité de la moyenne arithmétique des x_i sur leur moyenne géométrique.

Par ailleurs, si $a_1,\ldots,a_n,b_1,\ldots,b_n$ sont des nombres non négatifs et si $p>1, q>1$, avec $\frac{1}{p}+\frac{1}{q}=1$, alors on obtient l'*inégalité de Hölder*, à savoir $\sum_{i=1}^n a_i b_i \leqslant (\sum_{i=1}^p a_i^p)^{\frac{1}{p}} (\sum_{i=1}^n b_i^q)^{\frac{1}{q}}$: il suffit en effet d'utiliser la formule $x^{\frac{1}{p}} y^{\frac{1}{q}} \leqslant \frac{1}{p}x + \frac{1}{q}y$ pour $x = \frac{a_i^p}{\sum_{i=1}^n a_i^p}$ et $y = \frac{b_i^q}{\sum_{i=1}^n b_i^q}$ pour $i = 1,\ldots,n$, puis d'additionner membre à membre les n inégalités ainsi obtenues. En particulier, pour $p=q=2$, on retrouve l'*inégalité de Cauchy*

$$\sum_{i=1}^n a_i b_i \leqslant \sqrt{\sum_{i=1}^n a_i^2}\sqrt{\sum_{i=1}^n b_i^2}.$$

Comme la fonction $x\ln x$ est convexe sur $]0,+\infty[$, on a encore

$$(\sum_{i=1}^n \lambda_i x_i)^{\sum_{i=1}^n \lambda_i x_i} \leqslant \prod_{i=1}^n x_i^{\lambda_i x_i}.$$

Si $y_1,\ldots,y_n$ sont également des nombres positifs, on a de plus

$$\prod_{i=1}^n x_i^{\lambda_i} + \prod_{i=1}^n y_i^{\lambda_i} \leqslant \prod_{i=1}^n (x_i+y_i)^{\lambda_i} :$$

cela provient de la convexité de la fonction $f(z) = \ln(1+e^z)$ et de la substitution $z = \ln y - \ln x$.

Si, au surplus, les x_i sont dans $]0,1[$, on a

$$\sum_{i=1}^n \frac{\lambda_i}{1+x_i} \leqslant \frac{1}{1+\Pi_{i=1}^n x_i^{\lambda_i}} :$$

en effet, $f(z) = \frac{-1}{1+e^{-z}}$ est convexe sur $[0,+\infty[$ et l'on peut poser $z = \ln\frac{1}{x}$.

11. Convexité sur l'intérieur de l'intervalle pour des fonctions continues

On peut démontrer qu'une fonction convexe sur I est continue sur $\overset{\circ}{I}$: cela provient de l'existence, en tout point x^* de $\overset{\circ}{I}$, d'une dérivée à gauche $f'_-(x^*)$ et d'une dérivée à droite $f'_+(x^*)$ [6]. Pour vérifier qu'une fonction continue sur un intervalle fermé I est convexe sur I, il suffit de contrôler sa convexité sur $\overset{\circ}{I}$.

Proposition 11.1 *Soit f une fonction continue sur $[a,b]$; f est convexe sur $[a,b]$ si et seulement si elle est convexe sur $]a,b[$.*

Preuve

Supposons d'abord f convexe sur $]a,b[$. Pour tous $x,y \in\ ]a,b[$ et tout $\lambda \in\ ]0,1[$,

$$f\left[\lambda x + (1-\lambda)y\right] \leqslant \lambda f(x) + (1-\lambda)f(y).$$

Par continuité, on obtient pour $y \in\ ]a,b[$:

$$\lim_{x\to a^+} f\left[\lambda x + (1-\lambda)y\right] = f\left[\lambda a + (1-\lambda)y\right]$$

$$\leqslant \lim_{x\to a^+} \left[\lambda f(x) + (1-\lambda)f(y)\right] = \lambda f(a) + (1-\lambda)f(y).$$

De même, en passant à la limite pour y tendant vers b par valeurs inférieures, on trouve pour $x \in [a,b[$:

$$f\left[\lambda x + (1-\lambda)b\right] \leqslant \lambda f(x) + (1-\lambda)f(b).$$

En conséquence, f est bien convexe sur $[a,b]$.

La réciproque est évidente.

■

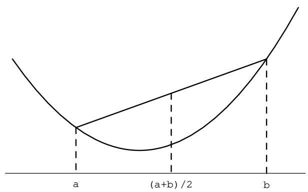
12. (J)-convexité pour des fonctions continues

Historiquement, la convexité a été introduite, tout au début de ce siècle, sous une forme plus faible que celle décrite dans les paragraphes précédents ; un des pionniers en la matière fut J.L. Jensen qui a considéré les fonctions f définies sur I et telles que, pour tous a, b de I,

$$f(\frac{a+b}{2}) \leqslant \frac{1}{2}f(a) + \frac{1}{2}f(b);$$

de telles fonctions seront appelées (J)-*convexes sur* I.

Géométriquement, la (J)-convexité consiste à exiger que le milieu de toute corde (joignant deux points du graphe) n'est jamais situé en-dessous du point de même abscisse sur le graphe.



Bien entendu, toute fonction convexe sur I est (J)-convexe sur I, mais on connaît des fonctions (J)-convexes qui ne sont pas convexes [4, p.443] ; de telles fonctions sont néanmoins très particulières, et notamment "méchamment discontinues" : on peut en effet constater que, pour des fonctions continues, les notions de convexité et de (J)-convexité sont équivalentes.

Proposition 12.1 *Soit f une fonction continue sur I ; f est convexe sur I si et seulement si elle est (J)-convexe sur I.*

Preuve

La condition est trivialement nécessaire.

Pour démontrer la réciproque, supposons f (J)-convexe sur I. Il est facile de vérifier (par récurrence) que, pour tout entier positif k et des points arbitraires $x_1, x_2, \ldots, x_{2^k}$ de I,

$$f(\frac{\sum_{i=1}^{2^k} x_i}{2^k}) \leqslant \frac{1}{2^k}\sum_{i=1}^{2^k} f(x_i).$$

Soit, à présent, un entier positif quelconque n ; nous nous proposons de démontrer que

$$f(\frac{\sum_{i=1}^{n} x_i}{n}) \leqslant \frac{1}{n}\sum_{i=1}^{n} f(x_i).$$

pour $x_1, \dots, x_n$ dans I.

Comme il existe un entier k tel que $n \leqslant 2^k$ et pour lequel l'inégalité ci-dessus est acquise, il suffit de prouver que, pour $x_1, \dots, x_N \in I$, $f(\frac{x_1+\cdots+x_N}{N}) \leqslant \frac{1}{N}\sum_{i=1}^{N} f(x_i)$ entraîne $f(\frac{x_1+\cdots+x_{N-1}}{N-1}) \leqslant \frac{1}{N-1}\sum_{i=1}^{N-1} f(x_i)$. C'est effectivement le cas puisque, pour $x_N = \frac{1}{N-1}(x_1+\cdots+x_{N-1}) = \frac{1}{N}(x_1+\cdots+x_N)$, l'hypothèse donne

$$f(x_N) \leqslant \frac{1}{N}\sum_{i=1}^{N-1} f(x_i) + \frac{1}{N} f(x_N),$$

d'où

$$f(x_N) \leqslant \frac{1}{N-1}\sum_{i=1}^{N-1} f(x_i).$$

Considérons à présent deux points arbitraires a, b de I, et des entiers positifs k et n, avec $k < n$. En vertu de ce qui précède, on obtient

$$\begin{aligned} f\left(\frac{a+\cdots+a+b+\cdots+b}{n}\right) &= f\left(\frac{ka}{n}+\frac{(n-k)b}{n}\right) \\ &\leqslant \frac{1}{n}\left[kf(a)+(n-k)f(b)\right], \end{aligned}$$

d'où, pour tout rationnel $t \in]0,1[$,

$$f\left[ta+(1-t)b\right] \leqslant tf(a)+(1-t)f(b).$$

Grâce à la continuité de f, l'inégalité précédente est encore valable pour t tendant vers un réel λ arbitraire de $]0,1[$, soit

$$f\left[\lambda a+(1-\lambda)b\right] \leqslant \lambda f(a)+(1-\lambda)f(b).$$

■

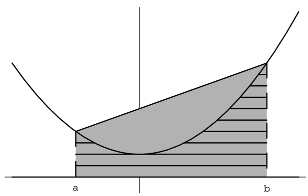
13. Inégalité sur les valeurs moyennes pour des fonctions continues

Pour des fonctions continues, la convexité peut s'exprimer à partir d'inégalités portant sur des intégrales.

Proposition 13.1 *Une fonction f continue sur un ouvert I est convexe sur I si et seulement si, pour tous a, b de I tels que $a < b$, est vérifiée l'inégalité suivante :*

$$\frac{1}{b-a}\int_a^b f(x)dx \leqslant \frac{1}{2}\left[f(a)+f(b)\right].$$

Interprétation. L'inégalité de l'énoncé exprime que la valeur moyenne de f sur $[a, b]$ est inférieure ou égale à la moyenne arithmétique des valeurs prises par f aux extrémités a, b. Géométriquement, si $f(x) \geqslant 0$ pour tout $x \in [a, b]$, cette formule traduit le fait que l'aire située sous la courbe, au-dessus de l'axe horizontal et entre les verticales $x = a, x = b$ est inférieure ou égale à l'aire du trapèze construit en remplaçant la courbe par la corde correspondante.



Preuve

Si f est convexe sur I, alors, pour tout $x \in [a, b] \subset I$, $f(x) \leqslant g(x)$, où $g(x) = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$, puisque la courbe est située sous ou sur la corde. En vertu de la propriété de conservation des inégalités par intégration, on a

$$\begin{aligned}\int_a^b f(x)dx \leqslant \int_a^b g(x)dx &= f(a)(b-a) + \frac{f(b)-f(a)}{b-a}\frac{(b-a)^2}{2} \\ &= (\frac{b-a}{2})\left[f(a)+f(b)\right].\end{aligned}$$

Réciproquement, raisonnons par l'absurde en supposant f non convexe sur I : il existe $a, b, c \in I$, avec $a < c < b$, tels que $f(c) > g(c)$. Grâce à la continuité de f, on peut supposer sans restriction que $f(x) > g(x)$ pour tout $x \in]a, b[$, quitte à restreindre l'intervalle $[a, b]$ de départ. Dans ces conditions, on conserve l'inégalité stricte par passage à l'intégrale entre a et

b, soit

$$\int_a^b f(x)dx > \int_a^b g(x)dx,$$

ce qui contredit l'hypothèse initiale.

■

14. Définition moderne sur toute la droite numérique

En convexité, une tendance actuelle (voir, par exemple, [5],[8],[10],[11], [14]) consiste à ne prendre en considération que des fonctions définies sur toute la droite numérique $\mathbb{R}$, mais à valeurs dans la droite numérique achevée $\bar{\mathbb{R}}$ (et même, souvent, dans $\mathbb{R} \cup \{+\infty\}$) : cela dispense de préciser à chaque fois le domaine de définition de la fonction.

Une fonction $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, non partout égale à $+\infty$, est dite appartenir à l'ensemble $Conv\mathbb{R}$ lorsqu'est vérifiée, au sein de la droite numérique achevée, l'inégalité

$$f\left[\lambda x + (1-\lambda)y\right] \leqslant \lambda f(x) + (1-\lambda) f(y)$$

pour toute paire $\{x, y\}$ de réels et tout réel $\lambda \in]0,1[$.

Bien entendu, lorsque $f \in Conv\mathbb{R}, \{x \in \mathbb{R} : f(x) \in \mathbb{R}\}$ est un intervalle I et f est convexe sur I. Réciproquement, si f est convexe sur un intervalle I, l'extension $\bar{f}$ définie par

$$\bar{f}(x) = \begin{cases} f(x) & \text{si} \;\; x \in I \\ +\infty & \text{si} \;\; x \notin I \end{cases}$$

appartient à $Conv\mathbb{R}$. En conséquence, cette nouvelle présentation de la convexité n'est pas fondamentalement différente de la définition initiale ; elle s'avère néanmoins commode, car elle ne demande pas de spécifier explicitement l'intervalle sur lequel est convexe la fonction considérée.

Cette version de la convexité est souvent utilisée en optimisation. A titre d'exemple, considérons une fonction f convexe sur I et cherchons à minimiser f sur un sous-intervalle J de I ; pour la fonction $\bar{f}$ définie sur $\mathbb{R}$ par

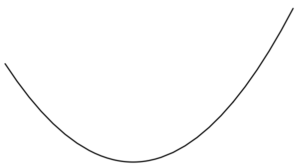
$$\bar{f}(x) = \begin{cases} f(x) & \text{si} \;\; x \in J \\ +\infty & \text{si} \;\; x \notin J, \end{cases}$$

on a visiblement $\bar{f} \in Conv\mathbb{R}$: le problème de minimisation posé est alors équivalent à la minimisation de $\bar{f}$ sur tout l'espace $\mathbb{R}$.

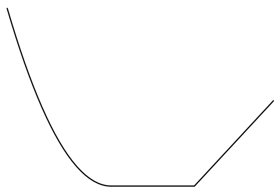
15. Notions apparentées

15.1. Stricte convexité

Une fonction f est *strictement convexe sur I* lorsque toute corde reste strictement au-dessus du graphe, ses extrémités exceptées ; cela revient à dire qu'une fonction strictement convexe sur I est convexe sur I et que son graphe ne contient aucun segment de droite (non réduit à un seul point) ; bien entendu, toute fonction strictement convexe est convexe, mais la réciproque n'est pas vraie.



strictement convexe



convexe (non strictement convexe)

Analytiquement, f est strictement convexe sur I lorsque, pour toute paire $\{a, b\}$ de points de I et pour tout réel $\lambda \in]0, 1[$,

$$f\left[\lambda a + (1-\lambda)b\right] < f(a) + (1-\lambda)f(b) :$$

il suffit en réalité de reprendre la définition donnée au paragraphe 2 en y remplaçant le signe $\leqslant$ par $<$. D'ailleurs, les énoncés des propositions 4.1, 5.1, 6.1, 6.2, 7.1, 8.1 et 10.1 peuvent être adaptés à la stricte convexité en changeant les signes d'inégalité large par les signes correspondants d'inégalité stricte (et notamment, en remplaçant l'hypothèse de croissance par celle de stricte croissance). La proposition 9 fait toutefois exception à cette règle ([2]) ; en effet, si $f''(x) > 0$ pour tout $x \in \mathring{I}$ (sauf, éventuellement, en un nombre fini de points), alors f est strictement convexe sur f, mais la réciproque n'est pas vraie : pour preuve, la fonction $f(x) = x^4$ est strictement convexe sur $\mathbb{R}$ alors que $f''(0) = 0$; en fait, on peut simplement affirmer que si f est

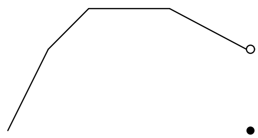
2. On trouve dans la littérature de nombreuses erreurs à ce sujet.

strictement convexe sur I (et deux fois dérivable sur $\mathring{I}$), alors tout intervalle ouvert inclus dans $\mathring{I}$ contient des points où la dérivée seconde de f est strictement positive.

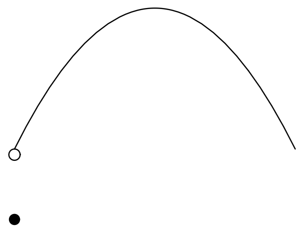
15.2. Concavité et stricte concavité

Une fonction f est dite *concave sur* I lorsque toute corde ne passe nulle part au-dessus du graphe ; elle est dite *strictement concave sur* I lorsque la corde en question reste strictement au-dessous du graphe, ses extrémités exceptées.

Il va sans dire qu'une fonction strictement concave est concave, la réciproque n'étant pas vraie.



concave
(pas strictement concave)



(strictement concave)

Le parallélisme très marqué entre l'étude des fonctions (strictement) concaves et celle des fonctions (strictement) convexes est la conséquence de cette remarque : f est concave (resp. strictement concave) sur I si et seulement si $-f$ est convexe (resp. strictement convexe) sur I.

Par passage à l'opposé, l'étude de la (stricte) concavité se déduit donc instantanément de celle de la (stricte) convexité. Contentons-nous d'en mentionner les faits les plus marquants. f est concave sur I si et seulement si $f\,[\lambda a + (1-\lambda)b] \geqslant \lambda f(a) + (1-\lambda)f(b)$ pour tous $a, b \in I$ et tout $\lambda \in \,]0,1[$; si, de plus, l'inégalité est stricte, f est strictement concave sur I. Une fonction aimable f sur I est concave (resp. strictement concave) sur I si et seulement si f' est décroissante (resp. strictement décroissante) sur $\mathring{I}$. Une fonction f continue sur I et deux fois dérivable sur $\mathring{I}$ est concave sur I si et seulement si $f''(x) \leqslant 0$ pour tout $x \in \mathring{I}$; pour une telle fonction f, si

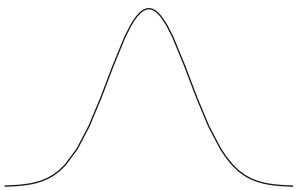
$f''(x) < 0$ pour tout $x \in \overset{\circ}{I}$ sauf au plus en un nombre fini de points, alors f est strictement concave sur I.

Par exemple, $ax + b$ et $-|x|$ sont concaves sur $\mathbb{R}$. Par ailleurs, sont strictement concaves : $\ln x$ sur $]0, +\infty[$; $\sin x$ sur $[2k\pi, (2k+1)\pi]$; $\cos x$ sur $\left[-\frac{\pi}{2} + 2k\pi, \frac{\pi}{2} + 2k\pi\right]$; $\operatorname{tg} x$ sur $\left]-\frac{\pi}{2} + k\pi, k\pi\right[$; $\operatorname{Arcsin} x$ sur $[-1, 0]$; $\operatorname{arctg} x$ sur $[0, +\infty[$; e^{-x^2} sur $\left[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right]$; x^m sur $[0, +\infty[$ pour $0 < m < 1$; x^m sur $]-\infty, 0[$ pour m rationnel supérieur à 1, à numérateur et dénominateur impairs, ou pour m rationnel inférieur à 1, à dénominateur impair et numérateur pair.

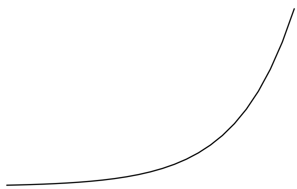
15.3. Quasi-concavité et notions dérivées

Si f est concave sur I, alors il est facile de vérifier que, pour tout réel α, l'ensemble $E(f, \alpha) = \{x \in I : f(x) \geqslant \alpha\}$ est un intervalle, donc convexe ; mais il est possible de trouver une fonction f non concave et un réel α tels que $E(f, \alpha)$ est convexe : c'est le cas, par exemple, pour la fonction e^{-x^2}.

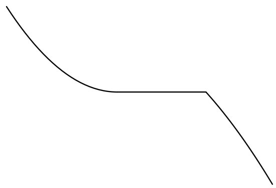
Une fonction f est dite *quasi-concave sur I* lorsque, pour tout réel α, $E(f, \alpha)$ est convexe ; toute fonction concave sur I est quasi-concave sur I, mais la réciproque n'est pas vraie.



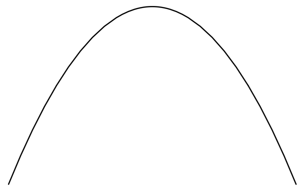
quasi-concave
non concave et
non convexe



quasi-concave
et convexe



quasi-concave
non concave
non convexe



concave, donc
quasi-concave

Une façon équivalente de présenter cette notion consiste à remarquer que f est quasi-concave sur I si et seulement si, pour tous $a, b \in I$ et tout $\lambda \in]0,1[$,

$$f\,[\lambda a + (1-\lambda)b] \geqslant \min\{f(a), f(b)\}.$$

Par exemple, toute fonction monotone (croissante ou décroissante) sur I est quasi-concave sur I, une fonction strictement monotone étant strictement quasi-concave. Si f est une fonction (strictement) quasi-concave sur I et si g est (strictement) croissante sur $f(I)$, alors $g \circ f$ est (strictement) quasi-concave sur I ; cette propriété permet de démontrer que e^{-x^2} est strictement quasi-concave sur $\mathbb{R}$ (car $-x^2$ est strictement concave et l'exponentielle strictement croissante) et que x^α, pour $\alpha > 0$ fixé, est strictement quasi-concave sur $]0, +\infty[$ (car $\ln x^\alpha$ est strictement concave et le logarithme népérien (donc l'exponentielle) est strictement croissant(e).).

De la même manière, on définit une fonction f *strictement quasi-concave sur* I par la condition

$$f\,[\lambda a + (1-\lambda)b] > \min\{f(a), f(b)\}$$

pour tous $a, b \in I$ et tout $\lambda \in]0,1[$; f est *quasi-convexe* (resp. *strictement quasi-convexe*) sur I lorsque $-f$ est quasi-concave (resp. strictement quasi-concave) sur I ou de façon équivalente, lorsque

$$f\,[\lambda a + (1-\lambda)b] \leqslant \max\{f(a), f(b)\}$$

(resp. $f\,[\lambda a + (1-\lambda)b] < \max\{f(a), f(b)\}$) pour tous $a, b \in I$ et tout $\lambda \in]0,1[$.

Les fonctions (strictement) quasi-concaves sont souvent exploitées, car elles permettent de modéliser de nombreuses situations concrètes (ainsi qu'en témoigne la courbe normale de Gauss définie par une fonction strictement quasi-concave, mais non concave ni convexe). Dans le domaine de l'optimisation, notamment, ces fonctions possèdent des propriétés fort intéressantes (assez semblables à celles des fonctions concaves).

15.4. Remarque finale

Il existe de nombreuses notions voisines des précédentes ; elles ont toutes leur utilité, leurs applications et leurs propriétés spécifiques : citons, par exemple, les fonctions approximativement convexes, pseudo-concaves, semi-concaves, invexes,...

Nous invitons le lecteur intéressé à consulter la littérature spécialisée.

Bibliographie

[1] Archinard G. - Guerrien B., *Analyse mathématique pour économistes*, Economica, 3ème édition, Paris, 1988.

[2] Avriel M. - Diewert W. - Schaible S. - Zang I., *Generalized concavity*, Plenum Press, New York, 1988.

[3] Bair J., *Mathématiques générales à l'usage des sciences économiques, de gestion et A.E.S.*, De Boeck Université, 3ème édition, Bruxelles, 1993.

[4] Beckenbach E.F., *Convex functions*, Bull. Amer. Math. Soc., vol. 54, 1948, pp. 439-460.

[5] Hiriart-Urruty J.B. - Lemaréchal C., *Convex analysis and minimization-algorithms I*, Springer-Verlag, Berlin-Heidelberg-New York, 1991.

[6] Janssen J., *Traité de mathématique appliquée pour l'assurance, l'économie et la finance*, tome 2 : *fonctions*, Office international de librairie, Bruxelles, 1993.

[7] Jongmans F. - Varlet J., *Notions de mathématique à l'usage des sciences humaines*, tome 1, notes de cours, Université de Liège, 1970.

[8] Moulin H. - Fogelman-Soulié F., *La convexité dans les mathématiques de la décision*, Hermann, Paris, 1979.

[9] Roberts A.W. - Varberg D.E., *Convex functions*, Academic Press, New York, 1973.

[10] Rockafellar R.T., *Convex analysis*, Princeton University Press, New Jersey, 1972.

[11] Valentine F.A., *Convex sets*, Mc Graw-Hill, New York, 1965.

[12] Van Tiel J., *Convex analysis : an introductory text*, J. Wiley and Sons, Chichester, 1984.

[13] Vodnev V. - Naoumovitch A et N., *Dictionnaire des mathématiques*, Academia et Editions Mir de Moscou, 1993.

[14] Wets R., *Grundlagen Konvexer Optimierung*, Springer-Verlag, Lecture Notes in Economics and Mathematical Systems, n° 137, Berlin-Heidelberg-New York, 1976.

Adresse des auteurs :

Université de Liège
Faculté d'Economie, de Gestion et de Sciences Sociales
Boulevard du Rectorat 7, Bât 31
4000 Liège
Belgique

Evaluation

M. Garin, J. Liesenborghs, P. Marlier, *Membres de la Commission Pédagogique de la SBPM*

1. Préliminaire

Ce texte a ses racines dans les travaux de la Commission Pédagogique de la SBPM au cours de l'année scolaire 1994-95.

Il ne paraît pas comme document de la Commission mais sous la signature et la responsabilité de ses principaux auteurs parce que l'unanimité n'a pas pu se faire sur certains points. Plutôt que de passer un temps considérable à le remettre en chantier pour arriver à un document plus satisfaisant pour les uns mais moins pour les autres, il a semblé préférable de le livrer à la réflexion de tous dans son état actuel et sa forme parfois abrupte. De toute manière, personne n'a la prétention de dire enfin la vérité définitive sur cette question difficile de l'évaluation, mais tous jugent utile de communiquer à leurs collègues un document de réflexion de gens de terrain.

> *"En réalité, le regard de l'autre ne me transforme en objet, et mon regard ne le transforme en objet, que si nous nous retirons dans le fond de notre nature pensante, si nous faisons l'un et l'autre regard inhumain, si chacun sent ses actions, non pas reprises et comprises, mais observées comme celles d'un insecte".*
>
> Merleau-Ponty

2. Evaluation certificative

2.1. Cadre et raisons d'être

L'évaluation certificative a pour but de vérifier si les objectifs attribués par le commanditaire (le plus souvent la "société") à un enseignement donné ont été atteints. Elle se manifeste par la sanction qu'un spécialiste (le professeur) transmet à la société (le client) sur la *qualité des travaux* réalisés par un apprenant dans un cadre donné.

LA SOCIETE souhaite protéger ses membres des incompétents ou des charlatans en instaurant des accès à la profession pour ceux qui, individuellement, collectivement ou par des entreprises, offrent des biens ou des services au public et en demandent rémunération. Idéalement, on souhaite qu'un médecin certifié soit un bon médecin, qu'un ingénieur ne construise pas de pont qui s'écroule, qu'un instituteur soit compétent dans les matières qu'il doit enseigner et dans la manière de transmettre le savoir, qu'un menuisier soit capable de réaliser un escalier bien construit, ...

L'ECOLE est une institution dont les finalités individuelles et sociales sont diverses : citons l'éducation à l'autonomie et l'apprentissage des contraintes inhérentes à la vie en société, le développement de la personnalité et la découverte de l'Autre, l'acquisition de connaissances, de compétences mais aussi les implications morales de toute activité. Les objectifs généraux de notre enseignement ont été définis par le Conseil de l'Education et de la Formation (CEF) :

Objectif 1 : "L'enseignement doit promouvoir le développement de la personne de chacun des élèves."

Objectif 2 : "L'enseignement, en amenant les jeunes à construire leur savoir, doit les conduire à prendre une place active dans la vie économique."

Objectif 3 : " L'enseignement doit amener les jeunes à être des citoyens responsables dans une société libre."

En pratique, plusieurs de ces objectifs resteront souvent implicites, et l'évaluation certificative ne pourra porter que sur un certain nombre d'entre eux. On peut même se demander si elle ne porte pas exlusivement sur la manière dont est atteint l'objectif 2, que plusieurs préféreraient d'ailleurs voir reporter en troisième position dans cette hiérarchie. Il est par conséquent inutile de reprocher à l'évaluation certificative de ne pas faire ce qui ne lui est pas assigné.

2.2.

1. *L'école fondamentale* a pour objetif d'amener les individus à maîtriser au moins le minimum nécessaire face à une société fortement organisée comme la nôtre : ceci expliquerait la tendance actuelle à vouloir considérer

que l'enfant recevra un enseignement de type fondamental jusqu'à 14 ans. A ce niveau, on comprend que cela implique une obligation de succès.

2. *L'enseignement de qualification* doit comporter des épreuves qui permettent de décider si les récipiendaires sont ou non aptes à assumer les fonctions pour lesquelles ils ont entrepris une formation.

3. *L'enseignement de transition* ([1]) doit certifier à certains moments que le récipiendaire pourra le plus probablement suivre avec fruit les travaux d'une étape ultérieure. Cela ne signifie nullement qu'il s'agisse d'une certitude : la certification porte sur le passé et sur des acquisitions ; c'est un pronostic d'évolution favorable. Elle ne peut être lue comme une sorte d'assurance anti-échec.

2.3. L'évaluation certificative intervient en fin de formation.

Elle ne constitue pas une aide pour l'apprenant et se fait sans son avis. Elle engage totalement la responsabilité du professeur à l'égard de la société. On peut comprendre l'angoisse humaine du professeur obligé de dire "ça ne va pas" en sachant bien que cela peut signifier l'écroulement de tout un projet de vie pour le récipiendaire, mais dans ce rôle, c'est à la société qu'il doit rendre des comptes.

L'évaluation certificative a certes pour but d'attester que des individus ont suivi avec fruit une formation donnée (fonction positive) mais une telle affirmation n'a de sens que si ladite certification n'est pas attribuée automatiquement à tout le monde, c'est-à-dire que si est ouverte la possibilité réelle d'élimination (fonction négative).

1. Historiquement, l'objectif de l'enseignement général de transition était la formation de l'"Honnête Homme", c'est-à-dire le développement strictement personnel sans aucun projet de préparation à l'exercice d'une profession. Dans un deuxième temps, cela est devenu la formation de spécialistes destinés à assumer des postes à responsabilités dans la société. Ceci a permis de développer l'idée que le jeune devait sacrifier beaucoup à ses études car il pourrait en recueillir les fruits plus tard. Actuellement, puisqu'on ne peut plus affirmer qu'une scolarité réussie va nécessairement déboucher sur une "bonne situation", il ne faut pas s'étonner d'entendre naître un discours du type "Qu'ils en profitent aujourd'hui car on ne sait pas de quoi demain sera fait".

2.4. L'évaluation certificative évalue des travaux.

La rédaction de bulletins est un moyen d'informer les apprenants et leurs parents de la *valeur des travaux* qui ont été réalisés dans le cadre de leur formation et il est important de savoir que ce n'est rien d'autre. Cependant, les travaux et épreuves certificatives ont été en principe conçus pour pouvoir fonder un jugement ayant valeur de pronostic sur la qualité des prestations futures du récipiendaire.

Il importe par ailleurs de bien distinguer le BULLETIN qui est un document PRIVE qui transmet de l'information sur l'avancement des apprentissages, du DIPLOME ou CERTIFICAT, qui est un document PUBLIC exigible par un employeur ou un responsable de formation ultérieure.

2.5. L'évaluation certificative concerne aussi le système d'enseignement.

S'il apparaît que les objectifs attribués à un enseignement ne sont atteints que par un nombre d'élèves jugé insuffisant, il importe de s'interroger sur l'organisation, le contenu et les moyens mis à la disposition de cet enseignement.

2.6. Conclusions.

1. Que le professeur soit un spécialiste, qu'il juge en son âme et conscience et que, sauf cas particulier à démontrer, son honnêteté ne puisse être mise en cause, ne signifie pas que des avis différents ne puissent être exprimés à propos d'un même élève. Ceci n'est pas une situation isolée dans la société : les tribunaux connaissent les batailles d'experts, les malades qui consultent plusieurs médecins avant de décider d'une intervention chirurgicale reçoivent souvent des avis divergents voire opposés, ... Il n'empêche que, de même que par la sentence du juge le prévenu devient innocent ou coupable par la décision d'un jury ou d'un professeur, l'élève acquiert une "*valeur intellectuelle sociale*" c'est-à-dire admise et reconnue par la société. Cette "valeur intellectuelle sociale" n'est pas la valeur intellectuelle de l'élève et encore moins sa valeur personnelle.

Par ailleurs, il ne faut pas confondre l'évaluation et son codage. Puisque le résultat de l'évaluation est un message transmis à l'intéressé et à la société, il importe qu'il soit codé de manière lisible par ceux à qui il est destiné. On peut donc modifier le codage (note littérale ou chiffrée, note unique ou rubriques multiples, ...) pour améliorer la lisibilité, mais ceci ne change rien au contenu du message.

2. L'évaluation certificative, traduisant concrètement les objectifs attribués à un enseignement, a certainement une influence sur cet enseignement lui-même. Il faut éviter que cette influence ne détermine complètement l'enseignement. L'évaluation certificative ne doit s'intéresser qu'aux compétences jugées indispensables. Elle ne peut toutefois pas empêcher l'enseignant de s'efforcer, dans ses cours, de développer des compétences qui débordent des objectifs minimaux pour autant que ces derniers soient atteints. Il est impératif de s'interroger sur la façon dont les objectifs minimaux sont déterminés.

3. L'évaluation certificative se situe en fin de parcours : les élèves ne peuvent recevoir leur diplôme (être certifiés) que s'ils maîtrisent les compétences leur permettant soit d'exercer une activité professionnelle, soit d'entamer d'autres étapes de formation.

4. Dans ce cadre se pose la question de savoir si on peut TOUT évaluer ; comment par exemple mesurer le degré d'acquisition de l'esprit d'analyse ou de synthèse ? Quelle place faire à de tels objectifs dans la liste des compétences minimales ?

5. Pour clôturer ces réflexions sur l'évaluation certificative, on ne peut manquer de se poser une question essentielle : QUE COMPTE FAIRE LA SOCIETE DES ELEVES QUI N'ONT ATTEINT QUE PEU OU PAS DU TOUT LES OBJECTIFS QU'ELLE AVAIT IMAGINES POUR EUX ?

Il existe deux types de réponses à cette question : ou bien on met tout en oeuvre pour leur donner une nouvelle chance, en permettant des redoublements si on ne dispose pas d'autres moyens, en imaginant et en remettant en place des remédiations plus efficaces, ou bien on les réoriente vers des domaines où ils se montreraient plus compétents, en évitant en ceci de rester coincés dans des stéréotypes hiérarchiques. Actuellement, toute réorientation est imaginée à sens unique du général vers le technique et du technique vers le professionnel. Si on se résigne à faire semblant, c'est-à-dire à laisser un nombre important de jeunes quitter l'école dans un état de non-formation et de non-qualification, on réalisera des économies à court terme, mais on risque de payer cette note plus tard en termes de chômage, de lutte contre la violence, de réadaptation, etc. ...

3. Evaluation formative

QUI ? QUOI ?

L'évaluation formative est une composante de la formation. Elle a deux fonctions essentielles : l'une est de permettre à l'élève de se situer lui-même par rapport aux objectifs (minimaux ou non) qu'il s'est fixés dans le type d'études qu'il a entrepris ; l'autre de permettre à l'enseignant de déterminer s'il peut progresser dans ses enseignements et à quelle vitesse. A partir de là, se posent un certain nombre de questions aussi bien pour l'enseignant que pour l'élève. Il est à souligner qu'une évaluation n'est JAMAIS NEUTRE mais dépend du projet poursuivi, et cela se traduit notamment dans le choix des épreuves. Par exemple, on peut se demander si l'élève se prépare à passer des examens ? Tels qu'ils sont ou tels qu'ils devraient être ? S'il vise l'acquisition de compétences transférables ailleurs qu'à l'école ?

On voit qu'une évaluation formative sérieuse doit déboucher sur l'auto-évaluation, c'est-à-dire sur l'autonomie, la liberté et la responsabilité. S'évaluer, c'est prendre en compte ses points forts, ses lacunes, ses progrès ; s'évaluer, c'est se construire des stratégies performantes ; s'évaluer aboutit à aborder avec plus de confiance et de lucidité les épreuves de certification et tirer des enseignements de l'analyse des résultats obtenus en vue d'une optimisation de ses capacités. Dans ce contexte, le professeur n'est pas l'expert qui juge l'apprenant au bénéfice de la société, mais une personne-ressource mise à sa disposition par la société pour lui transmettre des connaissances, l'aider à les structurer, à en rechercher de nouvelles, ... Il est aussi le gestionnaire du groupe qu'il dynamise, encourage et dont il oriente et stimule les recherches.

3.1. Comment ?

L'évaluation formative ne peut être qu'une évaluation en situation, c'est-à-dire l'observation de l'apprentissage en acte, des stratégies utilisées et des effets produits. Comme le dit Ph. MEIRIEU, il faudrait "organiser la classe pour que les élèves y travaillent, afin de pouvoir consacrer son attention à repérer sur quoi et comment on pourra engrener de nouvelles acquisitions. La stratégie d'un sujet est incontournable, et pourtant elle doit être dépassée, mais elle ne pourra l'être que si, dans un premier temps, on l'a d'abord respectée."

Un autre aspect du problème est qu'il est de la nature même de ce type d'évaluation d'être interactive. Elle requiert donc un engagement de la part de chacun des protagonistes. L'élève est en droit d'attendre une information significative en termes de but qu'il s'est assigné, de sa volonté de réussir ; le professeur doit pouvoir compter sur une information fiable, non truquée, pour assurer sa part de régulation du processus d'apprentissage.

3.2. Le bulletin

Le dialogue entre l'enseignant et chacun de ses élèves est donc essentiel à l'évaluation formative. Dans la plupart des cas, cette évaluation ne sera pas traduite en notes chiffrées. En fait, tout retour d'information de l'élève vers le professeur ou du professeur vers l'élève participe de l'évaluation formative ; ainsi en est-il de toute appréciation communiquée par le professeur à un élève concernant un travail qu'il vient de terminer ou qui doit être entrepris. L'évaluation formative est dans une très large mesure informelle : des explications complémentaires données en classe, des commentaires écrits sur un devoir à domicile, des conversations pesonnelles où interviennent des conseils positifs ou des mises en garde de maladresses à éviter, ... en relèvent évidemment. On voit qu'il ne s'agit pas de quelque chose de nouveau, mais il est sans doute utile de la sortir de l'implicite et d'apprendre à mieux en maîtriser les techniques.

3.3. Apprentissage

L'évaluation formative suppose de la part des enseignants une bonne connaissance des phénomènes d'apprentissage. C'est au professeur d'établir un dialogue avec son élève pour l'amener à devenir conscient de ses erreurs et des moyens de les corriger. Ce n'est pas toujours facile. L'évaluation formative suppose de la part du professeur une bonne maîtrise des diverses composantes (scientifique, pédagogique, psychologique) de son activité.

4. D'un point de vue pratique

On pourrait dire en résumé qu'il existe une évaluation pour aider (la formative) et une évaluation pour juger (la certificative). Mais dans les choses humaines, rien n'est simple : une évaluation formative mal appliquée peut devenir inhibitrice si elle ne respecte pas l'autonomie des élèves et les coince dans une démarche intellectuelle reproduisant celle du professeur. Par ailleurs une évaluation certificative dans ses intentions peut avoir des effets positifs si elle est vécue par l'élève comme un défi à relever, par exemple.

Etant donnés la complexité du problème et le cadre de notre système d'enseignement, notre objectif serait de donner aux maîtres l'envie et les moyens d'aller dans le sens d'une évaluation plus formative tout en composant avec des procédures traditionnelles (en attendant de pouvoir les modifier éventuellement).

Une formation à l'évaluation devrait être centrée sur la pratique. Il serait illusoire de donner des consignes rigides.

Parmi les tâches importantes, relevons :

- responsabiliser les élèves dans le projet de leur formation, c'est-à-dire les engager à se détacher de la tutelle de l'enseignant et à acquérir des bases d'autonomie.

- fournir aux élèves des outils d'auto-évaluation pour les rendre capables de mesurer leur progression.

- rendre les élèves actifs et inventifs, ou tout au moins favoriser les démarches qui encouragent et développent ces qualités.

- évaluer ce type de démarche plutôt que mettre l'accent sur leur habileté à reproduire un produit fini. (Bien sûr il est plus facile de faire reproduire que de faire comprendre.)

- éviter de réduire le cours de mathématique à un ensemble de procédures.

- ne pas tout contrôler, mais sélectionner des tâches qui font appel à plusieurs notions.

- revoir le statut de l'erreur qui dans le cadre de l'évaluation formative n'est pas considérée comme un manque ou un échec mais bien comme un indicateur d'intégration de l'apprentissage.

- rendre la parole aux élèves en les habituant à verbaliser pour communiquer entre pairs et pas uniquement avec le professeur.

- les inviter le plus souvent possible à exprimer leur propre pensée et à défendre leurs idées.

- les inciter à discuter de mathématique.

- pratiquer l'évaluation formative non pas sur un travail écrit et terminé mais bien en observant les élèves en action et en discutant avec eux au lieu de se contenter de poser des questions.

- privilégier la pédagogie du problème au lieu de la pédagogie de la réponse.

- mettre l'accent sur la reconstruction des mathématiques par chaque élève plutôt que sur la transmission de connaissances livresques.

Bibliographie

[1] Abrecht R., *L'évaluation formative. Une analyse critique*, Bruxelles, De Boek Université, Pédagogies en mouvement, 1991, 144 pages.

[2] Clausse A., *Philosophie et méthodologie d'un enseignement rénové*, Liège, Georges Thone, 1972

[3] Gasquet S., *Apprivoiser les maths*, l'école des parents, Syros Alternative.

[4] Marlier P., Ceci est-il bien évalué ? ou les arbitraires de l'évaluation, REVUE de l'Organisation des Etudes, 1988, 8, 21-38

[5] Meirieu Ph., *Apprendre ... oui mais comment ?*, Paris, ESF, 1992

[6] Meirieu Ph., *Enseigner, scénario pour un métier nouveau*, Paris, ESF, 1990.

Adresse des auteurs :

Michèle GARIN
Zeedijk 360
8670 KOKSIJDE

Jacqueline LIESENBORGHS
Rue de la Ramonerie 57
5032 BOSSIERE

Pierre MARLIER
Rue de Plainevaux 185/15
4100 SERAING

Bibliographie

J. Bair,

Théorie des graphes pour gestionnaires par J. BAIR. Editions Point de Vue (Sart Tilman, Bât. B7, 4000 LIEGE, Tél. : (041) 66.33.69), 148 pages, 1995

Extrait de l'introduction :

Le but premier de cet ouvrage est de présenter les notions fondamentales de la théorie des graphes : à savoir les définitions de base, les principaux résultats théoriques ainsi que différents algorithmes (parmi les plus simples et les plus efficaces).

Avec ce léger bagage théorique peuvent être étudiés de nombreux problèmes réels rencontrés dans le monde des affaires. Seront successivement abordés les thèmes suivants : la gestion scientifique séquentielle qui exploite le principe de programmation dynamique, l'ordonnancement pour planifier les différentes tâches composant un projet, l'affectation de candidats à des postes de travail, les flots (de marchandises) au sein d'un réseau de transport et la théorie d'aide multicritère à la décision.

L'accent est principalement mis sur les concepts et leur utilisation pratique, moins sur les développements théoriques. C'est pourquoi, certaines démonstrations, plus techniques, ont été volontairement omises. Une exception toutefois à cette règle : la preuve, pourtant longue, du théorème d'impossibilité d'ARROW est entièrement explicitée, car, outre la beauté du raisonnement, ce résultat est, du point de vue intellectuel, d'une importance capitale : il démontre notamment qu'il n'existe aucun système "démocratique" de votes et explique dès lors pourquoi sont développées autant de méthodes multicritères d'aide à la décision.

Insistons sur le fait que la lecture de ce livre ne nécessite pratiquement aucun prérequis en mathématique supérieure. Seules quelques connaissances algébriques élémentaires sont réclamées, ainsi que des notions rudimentaires sur les relations binaires, l'algèbre de Boole, le calcul matriciel, le calcul au sein de la droite numérique achevée et le calcul des probabilités, ces cinq matières étant d'ailleurs résumées dans les annexes.

L'objectif essentiel de l'ouvrage est de permettre au lecteur de résoudre des problèmes réellement rencontrés (ou qui pourraient l'être) en gestion. Pour l'y aider efficacement, de nombreux exemples sont entièrement traités dans le corps même du texte. De plus, de multiples exercices et problèmes à résoudre clôturent chaque chapitre : pour tous les énoncés proposés, la réponse est indiquée en fin de volume.

Introduction à la mathématique financière (200 exercices résolus - 50 cas réels différents traités en détail) par D. JUSTENS et J. ROSOUX.
Trosième édition, De Boeck (Université), 442 pages, 1995

Le mathématicien ne peut pas rester indifférent aux théories financières car, comme tout citoyen, il est confronté à des problèmes d'argent qu'il doit résoudre en connaissance de cause, et, en tant que scientifique, il peut trouver dans les activités financières de fort belles applications de mathématiques (les théories exploitées pouvant être fort sophistiquées).

Nous nous réjouissons donc de la parution de cet ouvrage qui est, assurément, un des rares ouvrages, écrit en langue française, qui présente la matière d'une manière claire, précise et rigoureuse et qui, de plus, offre le grand avantage de bien montrer comment les concepts théoriques sont utilisés dans la résolution de problèmes concrets et réels. Ajoutons à cela la présence de quelques résultats mathématiques originaux (par exemple, quelque jolies propriétés de la moyenne arithmétique de puissances d'un nombre réel positif), et leur exploitation inédite en finance (par exemple, pour justifier en détail la récente méthode proposée par le législateur belge pour calculer le TAEG).

Le texte présente toutes les notions fondamentales de la mathématique financière à savoir l'intérêt et l'escompte simples, l'intérêt composé, la capitalisation mixte, la capitalisation continue, les valeurs actuelle et acquise d'une suite de capitaux, ainsi que les notions de tableaux d'amortissements, d'usufruit et de nue-propriétés.

Comme l'indique le sous-titre de l'ouvrage, toutes les théories exposées sont illustrées par le traitement détaillé et très critique de cas réels, ainsi que pour de nombreux exercices proposés (avec les réponses finales).

De nombreux programmes informatiques (proposés en langage LDA et, pour les principaux, traduits en Q-Basic) sont également fournis pour permettre à tous les utilisateurs potentiels de résoudre leurs problèmes spécifiques.

Tout professeur de mathématique trouvera dans ce livre de nombreuses idées pour enrichir et illustrer ses cours.

J.BAIR

La loi de Poisson. La loi exponentielle et . . . la loi des séries de catastrophes

J. Wilmet,

La variable binomiale et celle de POISSON permettent déjà d'étudier beaucoup de situations de la vie courante. Cependant, on pourrait dire que le plus intéressant vient après. La variable exponentielle est une variable continue, modèle de l'intervalle de temps séparant deux événements accidentels. Son étude permet, par exemple, d'estimer la durée de vie de composants industriels mais aussi de savoir ce qu'il faut penser de la fameuse loi (maléfique !) des séries de catastrophes si souvent évoquées par les médias.

1. Rappels

Il peut être utile de faire quelques rappels avant d'aborder notre sujet principal : "la variable exponentielle et les séries de catastrophes". (Signalons de suite que ce mot doit être pris au sens du langage courant.)

Schéma de BERNOULLI

Le 33 sortira-t-il cette semaine au Lotto ? Vais-je réaliser un six en lançant le dé ? Parmi les centaines de pièces usinées qui sortent d'une chaîne de fabrication, celle que je vais tirer au hasard sera-t-elle bonne ou défectueuse ?

Tous ces exemples s'inscrivent dans un même schéma :
h : considérer une certaine expérience pouvant amener deux résultats contradictoires : r, appelé succès (parfois à tort !) de probabilité p et $\not{r}$, dit échec, de probabilité $1 - p = q$.

Ainsi, dans le premier exemple, h consiste à tirer les six numéros gagnants du Lotto (ne tenons pas compte du résultat complémentaire), r à ce que le 33 soit parmi les six et p vaut 1/7.

Variable de BERNOULLI

A ce schéma, on peut associer la variable X qui prend la valeur 1 en cas de succès r et la valeur 0 en cas d'échec $\not r$. De là,
la moyenne

$$\bar{X} = p \times 1 + q \times 0 = p$$

et la variance

$$\begin{aligned} \sigma^2(X) &= p(1-p)^2 + q(0-p)^2 \\ &= pq(q+p) = pq \end{aligned}$$

Variable binomiale

Partons d'un schéma de BERNOULLI et réalisons n épreuves de l'expérience h de manière indépendante. La variable binomiale :

$$x = \text{le nombre de succès } r \text{ dans ces } n \text{ épreuves.}$$

L'ensemble des résultats possibles pour x :

$$E = \{0, 1, \ldots, n\}$$

$$P_r(x = i) = C_n^i p^i q^{n-i}.$$

La variable binomiale x apparaît comme la somme de n variables de BERNOULLI indépendantes (une par épreuve). De là, la moyenne et la variante de x :

$$\begin{aligned} \bar{x} &= np \\ \sigma^2(x) &= npq. \end{aligned}$$

Notation : $x = B(n, p)$.

Variable de POISSON

Supposons maintenant que n tende vers l'infini, que p tende vers 0, le produit np restant toutefois constant : $np = \lambda > 0$.

Dans ces conditions,

— l'ensemble des résultats possibles devient $\mathbb{N}$

— pour i fixé,

$$\lim_{\substack{(n,p)\to(\infty,0)\\ np=\lambda}} P_r\ (x=i)$$

$$= \lim_{\substack{(n,p)\to(\infty,0)\\ np=\lambda}} C_n^i p^i q^{n-i}$$

$$= \lim_{n\to\infty} \frac{1}{i!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{i-1}{n}\right)\lambda^i\left(1-\frac{\lambda}{n}\right)^{n-i}$$

$$P_r(i) = \frac{\lambda^i}{i!}\,e^{-\lambda} \quad (i\in\mathbb{N})$$

c'est la probabilité de POISSON.

— à la limite, on obtient encore : $\bar{x} = \sigma^2(x) = \lambda$.

Dans la pratique, on supposera que n est assez grand, p assez petit, np étant constant. On remarque d'autre part que la masse probabiliste est concentrée dans [3] $[0, 3\lambda]$.

Notation : $x = P(\lambda)$.

Cette situation se rencontre couramment dans la vie de tous les jours, comme le montrent les exemples suivants.

Exemple 1.

Une compagnie d'assurances couvre le risque "incendie", pour 300.000 maisons. En moyenne, un logement sur 1200 brûle par an. Caractériser la variable :

$$x = \text{nombre annuel de sinistres.}$$

Construisons notre schéma de BERNOULLI : h consiste à observer une maison pendant un an. $x = 1$ si elle brûle. $p = 1/1200$.

On suppose que p est approximativement constant d'une maison à l'autre et que les logements subissent ou pas un incendie indépendamment les uns des autres. Alors,

$$\begin{aligned} x &= B(300.000;\ 1/1200) \\ &\cong P(250). \end{aligned}$$

Exemple 2.

Une chaîne de fabrication produit 2000 pièces à l'heure. Chaque pièce a une probabilité de 0,003 de présenter une défectuosité (non due à un disfonctionnement de la machine).

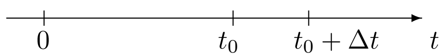
Le nombre de pièces défectueuses par heure est une variable de POISSON $P(6)$.

Ces deux exemples nous montrent qu'une variable de POISSON $P(\lambda)$ peut souvent être interprétée comme un nombre d'événements accidentels se produisant par unité de temps, λ étant le nombre moyen de tels événements.

2. Variable exponentielle

Les variables $B(n, p)$ et $P(\lambda)$ sont discrètes et sont étudiées dans l'enseignement secondaire. La variable exponentielle est une variable continue. Dans la pratique, elle représentera l'intervalle de temps séparant deux phénomènes accidentels.

Si nous disposons d'un chronomètre assez fin, nous pouvons admettre que deux événements ne sont jamais simultanés. L'ensemble des résultats possibles pour la variable exponentielle t est donc $\mathbb{R}_+^*$



0 est l'instant où le dernier phénomène accidentel s'est produit ; t indique le moment où l'accident *suivant* se produit, c'est-à dire aussi l'intervalle de temps séparant les deux instants.

Nous devons rechercher la densité de probabilité f de t ou encore la fonction de répartition F.

Pour rappel

$$F(t_0) = P_r\ (t \leqslant t_0) = \int_0^{t_0} f(t)dt.$$

Partons de l'hypothèse : λ est constant. Cela revient à dire que la probabilité qu'un accident (*un quelconque*, pas nécessaireent le suivant) se produise dans l'intervalle $]t_0, t_0 + \Delta t]$ est indépendante de t_0 mais est proportionnelle à la durée d'observation : si Δt est assez petit,

$$P_r \text{ (un accident dans }]t_0, t_0 + \Delta t]) = k\Delta t,$$

k étant constant. Dans le langage courant, cela revient à dire qu'il n'y a pas d'heure de pointe.

D'autre part, il est évident que la variable t a pour moyenne :

$$\bar{t} = \frac{1}{\lambda}.$$

Pour trouver F, calculons, de deux manières, la probabilité que *l'accident suivant* se produise dans $]t_0, t_0 + \Delta t]$:

— $(t \in]t_0, t_0 + \Delta t]) \Leftrightarrow$ (aucun accident dans $]0, t_0]$) et (un accident dans $]t_0, t_0 + \Delta t]$)

$P_r(t \in]t_0, t_0 + \Delta t]) = (1 - F(t_0)) \times k\Delta t$

— $]t_0, t_0 + \Delta t] =]0, t_0] \cup]t_0, t_0 + \Delta t]$

$F(t_0 + \Delta t) = F(t_0) + P_r(t \in]t_0, t_0 + \Delta t])$.

De là

$$F(t_0 + \Delta t) - F(t_0) = (1 - F(t_0))k\Delta t$$
$$\frac{F(t_0 + \Delta t) - F(t_0)}{\Delta t} = k(1 - F(t_0)).$$

A la limite, si $\Delta t \to 0$,

$$F'(t_0) = k(1 - F(t_0)).$$

(Rappelons que la fonction de répartition d'une variable continue est continûment dérivable).

F est donc solution de l'équation différentielle du premier ordre à variables séparées :

$$\frac{dF(t)}{1 - F(t)} = kdt.$$

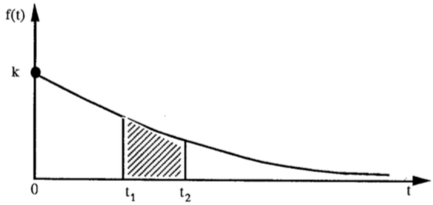
D'où

$$-\ln(1 - F(t)) = kt + cste$$

pour $t = 0$, il reste : $0 = cste$.

Finalement,

$$\begin{aligned} F(t) &= 1 - e^{-kt} \\ f(t) &= ke^{-kt} \end{aligned}$$



On obtient encore la probabilité que le phénomène accidentel suivant ait lieu en un temps t appartenant à $]t_1, t_2]$ après le précédent.

$$P_r(t \in]t_1 t_2]) = e^{-kt_1} - e^{-kt_2}$$

Valeur moyenne. Variance

$$\begin{aligned}\bar{t} &= \int_0^\infty tke^{-kt}dt = \frac{1}{k}\int_0^\infty ue^{-u}du \\ &= \frac{1}{k}\,\Gamma(2) = \frac{1}{k}.\end{aligned}$$

Nous avions trouvé plus haut : $\bar{t} = 1/\lambda$. Les paramètres k et t sont donc identiques. Désormais, on utilisera uniquement λ.

$$\begin{aligned}\sigma^2(t) &= \int_0^\infty t^2\lambda e^{-\lambda t}dt - \frac{1}{\lambda^2} \\ &= \frac{1}{\lambda^2}(\Gamma(3) - 1) = \frac{1}{\lambda^2}.\end{aligned}$$

En résumé, si le nombre de phénomènes accidentels par unité de temps est une constante λ, l'intervalle de temps séparant deux de ces phénomènes est

— une variable $t \in \mathbb{R}_+^*$

— dont la densité de probabilité est

$$f(t) = \lambda e^{-\lambda t}$$

— de moyenne $\bar{t} = 1/\lambda$ et de variance $\sigma^2(t) = 1/\lambda^2$.

3. Conséquences. Applications

Reprenons d'abord le premier exemple de variable de POISSON. Le nombre moyen d'incendies par an :

$$\lambda = 250.$$

Le temps moyen entre deux incendies :

$$E = \frac{1}{250} \cdot \text{an}.$$

Sa variance

$$\sigma^2(t) = \frac{1}{62.500}.$$

Si λ est grand, la variance de l'intervalle de temps est petite : les événements accidentels se produisent de manière assez régulière.

En revanche, considérons des événements beaucoup plus rares comme les tempêtes décennales, c'est-à-dire qui se produisent en moyenne tous les dix ans :

$$\lambda = \frac{1}{10} \qquad \bar{t} = 10 \qquad \sigma^2(t) = 100.$$

Pour les événements rares, la variance de t est donc grande, ce qui veut dire que ces événements peuvent se produire par rafales ! On peut observer plusieurs tempêtes décennales en cinq ou six ans mais attendre ensuite vingt, trente ou quarante ans avant d'essuyer la suivante.

Cette remarque explique sans doute le sentiment (la superstition ?) ancré dans la population que *les catastrophes se produisent par séries.*

Regardons les choses de manière plus précise en analysant la densité de probabilité de t dans le cas de catastrophes naturelles aux conséquences exceptionnellement graves. Par exemple, les tremblements de terre, inondations, éruptions volcaniques, vagues de chaleur,... faisant au moins trois cents morts.

Supposons

$$\lambda = 2 \text{ (catastrophes en moyenne par an)}$$

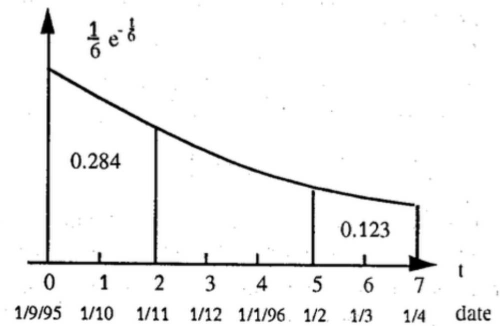
soit

$$\bar{t} = 6 \text{ mois}.$$

Si la dernière catastrophe s'est produite le 1er septembre 1995, demandez autour de vous vers quelle date se produira probablement la suivante. On vous répondra : "aux alentours du 1er mars 1996 !".

Pourtant, ce n'est pas la réponse donnée par le calcul des probabilités. En effet, choisissons le mois comme unité de temps. Alors,

$$f(t) = \frac{1}{6}\, e^{-\frac{1}{6}}$$



La probabilité que la catastrophe suivante ait lieu avant le 1er novembre 1995 vaut :

$$e^0 - e^{-\frac{2}{6}} \cong 0,284.$$

Celle qu'elle se produise entre le 1er février et le 1er avril 96 n'est plus que de :

$$e^{-\frac{5}{6}} - e^{-\frac{7}{6}} \cong 0,123.$$

En compensation, la probabilité pour qu'aucune autre catastrophe ne se produise avant le 1/9/97 vaut :

$$e^{-\frac{24}{6}} \cong 0,018.$$

Mais, dans ce cas, plus personne ne pensera aux catastrophes.

Application industrielle

La durée de vie d'un composant industriel est le laps de temps s'écoulant entre sa mise en service (c'est-à-dire la panne du précédent) et le moment où il tombe lui-même en panne. *A densité de panne constante*, c'est une variable exponentielle.

Une notion très importante est définie pour un tel composant : *sa fiabilité* après un temps t de fonctionnement, définie comme étant la probabilité que le composant soit encore en vie. Sa valeur est :

$$\text{fiabilité au temps } t = 1 - F(t) = e^{-\lambda t}.$$

Pour étudier, par exemple, la fiabilité du système de refroidissement d'une centrale nucléaire, il est nécessaire de prendre en compte plusieurs milliers de composants qui interviennent soit en série, soit en parallèle !

C'est encore la variable exponentielle qui intervient dans de nombreuses autres situations. Citons, sans que ce soit limitatif, la formation de files d'attente, le syndrome de l'auto-stoppeur, ...

4. Ce dont on n'a pas parlé

Tout au long de cet article, nous avons souligné l'importance de l'hypothèse : λ est constant.

Cette hypothèse restreint bien entendu le champ d'application de ce qui a été dit. Ainsi, la durée de vie du matériel qui s'use avec le temps (plaquettes de frein, ampoules électriques,...) ne suit pas une loi exponentielle. D'autre part, la variation possible de λ intéresse énormément des scientifiques (comme les climatologues) ou des entreprises (comme les compagnies d'assurances). Qu'on se rappelle simplement les articles parus dans la grande presse, au début de 1995, et soulignant que les primes d'assurance devraient être revues à la hausse car le nombre de catastrophes naturelles était en nette augmentation ! Pour ces sociétés, λ est donc croissant. Mais la vérification statistique de cette croissance éventuelle est une question difficile demandant de longues observations. Ce n'était pas l'objet de cet article !

Adresse de l'auteur :

Jean WILMET
rue du Bruliau 12
7120 Peissant

Revue des revues

P. Dalle Piane,

Math-Ecole (revue suisse), *n*°166- février 1995.

Au sommaire :

- **François Jaquet**, *Editorial*
- **André Calame,** *La conseillère fédérale Ruth Dreifuss interpelle les mathématiciens*
  L'article présente de larges extraits du discours d'ouverture prononcé par la Conseillère fédérale Ruth Dreifuss lors du Congès international des mathématiciens (ICM'94).
  Avant ce congrès, la conseillère avait envoyé trois questions à plus d'une douzaine de mathématiciens :
  1. comment les mathématiques pures peuvent-elles justifier leur art vis-à-vis de l'Etat qui les finance ?
  2. les mathématiques évitent-elles des discussions éthiques sur le rôle des mathématiques dans la société ?
  3. si la possibilité de créer dix nouvelles chaires dans les universités suisses existait, combien d'entre elles faudrait-il donner aux mathématiques et pourquoi ?

  L'article présente les réponses des mathématiciens.
- **Cosette Boillat-Juillerat,** *Des chiffres et des lettres ou déchiffrer des lettres*
  (Compte rendu des travaux d'un groupe du Colloque romand Mathématiques 93)
- **François Jaquet,** *Activité à géométrie variable*
  Compte rendu d'une activité organisée dans le cadre de la semaine de formation continue des maîtres de mathématiques valaisans. L'objectif de la séquence était une initiation au "problème ouvert". Le problème présenté était "taquins de pions". Le travail s'est déroulé en quatre étapes :
  – travail individuel,
  – travail par groupes de quatre participants,
  – débat en commun,
  – synthèse.
  La pratique du problème ouvert demande à celui qui la propose :
  – une analyse a priori,

– des observations en cours d'activité,
– une analyse a posteriori.
L'auteur présente les traces écrites de ces analyses. Cette "activité à géométrie variable" permet la mobilisation d'un vaste champ de connaissances (fonction affine, calcul des coefficients de son expression fonctionnelle, règles de calcul littéral, la démonstration par récurrence).

- **François Drouin,** *Jeu de l'oie, une activité ludique au collège*
  L'auteur a exploité une brochure de l'Irem de Lorraine décrivant comment fabriquer un jeu de l'Oie. Ce jeu peut être utilisé à l'école primaire et au collège.
- **François Jaquet,** *Moyens d'enseignement romands Mathématiques 1 à 4 : d'une édition à l'autre*
  Cet article est un extrait des futurs moyens d'enseignement romands de mathématiques pour l'école primaire.
- **Urs Moser,** *Que savent les élèves ? Etude du système scolaire suisse*
- **Luc-Olivier Pochon,** *Le jeu de GO*
  Les solutions des problèmes posés dans le numéro 164 de Math-écoles sont présentés.
- **Daniel Voiron,** *Mathématiques sans frontières*
  Ce championnat interclasses connaît un essor important en Europe et en Suisse en particulier. Des classes entières d'élèves de quinze et seize ans sont en lice. La classe s'organise pour résoudre en deux heures de douze à quinze problèmes et rend une réponse pour chacun d'eux. La solution d'un des problèmes doit être rédigée en langue étrangère. L'intérêt de cette épreuve réside dans la résolution de problèmes par groupes. Les énoncés sont originaux et attrayants. L'article présente neuf sujets de l'épreuve d'entraînement 94-95.
- *Courrier des lecteurs*

M. Frémal

APMEP (Association des Professeurs de Mathématiques de l'Enseignement Public-France), Bulletin $n°$399- juin 1995.

Ce bulletin est consacré, pour une bonne part, à la présentation des Journées Nationales APMEP à Grenoble, fin octobre 1995.

Ces journées avaient comme thèmes : Mathématiques : obstacle ou tremplin ?

En outre, on peut y lire

- Un article sur la réforme des classes préparatoires aux grandes écoles.
- Un court article, plutôt énumératif, sur l'écrit à l'Agrégation interne de Mathématique, de 1989 à 1994.
- Un autre article sur le même sujet.
- Dans le domaine de l'histoire des Mathématiques, vous y trouverez 2 textes. Le premier traite de "Quelques erreurs et supputations hasardeuses souvent répétées et parfois amplifiées" à propos des numérations chinoises. Le second est une réaction devant l'ouvrage de Georges Ifrah, "Histoire Universelle des Chiffres" et traite plus particulièrement de ce qui concerne la Mésopotanie.

  Cette livraison contient enfin les rubriques traditionnelles, "Nouvelles brèves", "Matériaux pour une documentation", "Avis de recherche" et "Les problèmes de l'A.P.M.E.P.".

APMEP (Association des Professeurs de Mathématiques de l'Enseignement Public-France), *n*° 400 - septembre 1995.

Comme le déclare le Président de l'APMEP, Jean-Paul Bardoulat, le 400ème numéro du Bulletin de l'APMEP méritait une attention toute particulière. Il est donc placé sous le thème "Qu'est-ce que faire des mathématiques ?".

Le Président termine son éditorial en rappelant que "l'activité mathématique est riche et variée, simple par les solutions qu'elle propose pour résoudre des problèmes complexes, ingrate quand on cherche, gratifiante lorsqu'on trouve.

C'est pour cette raison qu'elle suscite tant de réactions passionnées mais c'est aussi pour cela qu'elle est si formatrice ... pourvu qu'on s'y investisse suffisamment".

Tout naturellement, vous pourrez lire dans ce bulletin ...

- *En feuilletant le Bulletin (du n°1 au n°108)* par Paul Louis Hennequin.
- *Saisir l'Irrationnel : Dire, Montrer, Faire toucher, Tenir* par Evelyne Barbin.
- *Faire des Mathématiques* par D. Dacunha-Castelle.
- *La perte des sens, essence des Maths* par André Deledicq.
- *La mathématique, une invention* par Richard Pallascio.
- *La mathématique dans les baccalauréats généraux de quelques pays* par Pierre Legrand.

– *Les mathématiques à l'école* par Guy Brousseau.

Tous ces articles de fond, variés et documentés, ne peuvent que retenir l'attention du lecteur qui y trouvera peut-être des éléments de réponse à sa propre réflexion sur "Qu'est-ce que faire des Mathématiques ?".

Alors, bonne lecture !

Claude Villers

Mathematical Spectrum, Volume 26 $n°1$

– *Nombres premiers, primaires, factoriels et multifactoriels* par C. Caldwell et H. Dubner.

Il y a 2500 ans, Euclide démontra qu'il y a une infinité de nombres premiers en travaillant par l'absurde :

Supposons qu'il n'existe qu'un nombre fini de nombres premiers $p_1, p_2, \cdots p_n$.

Soit $N = p_1\, p_2 \cdots p_n + 1$.

N est $\neq 1$ et donc a au moins un diviseur premier k.

Mais k ne peut pas diviser $N - 1$, alors que, par hypothèse, tous les facteurs premiers divisent $N - 1$.

Il y a donc contradiction.

Les auteurs de l'article s'interrogent sur ces nombres $N = p_1\, p_2 \cdots p_n + 1$, qu'ils notent $p\sharp + 1$
où $p\sharp$ représente le produit de tous les nombres premiers $\leqslant p$. Ces nombres sont-ils toujours premiers, toujours composés, ou parfois composés ?

De même, qu'en est-il de $p\sharp - 1$, de $n! + 1$ et $n! - 1$?

Ces nombres ont été fréquemment étudiés :

Les nombres premiers de la forme $n! \pm 1$ sont appelés nombres premiers factoriels, et ceux de la forme $p\sharp \pm 1$, nombres premiers primaires.

Les auteurs résument les résultats connus et signalent qu'ils en ont ajoutés deux à la liste.

Ils généralisent la fonction factorielle en créant la multifactorielle (cf Mathematica) :

$$\begin{array}{lll} n!! & = n(n-2)\,(n-4)\ldots 1 & = n(n-2)!! \\ n!!! & = n(n-3)!!! & = n(n-3)\,(n-6)\ldots 1 \\ n!_k & = n(n-k)!_k & = n(n-k)\,(n-2k)\ldots 1 \end{array}$$

et étudient les nombres premiers ainsi obtenus.

Ils constatent sans pouvoir l'expliquer que les formes $n!!\pm 1$ n'engendrent qu'un nombre limité de facteurs premiers.

Ils utilisent les théorèmes suivants :

Théorème de Wilson : Soit p un entier > 1.
Alors p est premier ssi $(p-1)! \equiv -1(mod\ p)$ $\quad$ ($\Leftrightarrow$ ssi $(p-1)!+1$ est divisible par p).

Théorème de Fermat : Soit p un nombre premier ne divisant pas a.
Alors $a^{p-1} \equiv 1(mod\ p)$ $\quad$ ($\Leftrightarrow$ $a^{p-1}-1$ est divisible par p).

– *Extensions dans le domaine des racines carrées successives* par A. ABIAN et S. SVERCHKOV

Les auteurs donnent un développement de $1+x$, pour $x \in \mathbb{R}^+$, obtenu par itération infinie de racines carrées. Ils apportent la démonstration de ce résultat :

$$1+x = \lim_{n\to\infty} \sqrt{1+x\sqrt{1+(1+x)\sqrt{1+(2+x)\sqrt{\ldots\sqrt{1+(n+x)\sqrt{1+(n+1+x)}}}}}}.$$

Cela donne des résultats remarquables, en particulier pour x entier positif. Par exemple, si $x=1$, on obtient :

$$2 = \sqrt{1+\sqrt{1+2\sqrt{1+3\sqrt{1+4\sqrt{1+5\sqrt{1+\cdots}}}}}}.$$

– *Les nombres de l'Apocalypse* par C. Pickover

Le livre de l'Apocalypse de Saint Jean est le dernier du Nouveau Testament. Divers mystiques ont dépensé beaucoup d'énergie pour déchiffrer le nombre 666, utilisé par Saint Jean pour désigner le Nombre de la Bête, l'Antéchrist. Récemment encore, des mystiques de la droite fondamentaliste extrémiste des USA ont noté que chacun des mots du nom de Ronald Wilson Reagan comportait six lettres.

L'auteur a recherché par ordinateur, dans la suite de Fibonacci, un nombre apocalyptique, c'est-à-dire un nombre de 666 chiffres. C'est le cas du 3184ème qu'il cite in extenso.

Pour rappel, la suite de Fibonacci est

$$(1, 1, 2, 3, 5, 8 \ldots) \text{ telle que } t_n = t_{n-1} + t_{n-2} \quad \forall n > 2.$$

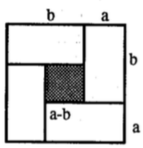
L'auteur pose alors diverses questions telles que :

— Existe-t-il d'autres nombres apocalyptiques dans la suite de Fibonacci ? Une note de l'éditeur montre qu'il en existe quatre autres.
— Existe-t-il un nombre apocalyptique premier ? Une seconde note de l'éditeur montre que oui.
— Est-ce une coïncidence si les 4 premiers chifrres et les 4 derniers qui composent le nombre apocalyptique donné constituent 1167 et 1163, c'est-à-dire deux dates du règne de Frédérik I d'Allemagne, qui intervint abondamment dans la politique papale ... ? Cette dernière question est restée sans réponse.

(Remarque : ma date de naissance n'est pas reprise dans les chiffres de ce nombre apocalyptique. Ouf !)

– *Une figure sous plusieurs aspects* par L. Short

Le but de l'auteur est de montrer l'unité sous-jacente aux mathématiques, en utilisant une même notion simple dans différents contextes.

a) 

Il utilise cette configuration pour démontrer géométriquement que :

* $(a + b)^2 - (a - b)^2 = 4ab$
* $\dfrac{a+b}{2} \geqslant \sqrt{ab}$ (l'égalité n'étant vraie que si $a = b$)

b) Opérons la substitution suivante :

$$\begin{aligned} a &= x_1 \qquad \text{avec } x_1 < x_2 \\ b &= x_2 \end{aligned}$$

Alors, si on augmente x_1 et diminue x_2 d'une même quantité positive d, on a que :

$$x_1 x_2 \leqslant (x_1 + d)(x_2 - d) \qquad \text{avec } 0 \leqslant d \leqslant \frac{x_2 - x_1}{2}.$$

Ce résultat conduit notamment à :

$$2 \times 10 \leqslant 3 \times 9 \leqslant 4 \times 8 \leqslant 5 \times 7 \leqslant 6 \times 6.$$

c) Soit $T_n = 1 + 2 + 3 + \cdots + n$.
T_n est appelé nombre triangulaire d'ordre n car on peut le représenter par

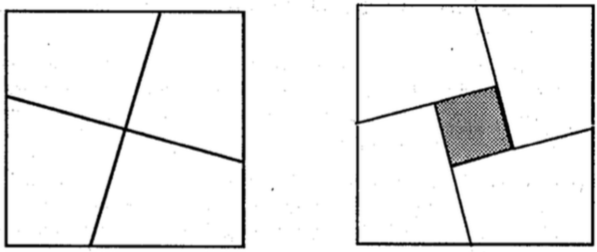
$$\begin{array}{lll} \bullet & & \\ \bullet & \bullet & \\ \bullet & \bullet & \bullet \\ & \vdots & \end{array}$$

L'utilisation de la même figure montre que :

$$8T_n + 1 = (2n + 1)^2$$

Ce résultat conduit à une partition du carré au moyen de rectangles tous égaux et d'un petit carré unitaire.

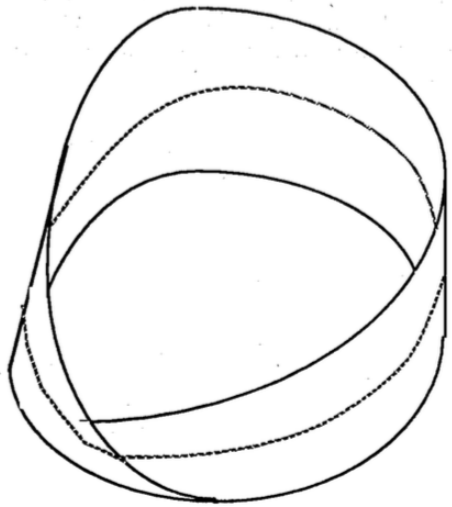
d) Si on découpe un carré suivant la figure (1) on obtient un puzzle de 4 pièces identiques qu'on peut disposer pour former le carré (2) d'apparence identique. Mais d'où vient le petit carré central ?



(1) (2)

L'auteur explique ce paradoxe grâce au théorème de Pythagore.

e) La figure ci-dessous représente un ruban de Möbius qui, comme on le sait, est une surface ayant un seul côté et une seule face et qu'on peut réaliser aisément au moyen d'un ruban.



Si on découpe ce ruban suivant la ligne pointillée et qu'on aplatit la surface obtenue, on obtient une figure semblable à celle qui est représentée en a).

f) On peut démontrer plusieurs identités algébriques ainsi que des propriétés relatives à la factorisation, et aux séries, par découpages géométriques.
L'idée de base est d'utiliser trois formats distincts de rectangles, que l'on combine différemment pour obtenir de grands rectangles. Cela permet aussi de "visualiser" la notion de variable, comme dimension d'un côté d'un rectangle.

g) L'auteur s'attarde sur le procédé de démonstration qui consiste à prouver que deux figures ont la même aire en procédant par dissection : R_1 et R_2 sont équivalentes par dissection si R_1 peut être découpé en pièces qui peuvent être rassemblées pour former R_2. De la sorte, il montre que, dans le *plan euclidien*, tout polygone peut êre découpé et rassemblé pour former un autre polygone de même aire. Ainsi, un octogone peut donner un carré, ou un hexagone, ou une croix grecque,... Mais dans l'espace, ou dans une géométrie non euclidienne, ceci n'est plus vrai. Par exemple, un tétraèdre régulier ne peut pas être découpé et rassemblé pour former un cube de même volume.

h) L'auteur donne aussi des exemples de pavage. Il termine en s'interrogeant sur l'extension possible de ses résultats à l'espace et donne une idée de l'exploitation de ceux-ci dans la recherche d'extrema, sans utilisation du calcul.

Pierrette DALLE PIANE

Olympiades

C. Festraets,

Voici les solutions des trois derniers problèmes proposés à l'Olympiade Internationale de Mathématique. Ils sont, à mon avis, nettement plus difficiles que les trois premiers.

4. *Trouver la plus grande valeur de x_0 pour laquelle il existe une suite $x_0, x_1, \ldots, x_{1995}$ de nombres réels strictements positifs vérifiant les deux conditions :*

(i) $x_0 = x_{1995}$

(ii) pour tout i, $1 \leqslant i \leqslant 1995$: $x_{i-1} + \dfrac{2}{x_{i-1}} = 2x_i + \dfrac{1}{x_i}$.

L'égalité

$$x_{i-1} + \frac{2}{x_{i-1}} = 2x_i + \frac{1}{x_i}$$

est équivalente à

$$2x_i^2 - x_i \left(x_{i-1} + \frac{2}{x_{i-1}} \right) + 1 = 0$$

ce qui donne

$$x_i = \frac{1}{2}\, x_{i-1} \quad (1) \quad \text{ou} \quad x_i = \frac{1}{x_{i-1}} \quad (2)$$

Pour tout i, $1 \leqslant i \leqslant 1995$, on a

$$x_i = 2^{k_i} x_0 \quad \text{ou} \quad x_i = 2^{k_i} \frac{1}{x_0}$$

avec k_i entier et $0 \leqslant |k_i| \leqslant 1995$.

En effet,

1) c'est vrai pour $i = 1$:

$$x_1 = \frac{1}{2}\, x_0 = 2^{-1} x_0 \quad \text{ou} \quad x_1 = \frac{1}{x_0} = 2^0 . \frac{1}{x_0}$$

2) si c'est vrai pour i, c'est vrai pour $i + 1$:

$$\text{on a soit} \quad x_{i+1} = \frac{1}{2}\, x_i \begin{cases} = \dfrac{1}{2}\, 2^{k_i} x_0 = 2^{k_i - 1} x_0 \\[2ex] = \dfrac{1}{2}\, 2^{k_i} \dfrac{1}{x_0} = 2^{k_i - 1} \dfrac{1}{x_0} \end{cases}$$

$$\text{soit} \quad x_{i+1} = \frac{1}{x_i} \begin{cases} = \dfrac{1}{2^{k_i} x_0} = 2^{-k_i} \dfrac{1}{x_0} \\ \text{ou} \\ = \dfrac{1}{2^{k_i} \frac{1}{x_0}} = 2^{-k_i} . x_0 . \end{cases}$$

On a donc

$$x_{1995} = 2^K . x_0 \quad \text{ou} \quad x_{1995} = 2^K . \frac{1}{x_0} \quad (\text{avec } K = k_{1995}).$$

Remarquons que, pour passer de x_0 à x_{1995}, on effectue K fois la transformation (1) et $(1995 - K)$ fois la transformation (2).

On aura

$$x_{1995} = 2^K . x_0$$

lorsque $(1995 - K)$ est pair, donc lorsque K est impair et alors la condition $x_{1995} = x_0$ ne peut être remplie car elle entraîne $2^K = 1$, ce qui est impossible pour K impair.

On aura

$$x_{1995} = 2^K . \frac{1}{x_0}$$

lorsque $(1995 - K)$ est impair, donc lorsque K est pair. Dans ce cas, la condition $x_{1995} = x_0$ implique $x_0^2 = 2^K$ et la plus grande valeur de x_0 est $x_0 = 2^{997}$; elle est obtenue pour $K = 1994$ et correspond à la suite

$$\left(x_0,\ x_1 = \frac{1}{2} x_0,\ x_2 = \frac{1}{2} x_1, \ldots, x_{1994} = \frac{1}{2} x_{1993},\ x_{1995} = \frac{1}{x_{1994}} \right)$$
$$= \left(2^{997}, 2^{996}, 2^{995}, \ldots, \frac{1}{2^{997}}, 2^{997} \right)$$

5. *Soit $ABCDEF$ un hexagone convexe tel que*

$$\begin{aligned} AB &= BC = CD \\ DE &= EF = FA \end{aligned}$$

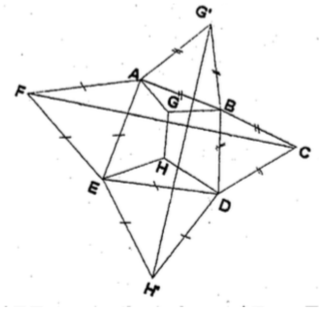
et

$$\measuredangle BCD = \measuredangle EFA = 60°$$

Soient G et H deux points intérieurs à l'hexagone tels que

$$\measuredangle AGB = \measuredangle DHE = 120°.$$

Montrer que $AG + GB + GH + DH + HE \geqslant CF$.



Le triangle AFE est équilatéral car $AF = EF$ et $\measuredangle EFA = 60°$; de même, le triangle BCD est équilatéral.

Construisons à l'extérieur du quadrilatère $ABDE$ les triangles équilatéraux $AG'B$ et $DH'E$. G est sur le cercle circonscrit au triangle $AG'B$ car $\measuredangle AGB + \measuredangle AG'B = 120° + 60° = 180°$; de même, H est sur le cercle circonscrit au triangle $DH'E$.

On a donc, en vertu du théorème de Ptolémée,

$$AB.GG' = AG.BG' + AG'.BG$$

d'où

$$\begin{aligned} GG' &= AG + BG. \\ DE.HH' &= DH.EH' + DH'.EH \end{aligned}$$

d'où

$$HH' = DH + EH.$$

Et il vient

$$\begin{aligned} AG + GB + GH + DH + HE &= GG' + GH + HH' \\ &\geqslant G'H'. \end{aligned}$$

Or la symétrie d'axe BE, médiatrice de $[AD]$, applique F sur H' et C sur G', donc $[FC]$ sur $[H'G']$. On a dès lors l'inégalité proposée.

6. *Soit p un nombre premier impair. Trouver le nombre de sous-ensembles A de l'ensemble $\{1, 2, \ldots, 2p\}$ tels que :*

(i) A contient exactement p éléments ;

(ii) la somme de tous les éléments de A est divisible par p.

Dans ce qui suit, on désignera par $s(E)$ la somme des éléments d'un ensemble E et par a' le reste du nombre entier a modulo p.

Soit $X = \{1, 2, \dots, p\}$ et $Y = \{p+1, p+2, \dots, 2p\}$.

$$\begin{aligned} s(X) &= 1 + 2 + \cdots + p = p.\frac{p+1}{2} \equiv 0 \ (mod\ p) \\ s(Y) &= (p+1) + (p+2) + \cdots + 2p = p.\frac{3p+1}{2} \equiv 0 \ (mod\ p). \end{aligned}$$

X et Y sont deux sous-ensembles satisfaisant les conditions (i) et (ii).

Construisons p sous-ensembles de p éléments de $\{1, 2, \dots, p\}$ de la manière suivante : on choisit arbitrairement n éléments $y_1, y_2, \dots, y_n$ de Y et $p - n$ éléments $x_1, x_2, \dots, x_{p-n}$ de X et on forme

$$\begin{aligned} A_0 &= \{y_1, y_2, \dots, y_n, x_1, x_2, \dots, x_{p-n}\} \\ A_1 &= \{y_1, y_2, \dots, y_n, (x_1+1)', (x_2+1)', \dots, (x_{p-n}+1)'\} \\ A_2 &= \{y_1, y_2, \dots, y_n, (x_1+2)', (x_2+2)', \dots, (x_{p-n}+2)'\} \\ &\vdots \\ A_{p-1} &= \{y_1, y_2, \dots, y_n, (x_1+p-1)', (x_2+p-1)', \dots, (x_{p-n}+p-1)'\} \end{aligned}$$

Si $s(A_0) \equiv \alpha \ (mod\ p)$, alors

$$\begin{aligned} s(A_1) &\equiv \alpha + (p-n).1 \ (mod\ p) \\ s(A_2) &\equiv \alpha + (p-n).2 \ (mod\ p) \\ &\vdots \\ s(A_{p-1}) &\equiv \alpha + (p-n)(p-1) \ (mod\ p) \end{aligned}$$

et, par le théorème de Gauss, ces restes modulo p sont, abstraction faite de l'ordre, égaux à $0, 1, 2, \dots, p-1$. Donc, un et un seul de ces p sous-ensembles satisfait les conditions (i) et (ii).

Le nombre de choix possibles des n éléments de Y et des $p - n$ éléments de X est

$$\binom{p}{n} \cdot \binom{p}{p-n} = \binom{p}{n}^2.$$

n peut prendre toutes les valeurs comprises entre 1 et $p-1$; il y a donc

$$\binom{p}{1}^2 + \binom{p}{2}^2 + \cdots + \binom{p}{p-1}^2 = \binom{2p}{p} - 2$$

sous-ensembles formés comme ci-dessus (on remarquera que l'on obtient tous les sous-ensembles de p éléments de $\{1, 2, \ldots, 2p\}$ sauf X et Y) et parmi ceux-ci

$$\frac{1}{p}\left(\binom{2p}{p} - 2\right)$$

ont la somme de leurs éléments divisible par p. Si, à ces sous-ensembles, on adjoint X et Y, le nombre total demandé est

$$\frac{1}{p}\left(\binom{2p}{p} - 2\right) + 2.$$

Des problèmes et des jeux

C. Festraets,

$\boxed{\text{Reste}}$ problème 166 de M. et P. $n°$ 103

Soit $g(x) = x^5 + x^4 + x^3 + x^2 + x + 1$.

Quel est le reste de la division de $g(x^{12})$ *par* $g(x)$ *?*

Solution de C. VAN HOOSTE de Marbaix-la-Tour.

On a

$$\begin{aligned} g(x^{12}) &= x^{60} + x^{48} + x^{36} + x^{24} + x^{12} + 1 \\ &= (x^{60} - 1) + (x^{48} - 1) + (x^{36} - 1) + (x^{24} - 1) \\ &\quad +(x^{12} - 1) + 6 \end{aligned} \tag{1}$$

Or, quel que soit $k \in \mathbb{N}_0$,

$$x^{6k} - 1 = (x^6 - 1)(x^{6(k-1)} + x^{6(k-2)} + \cdots + 1)$$

ce qui montre que $x^{6k} - 1$ est divisible par $x^6 - 1$.

Par conséquent, dans l'égalité (1), chaque terme mis entre parenthèses est divisible par $x^6 - 1$. Ainsi, le reste de la division de $g(x^{12})$ par $g(x)$ est égal à 6.

Généralisation de H.-J. SEIFFERT de Berlin.

Soit k et n des entiers positifs tels que $k \geqslant 2$ et n est divisible par k.

Les zéros $z_j = \exp(2\pi ji/k)$, $j = 1, 2, \ldots, k-1$, du polynôme

$$g(x) = x^{k-1} + x^{k-2} + \cdots + x + 1 = \frac{x^{k-1}}{x - 1}$$

sont simples et sont aussi des zéros du polynôme $g(x^n) - k$ car $z_j^k = 1$ pour $j = 1, 2, \ldots, k-1$; k est un diviseur de n, donc $z_j^n = 1$ et $g(1) = k$.

Dès lors, par le lemme de Gauss, on a

$$g(x^n) - k = g(x).q(x)$$

où $q(x)$ est un polynôme à coefficients entiers et comme $g(x)$ est de degré $k-1 \geqslant 1$, il s'ensuit que le polynôme constant $r(x)=k$ est le reste de la division de $g(x^n)$ par $g(x)$.

Dans le cas particulier proposé par l'énoncé, $k=6$, $n=12$ et le reste demandé est $r(x)=6$.

Bonnes solutions de P. DASSY de Liège, J. FINOULST de Diepenbeek, F. GLINEUR de Quiévrain, J. JANSSEN de Lambermont, B. LOISEAU de Mouscron et J. RONDOU de Heverlee.

$\boxed{\text{Intersection}}$ problème 167 de M. et P. $n°$ 103

Soit $K=\{a,b,c,d,e\}$ et F un ensemble comprenant 16 sous-ensembles de K tels que trois quelconques d'entre eux ont une intersection non vide.

Démontrer que tous les éléments de F ont au moins un élément en commun.

Solution de J. JANSSEN de Lambermont

Les sous-ensembles non vides de K se répartissent ainsi

$$\begin{cases} 5 & \text{singletons} \\ 10 & \text{paires} \\ 10 & \text{triplets} \\ 5 & \text{quadruples} \\ K & \end{cases}$$

Démontrons que F, qui comprend 16 de ces sous-ensembles, doit toujours comprendre un et un seul singleton, l'élément de celui-ci étant évidemment commun aux 16 sous-ensembles pour éviter les intersections vides.

1) F ne peut comprendre deux singletons puisque leur intersection est forcément vide.

2) Soit le singleton $\{a\}$. Il y a exactement 16 sous-ensembles de K qui comprennent a : $\{a\}$, 4 paires, 6 triplets, 4 quadruplets et K.

F peut donc comprendre ces 16 sous-ensembles.

3) Si F ne comprend aucun singleton, alors

– F ne peut pas comprendre 2 paires (dont l'intersection est $\{a\}$ par exemple), car on devrait alors avoir 14 autres sous-ensembles comprenant a, or il n'en reste plus que 13 !

– F ne peut contenir une paire unique, $\{a, b\}$ par exemple, car après avoir écarté le triplet $\{c, d, e\}$ qui lui est disjoint, on doit prendre tous les triplets restants y compris $\{a, d, e\}$ et $\{b, d, e\}$ qui vont, avec $\{a, b\}$, donner une intersection vide.

– si F ne comprend aucune paire, F doit comprendre les 16 sous-ensembles restants dont $\{a, b, c\}$, $\{a, d, e\}$ et $\{b, c, d\}$ qui ont une intersection vide.

Généralisation par B. LOISEAU de Mouscron

Soit K un ensemble à n éléments ($n > 0$) et F un ensemble formé de 2^{n-1} sous-ensembles de K tels que trois quelconques d'entre eux ont une intersection non vide. Alors, tous les éléments de F ont exactement un élément commun, soit x_0, et $F = \{A \subseteq K \mid x_0 \in A\}$.

Rappelons les définitions suivantes :

1) un ensemble F non vide de parties de K est un filtre sur K ss'il possède les propriétés suivantes :

(0) la partie vide de K n'appartient pas à F ;
(1) l'intersection de deux éléments de F appartient à F ;
(2) toute partie de K contenant un élément de F appartient à F.

2) un filtre F sur K est un ultrafiltre ss'il est maximal, c'est-à-dire ss'il n'existe aucun autre filtre sur K contenant F ; on montre que le filtre F est un ultrafiltre ssi

(3) pour tout $A \subseteq K$, ou bien $A \in F$, ou bien $K \setminus A \in F$.

Nous allons montrer que, dans les conditions de l'énoncé, c'est-à-dire F est formé de 2^{n-1} sous-ensembles de K tels que trois quelconques d'entre eux ont une intersection non vide, F est un ultrafiltre sur K.

La propriété (0) est triviale.

Montrons ensuite que F vérifie la propriété (3). Si $A \in F$, alors $K \setminus A \notin F$ car $A \cap (K \setminus A) = \emptyset$. L'injection

$$\mathcal{P}_K \to \mathcal{P}_K : A \mapsto K \setminus A$$

restreinte aux 2^{n-1} éléments de F nous fournit donc 2^{n-1} éléments de K n'appartenant pas à F. $F \cup \{K \setminus A \mid A \in F\}$ "couvre" donc $\mathcal{P}_K$ car $\mathcal{P}_K$ a 2^n éléments et on a bien pour tout $A \subseteq K$, ou bien $A \in F$, ou bien $K \setminus A \in F$.

Mais alors, si $A, B \in F$ et si $A \cap B \notin F$, par (3) $K \setminus (A \cap B) \in F$ et on aurait $A \cap B \cap (K \setminus (A \cap B)) \neq \emptyset$, ce qui est faux. La propriété (1) est donc vérifiée.

Et enfin, si $A \in F$ et $A \subseteq B$, alors $B \in F$ car sinon $K \setminus B \in F$, et on aurait $A \cap (K \setminus B) \neq \emptyset$, ce qui est impossible puisque $A \subseteq B$. Ce qui démontre la propriété (2).

F est bien un ultrafiltre sur K. Or K est un ensemble fini et tout ultrafiltre sur K est un ultrafiltre principal engendré par un élément x_0 de K. On a donc

$$F = \{A \subseteq K \mid x_0 \in A\}.$$

De bonnes solutions ont été envoyées par J. FINOULST de Diepenbeek, F. GLINEUR de Quiévrain et C. VAN HOOSTE de Marbaix-la-Tour.

$\boxed{\text{Fonction élémentaire}}$ problème 168 de M. et P. n° 103

Déterminer toutes les fonctions f de $\mathbb{R}$ dans $\mathbb{R}$, définies pour tout réel x et telles que

$$\forall x,\ y \in \mathbb{R} : f(x^2 - y^2) = f^2(x) - f^2(y).$$

Solution de C. VAN HOOSTE de Marbaix-la-Tour.

1) *Généralités*

a) $f(0) = 0$.
De fait, x étant un réel quelconque, on a

$$f(0) = f(x^2 - x^2) = f^2(x) - f^2(x) = 0.$$

b) Pour tout réel x, $f(x^2) = f^2(x)$.

$$f(x^2) = f(x^2 - 0) = f^2(x) - f^2(0) = f^2(x).$$

c) Pour tout réel x, $f(-x^2) = -f^2(x)$.

$$f(-x^2) = f(0 - x^2) = f^2(0) - f^2(x) = -f^2(x).$$

d) La fonction f est positive dans $\mathbb{R}^+$ et négative dans $\mathbb{R}^-$.
C'est une conséquence des deux propriétés précédentes.

e) La fonction f est impaire.
De fait, en rapprochant les propriétés b) et c), on obtient

$$f(-x^2) = -f(x^2)$$

pour tout réel x.

f) Soit $f(1) = 0$, soit $f(1) = 1$.
En effet, d'après la propriété b), on a

$$f(1) = f(1^2) = f^2(1).$$

2) *Etude du cas* $f(1) = 0$

a) Pour tout réel x, $f(x-1) = f(x)$.
Si x est positif, il existe un réel y tel que $x = y^2$. Ainsi, on a

$$f(x-1) = f(y^2-1) = f^2(y) - f^2(1) = f(y^2) = f(x).$$

Si x est négatif, on a

$$f(x-1) = -f(-x+1) = -f(-x) = f(x)$$

en tenant compte que la fonction f est impaire et que $1-x$ est positif.

b) Pour tout réel $x \in [0,1[$, $f(x) = 0$.
D'une part, comme x est positif, on a $f(x) \geqslant 0$.
D'autre part, par la propriété 2a) et par le fait que $x-1$ est négatif, on a aussi

$$f(x) = f(x-1) \leqslant 0.$$

Il s'ensuit que $f(x) = 0$.

c) Pour tout réel x, $f(x) = 0$.
De fait, il existe un entier n tel que $n \leqslant x < n+1$. Dès lors, grâce aux deux propriétés précédentes, on a

$$f(x) = f(x-1) = f(x-2) = \cdots = f(x-n) = 0$$

car $0 \leqslant x-n < 0$.
Cette propriété exprime que f *est la fonction nulle.*

3) *Etude du cas* $f(1) = 1$.

Considérons la fonction g telle que, pour tout réel x, on a

$$g(x) = f(x) - x.$$

En premier lieu, quels que soient les réels x et y, on a

$$\begin{aligned} g(x^2-y^2) &= f(x^2-y^2)-(x^2-y^2) \\ &= f^2(x)-f^2(y)-x^2+y^2 \\ &= (f(x^2)-x^2)-(f(y^2)-y^2) \\ &= g(x^2)-g(y^2). \end{aligned}$$

En second lieu, on a

$$g(1)=f(1)-1=0.$$

Dès lors, g est une fonction de la même famille que f. De plus, son image en 1 étant 0, elle s'apparente au cas étudié précédemment. Par conséquent, g est la fonction nulle et, par la suite, *f est la fonction identique.*

4) *Conclusion*

Seules la fonction nulle et la fonction identique vérifient la condition imposée.

Bonnes solutions de F. GLINEUR de Quiévrain, de B. LOISEAU de Mouscron et de H. SEIFFERT de Berlin. Un autre lecteur m'envoie une solution partielle, il a en effet supposé dans sa démonstration que f est une fonction polynôme, ce qui n'est pas une donnée de l'énoncé.

175. Preuve par $a+b+c$

Dans un triangle ABC, l'angle A est double de l'angle B, l'angle C est obtus et les longueurs a, b, c des trois côtés sont des nombres entiers (positifs, non nuls). Déterminer le plus petit périmètre de ce triangle.

176. Echec et math

On consière un échiquier $m \times n$. Pour quelles valeurs de m et de n est-il possible de le recouvrir entièrement avec des dominos 4×1 ?

177. Ha ha !

Déterminer le plus petit entier positif a tel que $11^{3n+1}+a.5^n$ soit divisible par 13 et $23^{2n+1}+2^{n+a}$ soit divisible par 31, quel que soit n entier positif.

Bibliographie

G. Haesbroeck,

La pensée hors langage : à la rencontre d'adolescents autistes par Frédérique et Georges PAPY, Adriana SCHULER. Bayard Editions, Paris, 1995, 258 pages

Ce livre relate deux expériences différentes et complémentaires impliquant chacune un adolescent autiste sans langage.

La première approche, celle de la psycholinguiste A. Schuler, est évaluative : grâce à un protocole d'examen sophistiqué, des anomalies spécifiques du traitement des informations sont repérées chez un jeune patient.

La deuxième approche, réalisée par Frédérique et Georges Papy, est pédagogique. Les mathématiciens font découvrir à un jeune autiste les mathématiques et leur application. Ils montrent que l'apprentissage des mathématiques a permis à un jeune garçon de donner un sens à sa perception du monde et que la découverte du nombre peut tenir lieu d'agent de la mutation de la pensée à la place du langage courant.

Ce livre offre ainsi un sujet de réflexions à tous ceux qui s'intéressent à la pensée, à ses troubles et aux moyens de les évaluer et de les traiter.

J. BAIR

Analyse Infinitésimale : Pour HEC et ingénieurs commerciaux par Camille DEBIEVE et Yves FELIX. Bibliothèque des universités : série mathématiques, 240 pages

Ce cours d'analyse mathématique traite en 5 chapitres l'étude des fonctions à une variable réelle, et présente en 2 chapitres complémentaires les éléments nécessaires à une bonne introduction aux fonctions à plusieurs variables. Présentation et style sont clairs et rigoureux. Les théorèmes essentiels sont démontrés, interprêtés, illustrés par des exemples et agrémentés de représentations graphiques.

Chaque chapitre reprend une panoplie d'exercices classiques, avec suggestions ou solutions. Bref, un bon livre rassemblant les notions d'analyse nécessaires au seuil de l'enseignement supérieur.

Il est cependant regrettable que seul le chapitre consacré à la dérivation reprenne des exemples économiques. Vu le public spécifique visé, il aurait été intéressant et très utile que ce manuel présente, tant en théorie qu'en exercices, d'avantage d'applications économiques et commerciales.

G. HAESBROECK