Jean-Paul Delahaye

Optimisation impossible

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 19, n° 2 (1985), p. 213-233

<http://www.numdam.org/item?id=M2AN_1985__19_2_213_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(vol 19, n° 2, 1985, p 213 à 233)

OPTIMISATION IMPOSSIBLE

par Jean-Paul DELAHAYE (*)

Communiqué par F ROBERT

Résumé — *En utilisant le seul fait que dans un algorithme d'optimisation chaque point généré l'est apres un nombre fini d'evaluations de la fonction economique et de ses derivees partielles, on montre que pour certaines familles de fonctions aucun algorithme d'optimisation ne peut exister Les résultats assez élémentaires établis dans ce travail constituent des sortes de démonstrations de principes intuitivement connus des optimiseurs*

Abstract — *Using only that every point generated by an optimization algorithm needs the evaluation of the economic function or of its derivatives on a finite number of points, it is shown that for certain families of functions no optimization algorithm can exist Established results are rather elementary and may be considered as mathematical proofs of well known principles of optimization theory.*

INTRODUCTION

Il est bien clair pour quiconque s'est posé quelques problèmes d'optimisation qu'en ce domaine tout n'est pas possible. Par exemple : il n'est pas raisonnable d'espérer construire un algorithme d'optimisation qui puisse pour toute fonction définie sur $\mathbb{R}$ atteignant son maximum, engendrer une suite convergente (x_n) dont la limite x vérifierait $f(x) = \max_{t \in \mathbb{R}} f(t)$; on sent bien que savoir d'une fonction qu'elle atteint son maximum n'est pas suffisant, il faudrait au moins savoir qu'elle est continue ; c'est intuitivement évident.

Il est bien clair..., il n'est pas raisonnable d'espérer..., on sent bien que..., c'est intuitivement évident... Oui, mais comment donner un sens rigoureux à tout cela ? Comment démontrer mathématiquement des résultats qui justifieraient ces impressions ?

C'est là le but de ce travail dans lequel, délibérément, nous chercherons uniquement à établir quelques principes élémentaires d'optimisation.

(*) Reçu en mars 1984

([1]) Université des Sciences et Techniques de Lille, U E R d'I E E A Laboratoire d'Informatique Fondamentale de Lille (L A 369 CNRS) 59655 Villeneuve d'Ascq Cedex

Ces principes n'étonneront pas ceux qui sont préoccupés d'optimisation, ils les connaissent et même les utilisent pour guider leurs recherches, construire leurs algorithmes et choisir les hypothèses de leurs théorèmes ; cette connaissance intuitive et parfois imprécise va être basée sur des résultats rigoureusement démontrables. D'autres principes plus fins pourront peut-être grâce aux méthodes proposées ici, être énoncés et établis ultérieurement.

Le premier problème, bien sûr, est de fixer les définitions (§ 1) et en particulier la définition d'algorithme d'optimisation, c'est essentiel car notre définition ne doit être ni trop large (nous ne pourrions pas établir de résultats de limitation) ni trop stricte (on nous accuserait de fournir une modélisation incomplète et nos résultats seraient sans intérêt). La méthode que nous utilisons pour définir ce que nous appelons générateurs de suites pour problèmes d'optimisation est inspirée de travaux concernant les transformations de suites [3, 4, 5, 6, 7]. Cette méthode, tout en évitant les lourds outils de la théorie des fonctions récursives (qui, au premier abord, est la voie naturelle pour démontrer des résultats de limitation concernant des algorithmes [1, 2, 9]) permet d'obtenir relativement facilement des résultats intéressants.

Le § 2 contient les résultats principaux de ce travail, c'est-à-dire la démonstration d'un certain nombre de principes élémentaires d'optimisation.

Le § 3 propose des généralisations et extensions de la méthode adoptée et fournit (sans démonstration) quelques résultats complémentaires.

I. DÉFINITIONS

Soit E un espace métrique.

On notera $\mathscr{A}(E)$ l'ensemble de toutes les applications de E dans $\mathbb{R}$, et $\mathscr{M}(E)$ l'ensemble de toutes les applications de E dans $\mathbb{R}$ possédant un maximum, c'est-à-dire telles que :

$$\exists \hat{x} \in E, \quad \forall x \in E : f(\hat{x}) \geqslant f(x) .$$

Pour tout $f \in \mathscr{M}(E)$, nous poserons

$$\operatorname{opt}(f) = \{ x \in E \mid \forall y \in E \ f(x) \geqslant f(y) \} .$$

Soit $f \in \mathscr{M}(E)$, nous envisagerons les deux problèmes d'optimisation suivants :

— Problème de la recherche du maximum pour f :

$$\left.\begin{array}{l} \text{trouver } M \in \mathbb{R} \text{ tel que :} \\ M = \max \{ f(x) \mid x \in E \} \end{array}\right\} . \tag{a}$$

— Problème de la recherche d'une solution optimale pour f :

$$\left.\begin{array}{l}\text{trouver } \hat{x} \in E \text{ tel que :} \\ f(\hat{x}) = \max \{ f(x) \mid x \in E \} \quad (\text{i.e. } \hat{x} \in \operatorname{opt}(f))\end{array}\right\} \qquad (b)$$

Remarque : Bien d'autres problèmes d'optimisation pourraient être envisagés. Par exemple :

- Pour $F \subset E$, trouver $M = \max \{ f(x) \mid x \in F \}$;
- Trouver tous les $x \in E$ tels que : $f(\hat{x}) = \max \{ f(x) \mid x \in E \}$;
- Trouver $\hat{x} \in E$ tel qu'il existe un voisinage V de x, pour lequel : $f(\hat{x}) = \max \{ f(x) \mid x \in V \}$;
- Trouver $\hat{x} \in E$ tel que $| f(\hat{x}) - \max \{ f(x) \mid x \in E \} | \leqslant \varepsilon$; etc...

En réalité, en optimisation, on recherche des méthodes pouvant être efficaces pour plusieurs fonctions.

Une formulation plus exacte des problèmes (a) et (b) est donc : Soit $\mathcal{S} \subset \mathcal{M}(E)$.

— Problème de la recherche du maximum sur $\mathcal{S}$:

$$\left.\begin{array}{l}\text{Pour toute } f \in \mathcal{S} \text{ trouver } M(f) \in \mathbb{R} \text{ tel que :} \\ M(f) = \max \{ f(x) \mid x \in E \}\end{array}\right\} \qquad (A)$$

— Problème de la recherche d'optimum sur $\mathcal{S}$:

$$\left.\begin{array}{l}\text{Pour toute } f \in \mathcal{S} \text{ trouver } x(f) \in \mathbb{R} \text{ tel que :} \\ f(x(f)) = \max \{ f(x) \mid x \in E \} \quad (x(f) \in \operatorname{opt}(f))\end{array}\right\} \qquad (B)$$

« Trouver α » ne signifie pas « calculer exactement α » (car on sait bien que le plus souvent cela n'est pas réaliste, ou même n'a pas de sens) mais signifie seulement « proposer une suite (α_n) convergeant vers α » (ou même « proposer une suite (α_n) dont toutes les valeurs d'adhérence soient solutions du problème », mais nous n'envisagerons ce point de vue qu'au § 3).

On cherche donc une méthode de calcul qui, pour toute fonction $f \in \mathcal{S}$ construise (génère) une suite $M_n(f)$ (problème (A)) ou $x_n(f)$ (problème (B)), convergeant vers $M(f)$ ou vers $\hat{x}(f) \in \operatorname{opt}(f)$. Pour tout ensemble X nous désignons par $X^{\mathbb{N}}$ l'ensemble des suites infinies d'éléments de X.

DÉFINITION : *On appelle* générateur de suites (*ou plus simplement* générateur) pour le problème (A) (*resp.* le problème (B)) *sur* $\mathcal{S}$, *toute application* M (*resp.* x)

$$M : \mathcal{S} \to \mathbb{R}^{\mathbb{N}} \; ; \; f \mapsto (M_n(f))$$
$$x : \mathcal{S} \to E^{\mathbb{N}} \; ; \; f \mapsto (x_n(f)) \, .$$

On remarquera que l'image de f par M est notée $(M_n(f))$, la notation $M(f)$ étant réservée pour désigner le maximum de f lorsque $f \in \mathcal{M}(E)$.

Nous dirons que le générateur M (resp. x) est *convergent* sur $\mathcal{S}$ si :

pour tout $f \in \mathcal{S}$ la suite $M_n(f)$ (resp. $(x_n(f))$) converge vers $M(f)$ (resp. vers un $x \in \operatorname{opt}(f)$).

Il est bien clair cependant que pour qu'un générateur soit utilisable, il faut qu'il soit « algorithmique », (« calculable »), c'est-à-dire il faut que les différents points de la suite $M_n(f)$ ou $(x_n(f))$ puissent être obtenus successivement, chacun n'utilisant qu'une information finie sur f (par exemple des évaluations de f en un nombre fini de points) cette information subissant un traitement numérique fini.

Un de nos buts étant d'obtenir des résultats de limitation (impossibilité de la résolution du problème (A) ou (B) sur la classe $\mathcal{S}$) il nous faut donc exprimer cette condition d'« algorithmicité » (de « calculabilité »).

Nous choisissons volontairement de n'exprimer cette condition que partiellement (les discussions nécessaires pour l'exprimer entièrement pourraient d'ailleurs être l'objet de livres entiers !).

Cette attitude ne fera que donner plus de force à nos résultats négatifs car tous les problèmes que nous allons montrer être impossibles à résoudre avec notre notion de générateur le sont a fortiori avec une notion de générateur exprimant plus strictement la condition d'« algorithmicité ».

DÉFINITION : *Nous dirons que le générateur M (resp. x) pour le problème (A) (resp. pour le problème (B)) est à information finie d'ordre* 0 *sur $\mathcal{S}$ si* :

$$\left.\begin{array}{l}\forall f \in \mathcal{S}\,, \quad \forall n \in \mathbb{N}\,, \quad \exists k(n) \in \mathbb{N}\,, \quad \exists y_1^n, \dots, y_{k(n)}^n \in E\,, \quad \forall g \in \mathcal{M}(E) : \\ (f(y_1^n) = g(y_1^n), \dots, f(y_{k(n)}^n) = g(y_{k(n)}^n)) \Rightarrow M_n(f) = M_n(g) \ (\textit{resp. } x_n(f) = x_n(g))\,. \end{array}\right\} (*)$$

(On trouvera dans [5] la définition de transformation à information finie qui a servi de modèle à celle-ci).

Lorsque M (resp. x) est un générateur à information finie d'ordre 0 sur $\mathcal{S}$, pour chaque $f \in \mathcal{S}$ nous considérons la suite $(k(n))$ où $k(n)$ désigne le plus petit entier tel que :

$$\exists y_1^n, \dots, y_{k(n)}^n \in E\,; \quad \forall g \in \mathcal{M}(E) :$$
$$f(y_1^n) = g(y_1^n), \dots, f(y_{k(n)}^n) = g(y_{k(n)}^n) \Rightarrow M_n(f) = M_n(g) \text{ (resp. } x_n(f) = x_n(g))\,.$$

Nous choisissons aussi un $k(n)$-uplet tel que la relation soit vraie et nous notons :

$$X_n(f) = \{ y_i^j \mid j \in \{ 0, \dots, n \}, i \in \{ 1, \dots, k(n) \} \}$$
$$X(f) = \bigcup_{n \in \mathbb{N}} X_n(f)\,.$$

Avec ces notations, la condition (*) signifie que, pour toute fonction g qui coïncide avec f sur $X_n(f)$ (qui est fini) le générateur M (resp. x) réagit de la même façon que pour f : la connaissance de f sur $X_n(f)$ détermine $M_n(f)$, (resp. $x_n(f)$).

Un générateur à information finie d'ordre 0 est donc un générateur qui, pour chaque calcul de $M_n(f)$, (resp. $x_n(f)$) n'utilise qu'un nombre fini d'évaluations de f; ce nombre, et les points évalués (les y_i^j) pouvant dépendre de f.

Remarques

1) Lorsque E est un ouvert de $\mathbb{R}^n$ par exemple, on pourrait définir de manière analogue la notion de *générateur à information finie d'ordre h*, ($h \in \mathbb{N}$) en modifiant (*) de façon à exiger la coïncidence de f et de g ainsi que de leurs dérivées partielles jusqu'à l'ordre h sur $y_1^n, \ldots, y_{k(n)}^n$ pour qu'il y ait $M_n(f) = M_n(g)$.

Pour des raisons de simplicité dans les formulations et les démonstrations, nos résultats seront toujours énoncés à l'ordre 0 et nous indiquerons en remarque les généralisations possibles à l'ordre h.

A la place de générateur à information finie d'ordre 0, nous dirons *générateur à information finie.*

On peut dire que tous les algorithmes d'optimisation (une fois mis en forme et les paramètres fixés) constituent des générateurs à information finie d'ordre 0, 1, ou 2. C'est le cas pour la méthode de Newton, la méthode du gradient conjugué et ses différentes variantes.

2) Bien souvent déjà en optimisation des études générales de la notion d'algorithme ont été proposées [6, 10, 11, 13] mais en fait il s'agit dans ces travaux non pas de la formalisation de la notion d'algorithme mais de la définition de *schémas généraux d'algorithmes*; le but recherché étant d'obtenir simultanément des résultats de convergence et de semi-convergence pour des algorithmes différents. Les méthodes utilisées dans ces travaux sur les schémas généraux d'algorithmes (en particulier la notion de fonction multivoque) ne fournissent en aucune manière une formalisation de la notion d'algorithme « calculable » dont nous avons besoin pour établir des résultats d'impossibilité. C'est pourquoi nous n'avons pas cherché à utiliser les outils de cette branche de l'optimisation.

Certaines études sur la notion d'algorithmes optimaux [12] auraient pu servir de base à notre système de définition. Cependant nous avons préféré introduire la définition de générateur à information finie car elle évite de choisir un formalisme particulier et permet donc à nos résultats d'être valables quelle que soit la définition (« raisonnable ») utilisée pour formaliser la notion d'algorithme itératif infini d'optimisation.

II. RÉSULTATS

A partir des définitions simples données au § 1, nous allons maintenant énoncer et démontrer un certain nombre de résultats qui possèdent chacun une interprétation correspondant à un principe plus ou moins intuitif, souvent connu des optimiseurs.

a) Semi-continuité et recherche du maximum

Soit E un espace métrique, nous désignerons par $\mathcal{M}_{\text{s.c.i.}}(E)$ (resp. $\mathcal{M}_{\text{s.c.s.}}(E)$) l'ensemble des applications f de E dans $\mathbb{R}$ ayant un maximum ($f \in \mathcal{M}(E)$) et semi-continues inférieurement (resp. semi-continues supérieurement) c'est-à-dire telles que :

$$\forall x \in E, \quad \forall \varepsilon \in \mathbb{R}^{+*}, \quad \exists \eta \in \mathbb{R}^{+*}, \quad \forall y \in E :$$
$$d(x, y) \leqslant \eta \Rightarrow f(y) \geqslant f(x) - \varepsilon \,(\text{resp.}\, f(y) \leqslant f(x) + \varepsilon)$$

THÉORÈME 1 : *Soit E une partie compacte de* $\mathbb{R}^m$, *d'intérieur non vide :*

(i) *il existe des générateurs convergents à information finie pour le problème* (A) *sur* $\mathcal{M}_{\text{s.c.i.}}(E)$,

(ii) *il n'existe pas de générateurs convergents à information finie pour le problème* (A) *sur* $\mathcal{M}_{\text{s.c.s.}}(E)$.

Interprétation : Ce sont des conditions de semi-continuité inférieure qui rendent possible la recherche du maximum d'une fonction et non pas des conditions de semi-continuité supérieure.

Démonstration du théorème 1 (i) : Soit $\{ a_0, a_1, \ldots, a_n, \ldots \}$ une partie dénombrable dense de E (une telle partie existe toujours, il suffit de prendre $\mathbb{Q}^m \cap E$).

On pose $M_n(f) = \max \{ f(a_i) \mid 0 \leqslant i \leqslant n \}$.

Soit $(a_{\alpha(n)})_{n \in \mathbb{N}}$ une sous-suite de (a_n) convergeant vers $a \in \operatorname{opt}(f)$. On a :

$$\begin{aligned} M(f) \geqslant \limsup_{n \to \infty} M_n(f) &\geqslant \liminf_{n \to \infty} M_n(f) \\ &\geqslant \liminf_{n \to \infty} f(a_{\alpha(m)}) \\ &\geqslant f(a) = M(f) \end{aligned}$$

donc $M_n(f)$ converge vers $M(f)$.

Remarques

1) Le générateur proposé est bien sûr à information finie mais en fait il est beaucoup plus : concevoir un algorithme pour le calcul successif des $M_n(f)$ ne pose aucun problème ; quel que soit le sens (restrictif ou pas) qu'on donne au mot algorithme, notre générateur est donc algorithmique.

2) Le calcul des $M_n(f)$ ici nécessite « à la limite » l'évaluation de f sur une partie dense de E. Ceci rend inutilisable en pratique un tel générateur dès que la dimension de E dépasse quelques unités. Nous verrons plus loin (Théorème 3) que ce défaut pour $\mathcal{M}_{\text{s.c.i.}}(E)$ est irrémédiable.

3) La démonstration de (i) s'étend bien sûr à tout espace métrique séparable (i.e. possédant une partie dénombrable dense).

Démonstration du théorème 1 (ii) : Supposons qu'on ait un générateur M convergent à information finie pour le problème (A) sur $\mathcal{M}_{\text{s.c.s.}}(E)$.

Soit $f : E \to \mathbb{R}$; $x \mapsto f(x) = 0$. Bien sûr $f \in \mathcal{M}_{\text{s.c.s.}}(E)$.

Soit $X(f)$ l'ensemble des points utilisés par M pour le calcul de la suite $(M_n(f))$. L'ensemble $X(f)$ est dénombrable (voir la définition de $X(f)$ au § 1) et E ne l'est pas ; il existe donc $z \in E$, $z \notin X(f)$.

On définit $g : E \to \mathbb{R}$, en posant :

$$g(x) = 0 \quad \text{si} \quad x \neq z$$
$$g(z) = 1 \,.$$

L'application g est dans $\mathcal{M}_{\text{s.c.s.}}(E)$.

Puisque le générateur M est convergent sur $\mathcal{M}_{\text{s.c.s.}}(E)$, on a :

$$\lim_{n \to \infty} M_n(g) = M(g) = 1$$

$$\lim_{n \to \infty} M_n(f) = M(f) = 0 \,.$$

Ceci est impossible car par définition de $X(f)$ les deux suites $(M_n(f))$ et $(M_n(g))$ sont identiques.

Remarques

1) Cette première démonstration d'un résultat négatif est très simple puisque seules deux fonctions suffisent à « piéger » l'éventuel générateur. Les autres démonstrations de résultats négatifs seront moins simples.

2) La démonstration de (ii) s'étend bien sûr à tout espace métrique non dénombrable.

b) Recherche du maximum et recherche d'une solution optimale

Soit E un espace métrique, nous désignerons par $\mathcal{M}_c(E)$ l'ensemble des applications f de E dans $\mathbb{R}$ ayant un maximum ($f \in \mathcal{M}(E)$) et continues :

$$\mathcal{M}_c(E) = \mathcal{M}_{sci}(E) \cap \mathcal{M}_{scs}(E) .$$

THÉORÈME 2 : *Soit E une partie compacte de $\mathbb{R}^m$ d'intérieur non vide .*

(i) *il existe des générateurs convergents à information finie pour le problème (A) sur $\mathcal{M}_c(E)$,*

(ii) *il n'existe pas de générateurs convergents à information finie pour le problème (B) sur $\mathcal{M}_c(E)$.*

Interprétation : Il est évident que pour toute famille de fonctions $\mathcal{F} \subset \mathcal{M}_c(E)$, s'il existe un générateur convergent à information finie pour le problème (B) sur $\mathcal{F}$, alors il existe un générateur convergent à information finie pour le problème (A) sur $\mathcal{F}$.

On peut donc dire que : la recherche du maximum d'une fonction est un problème *strictement* plus facile que la recherche de points réalisant ce maximum.

Démonstration du théorème 2 (i) : Immédiat à partir du théorème 1 (i).

Démonstration du théorème 2 (ii) : Nous démontrons le résultat uniquement dans le cas $E = [0, 1]$; l'adaptation au cas général se fait facilement.

Supposons donné un générateur x convergent à information finie pour le problème (B) sur $\mathcal{M}_c(0, 1])$.

Soit $f_0 \in \mathcal{M}_c([0, 1])$ définie par :

$$\forall t \in [0, 1] : f_0(t) = 0 .$$

Soit $x_0(f_0)$ la première réponse donnée par x pour f_0.

Puisque l'ensemble $X_0(f_0)$ des points utilisés par x pour calculer $x_0(f_0)$ est fini, on peut trouver un intervalle $[a_0, b_0]$ de longueur non nulle ne rencontrant pas $X_0(f_0)$ et tel que :

$$[a_0, b_0] \subset [0, 1/4] \quad \text{si} \quad x_0(f_0) \in [1/2, 1] ,$$

$$[a_0, b_0] \subset [3/4, 1] \quad \text{si} \quad x_0(f_0) \in [0, 1/2] .$$

Nous supposerons par la suite que c'est le premier cas qui se produit. On définit alors f_1 en posant :

$$a'_0 = (2\,a_0 + b_0)/3\,, \quad b'_0 = (a_0 + 2\,b_0)/3$$
$$\forall t \in [a'_0, b'_0] : \quad f_1(t) = 1/2$$
$$\forall t \in [0, 1] - [a_0, b_0] : \quad f_1(t) = f_0(t)$$
$$f_1 \text{ linéaire sur } [a_0, a'_0], [b'_0, b_0]\,.$$

Figure 1.

Puisque $[a_0, b_0]$ ne rencontre pas $X_0(f_0)$ on a :

$$x_0(f_0) = x_0(f_1)\,.$$

On pose $n_0 = 0$.

Puisque x est convergent et que opt $(f_1) \subset [0, 1/4]$ il existe $n_1 > n_0$ tel que :

$$x_{n_1}(f_1) \in [0, 3/8]\,.$$

Puisque l'ensemble $X_{n_1}(f_1)$ des points utilisés par x pour calculer $x_{n_1}(f_1)$ est fini, on peut trouver un intervalle de longueur non nulle $[a_1, b_1]$ ne rencontrant pas $X_{n_1}(f_1)$, contenu dans $[3/4, 1]$.

On définit alors f_2 en posant :

$$a'_1 = (2\,a_1 + b_1)/3\,, \quad b'_1 = (a_1 + 2\,b_1)/3\,,$$
$$\forall t \in [a'_1, b'_1] : \quad f_2(t) = 1 - 1/4\,,$$
$$\forall t \in [0, 1] - [a_1, b_1] : \quad f_2(t) = f_1(t)\,,$$
$$f_2 \text{ linéaire sur } [a_1, a'_1], [b'_1, b_1]\,.$$

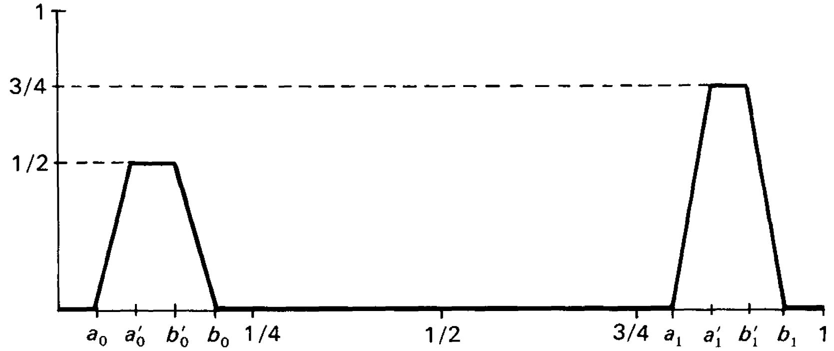

Figure 2.

Puisque $[a_1, b_1]$ ne rencontre pas $X_1(f_1)$, on a :

$$x_1(f_0) = x_1(f_1) = x_1(f_2).$$

Puisque x est convergent et que $\operatorname{opt}(f_2) \subset [3/4, 1]$ il existe $n_2 > n_1$ tel que

$$x_{n_2}(f_2) \in [5/8, 1].$$

Puisque l'ensemble $X_{n_2}(f_1)$ des points utilisés par x pour calculer $x_{n_2}(f_2)$ est fini, on peut trouver un intervalle $[a_2, b_2]$ de longueur non nulle, ne rencontrant pas $X_{n_2}(f_2)$, contenu dans $[a'_0, b'_0]$.

On définit alors f_3 en posant :

$$a'_2 = (2\,a_2 + b_2)/3\,, \quad b'_2 = (a_2 + 2\,b_2)/3\,,$$
$$\forall t \in [a'_2, b'_2] : \ f_3(t) = 1 - 1/2^3\,,$$
$$\forall t \in [0, 1] - [a_2, b_2] : \ f_3(t) = f_2(t)\,,$$
$$f_3 \text{ linéaire sur } [a_2, a'_2], [b'_2, b_2]\,.$$

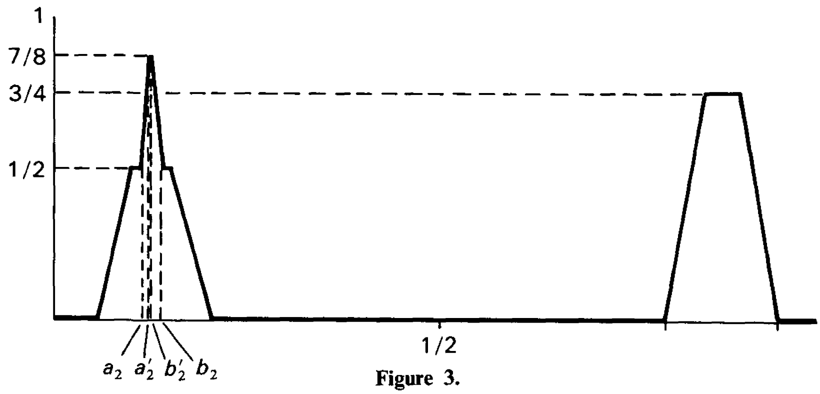

Figure 3.

On continue de la même façon, définissant ainsi des fonctions continues f_i, des intervalles $[a_i, b_i]$, $[a'_i, b'_i]$, des indices n_i tels que, pour tout $i \in \mathbb{N}$:

$$\left. a'_i = (2\,a_i + b_i)/3\,, \quad b'_i = (a_i + 2\,b_i)/3\,; \right\} \qquad (A_i)$$

$$\left. \begin{array}{l} [a_{2i}, b_{2i}] \supset [a_{2i+2}, b_{2i+2}]\,, \\ [a_{2i+1}, b_{2i+1}] \supset [a_{2i+3}, b_{2i+3}]\,. \end{array} \right\} \qquad (B_i)$$

$$\left. \begin{array}{l} \forall t \in [a'_i, b'_i] : \ f_{i+1}(t) = 1 - 1/2^{i+1}\,, \\ \forall t \in [0, 1] - [a_i, b_i] : \ f_{i+1}(t) = f_i(t)\,, \end{array} \right\} \qquad (C_i)$$

$$f_{i+1} \text{ linéaire sur } [a_i, a'_i], [b'_i, b_i]$$

$$\left. \begin{array}{l} n_{i+1} > n_i\,, \\ [a_i, b_i] \cap X_{n_i}(f_i) = \varnothing\,, \end{array} \right\} \qquad (D_i)$$

$$\left. \begin{array}{l} x_{n_i}(f_i) \subset [0, 3/8] \quad \text{si} \quad i \text{ est impair}\,, \\ x_{n_i}(f_i) \subset [5/8, 1] \quad \text{si} \quad i \text{ est pair}\,. \end{array} \right\} \qquad (E_i)$$

Les formules (A_i), (B_i), (C_i) montrent que la suite de fonctions (f_i) est uniformément convergente vers une fonction f continue et possédant 1 comme maximum, ce maximum étant atteint au moins une fois dans l'intervalle $[0, 1/4]$ et au moins une fois dans l'intervalle $[3/4, 1]$. Les formules (D_i) montrent que, pour tout $i \in \mathbb{N}$,

$$X_{n_i}(f_i) = X_{n_i}(f)$$

et que, en conséquence, pour tout $i \in \mathbb{N}$:

$$x_{n_i}(f) = x_{n_i}(f_i)\,.$$

La suite $(x_{n_i}(f))$ d'après les formules (E_i) n'est pas convergente ce qui, puisque $f \in \mathscr{M}_c[0, 1]$, est contraire aux hypothèses faites sur x.

Remarques

1) Ce qui nous a permis de « piéger » l'éventuel générateur x c'est la possibilité de lui proposer des fonctions ayant plusieurs solutions optimales. Nous verrons au théorème 4 que si on impose aux fonctions envisagées de n'avoir qu'une seule solution optimale (c'est-à-dire si on réduit la famille $\mathscr{M}_c(E)$) alors le résultat du théorème 2 n'est plus vrai.

2) Si on avait exigé du générateur x d'être seulement semi-convergent (c'est-à-dire que toutes les valeurs d'adhérence des $x_n(f)$ soient dans opt (f), pout tout $f \in \mathscr{M}_c(E)$) la démonstration proposée ne marcherait plus. Nous reviendrons sur cette question au paragraphe suivant.

3) Bien sûr, le résultat négatif (partie (ii) du théorème 2) vaut pour toute famille contenant $\mathcal{M}_c(E)$, et même pour toute famille contenant toutes les fonctions qui interviennent dans la démonstration.

4) La partie (ii) du théorème s'étend à tout espace métrique possédant une partie homéomorphe à un compact d'intérieur non vide de $\mathbb{R}^n$. La partie (i), elle, vaut pour tout espace métrique séparable.

5) Nous désignerons par $\mathcal{M}_{c.p.}(E)$ ($p \in \mathbb{N}$) l'ensemble des applications de E dans $\mathbb{R}$ ayant un maximum et p fois différentiables sur E. Voici un problème non résolu :

$\left\{\right.$ Existe-t-il des générateurs convergents à information finie d'ordre $k (k \leqslant p)$ pour le problème (B) sur $\mathcal{M}_{c.p.}(E)$?

En particulier, on ne sait pas répondre à la question :

$\left\{\right.$ Existe-t-il des générateurs convergents à information finie d'ordre 1 pour le problème (B) sur $\mathcal{M}_{c.1.}([0, 1])$?

ce qui, de façon plus explicite, signifie encore :

$\left\{\right.$ Peut-on trouver un algorithme qui, pour toute fonction f continue et dérivable sur $[0, 1]$ en utilisant à chaque étape de ses calculs des évaluations de f et f' en un nombre fini de points, construirait une suite convergente $x_n(f)$ ayant pour limite un point où f atteint son maximum ?

La démonstration de la partie (ii) du théorème 2 ne s'adapte pas (du moins de façon évidente) car les fonctions construites pour piéger l'éventuel générateur ne sont pas dérivables et même si on modifie un peu les f_n pour qu'elles soient dérivables (ce qui est facile) malheureusement la fonction f obtenue à la limite n'est plus dérivable.

c) Densité de l'ensemble des points d'évaluation

THÉORÈME 3 : *Soit E une partie compacte de $\mathbb{R}^m$ d'intérieur non vide. Soit M un générateur convergent à information finie pour le problème (A) sur $\mathcal{M}_{s.c.i.}(E)$ (resp. $\mathcal{M}_{c.p.}(E)$). Alors, pour tout $f \in \mathcal{M}_{s.c.i.}(E)$ (resp. $f \in \mathcal{M}_{c.p.}(E)$) $X(f)$ est dense dans E.*

Interprétation : L'ensemble des points où il est nécessaire d'évaluer une fonction s.c.i. ou c.p. pour connaître son maximum est dense dans l'ensemble de définition.

Démonstration : Soit M un générateur convergent à information finie pour le problème (A) sur $\mathcal{M}_{s.c.i.}(E)$ (resp. $\mathcal{M}_{c.p.}(E)$).

Supposons qu'il existe $f \in \mathcal{M}_{\text{s.c.i.}}(E)$ (resp. $\mathcal{M}_{\text{c.p.}}(E)$) telle que $X(f)$ ne soit pas dense dans E.

Il existe donc $a \in E$, $\varepsilon \in \mathbb{R}^{+*}$ tels que la boule ouverte de centre a et de rayon ε, $B(a, \varepsilon)$, ne rencontre pas $X(f)$.

Il est facile de construire une fonction $g \in \mathcal{M}_{\text{s.c.i.}}(E)$ (resp. $\mathcal{M}_{\text{c.p.}}(E)$) telle que :

$$\forall x \in E - B(a, \varepsilon) \quad f(x) = g(x)$$

$$M(f) \neq M(g).$$

Par définition $X(f) = X(g)$ et $(x_n(f)) = (x_n(g))$.

Le générateur M ne sera donc pas convergent pour g.

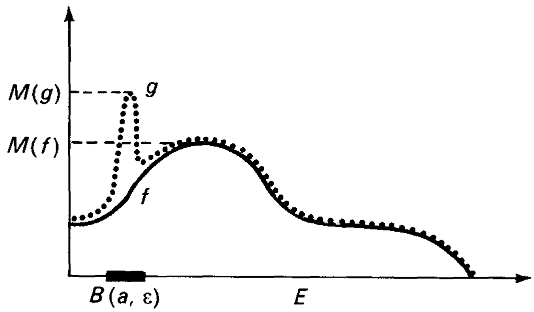

Figure 4.

Remarque : L'extension du théorème 3 aux générateurs d'ordre $k \in \mathbb{N}^*$ est évidente.

d) Solutions optimales et solution optimale unique

Soit E un espace métrique, nous désignerons par $\mathcal{MU}_c(E)$ l'ensemble des applications continues f de E dans $\mathbb{R}$ ayant un maximum atteint en un seul point, c'est-à-dire telle que : $\operatorname{card}(\operatorname{opt}(f)) = 1$.

THÉORÈME 4 : *Soit E une partie compacte de $\mathbb{R}^m$ d'intérieur non vide :*

(i) *il existe des générateurs convergents à information finie pour le problème (B) sur $\mathcal{MU}_c(E)$,*

(ii) *il n'existe pas de générateurs convergents à information finie pour le problème (B) sur $\mathcal{M}_c(E)$.*

Interprétation : Le fait d'avoir affaire à des fonctions dont l'optimum est unique peut être décisif pour la convergence des méthodes d'optimisation : sans la condition d'unicité, il se peut qu'aucun algorithme ne puisse converger alors qu'avec la condition d'unicité les algorithmes peuvent converger.

Démonstration : La partie (ii) du théorème 4 qui est la même que la partie (ii) du théorème 2 a déjà été démontrée. Établissons la partie (i).

Soit $\{ a_0, a_1, \dots, a_n, \dots \}$ une partie dénombrable dense de E.

On pose :

$$x_n(f) = a_{i_n}$$

avec

$$i_n = \min \{ i \in \{ 0, 1, \dots, n \} \mid f(a_i) = \max \{ f(a_j) \mid 0 \leqslant j \leqslant n \} \}.$$

Comme dans le théorème 1 (i) on montre que :

$$\lim_{n \to \infty} f(x_n(f)) = M(f)$$

ce qui, puisque E est compact, entraîne que :

$$\lim_{n \to \infty} x_n(f) \in \text{opt}\,(f)\,.$$

Remarques

1) La partie (i) s'étend à tout espace métrique séparable et la partie (ii) à tout espace métrique non dénombrable.

2) Lorsque E n'est pas compact, nous verrons, au théorème 5, que l'unicité de la solution optimale n'est pas suffisante.

e) Compacité et non-compacité

THÉORÈME 5 : *Soit E une partie compacte de $\mathbb{R}^m$ d'intérieur non vide* :

(i) *il existe des générateurs convergents à information finie pour le problème* (B) *sur* $\mathcal{M}\mathcal{U}_c(E)$,

(ii) *il n'existe pas de générateurs convergents à information finie pour le problème* (B) *sur* $\mathcal{M}\mathcal{U}_c(\mathbb{R}^m)$.

Interprétation : La compacité de l'ensemble de définition des fonctions que l'on cherche à optimiser peut être décisive ; avec elle on peut trouver des méthodes d'optimisation, sans elle, il est possible qu'aucun algorithme (convergent) n'existe.

Démonstration : La partie (i) a déjà été établie au théorème 4.

Montrons la partie (ii) dans le cas de $\mathbb{R}$ pour plus de simplicité. Supposons donné un générateur x convergent à information finie pour le problème (B) sur $\mathcal{MU}_c(\mathbb{R})$.

Soit $f_0 \in \mathcal{MU}_c(\mathbb{R})$ définie par :

$$\forall t \in [-2, -1] : \; f_0(t) = 1/2$$
$$f_0(1) = 1 - 1/2^2 , \; f_0(0) = 0$$
$$\forall t \in \mathbb{R} - [-3, 2] : \; f_0(t) = 0$$
$$f_0 \text{ linéaire sur } [-3, -2], [-1, 0], [0, 1], [1, 2] .$$

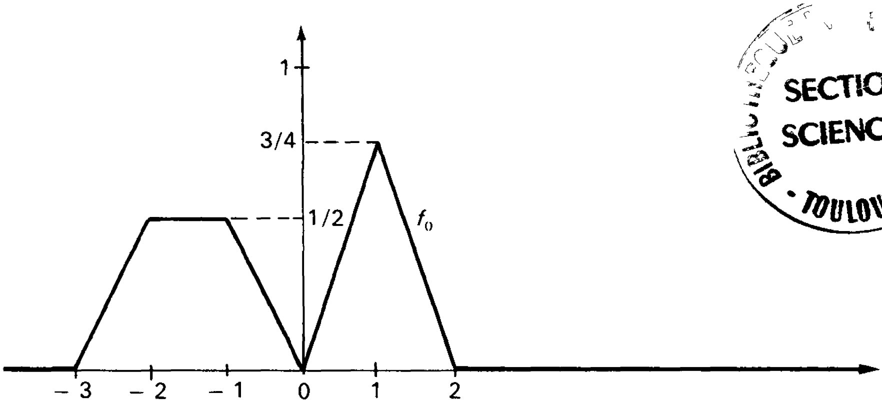

Figure 5.

Soit $(x_n(f_0))$ la suite des réponses données par x pour f_0.

Puisque x est convergent et que opt $(f_0) = \{ 1 \}$ il existe n_0 tel que

$$x_{n_0}(f_0) \in [0, 2] .$$

Puisque l'ensemble $X_{n_0}(f_0)$ des points utilisés par x pour calculer $x_{n_0}(f_0)$ est fini, on peut trouver un intervalle de longueur non nulle $[a_1, b_1]$ ne rencontrant pas $X_{n_0}(f_0)$ et contenu dans $[-2, -1] = [a_0, b_0]$. De même on peut trouver un intervalle de longueur non nulle $[\alpha_1, \beta_1]$ ne rencontrant pas $X_{n_0}(f_0)$ et contenu dans $[2, 4]$.

On définit alors f_1 en posant :

$$a'_1 = (2\,a_1 + b_1)/3\,, \quad b'_1 = (a_1 + 2\,b_1)/3\,,$$
$$\forall t \in [a'_1, b'_1] : \; f_1(t) = 1 - 1/2^2\,,$$
$$f_1((\alpha_1 + \beta_1)/2) = 1 - 1/2^3\,,$$
$$\forall t \in]-\infty, a_1] \cup [b_1, \alpha_1] \cup [\beta_1, +\infty[\, : f_1(t) = f_0(t)$$
$$f_1 \text{ linéaire sur } [a_1, a'_1], [b'_1, b_1], [\alpha_1, (\alpha_1 + \beta_1)/2]\,,$$
$$[(\alpha_1 + \beta_1)/2, \beta_1]\,.$$

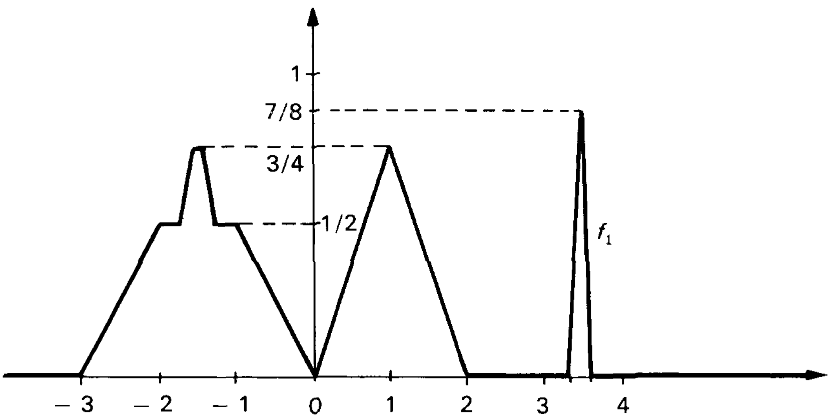

Figure 6.

Soit $(x_n(f_1))$ la suite des réponses données par x pour f.
Par construction de f_1 cette suite coıncide avec $(x_n(f_0))$ pour $n \leqslant n_0$.

Puisque $f_1 \in \mathcal{MU}_c(\mathbb{R})$, que x est convergent sur $\mathcal{MU}_c(\mathbb{R})$ et que

$$\mathrm{opt}(f_1) = (\alpha_1 + \beta_1)/2\,,$$

il existe $n_1 > n_0$ tel que :

$$x_{n_1}(f_1) \in [2, 4]\,.$$

Puisque l'ensemble $X_{n_1}(f_1)$ des points utilisés par x pour calculer $x_{n_1}(f_1)$ est fini, on peut trouver un intervalle de longueur non nulle $[a_2, b_2] \subset [a'_1, b'_1]$ et ne rencontrant pas $X_{n_1}(f_1)$. De même on peut trouver un intervalle de longueur non nulle $[\alpha_2, \beta_2]$ ne rencontrant pas $X_{n_1}(f_1)$ et contenu dans [4, 6].

On définit alors f_2 en posant :

$$a'_2 = (2\,a_2 + b_2)/3\,, \quad b'_2 = (a_2 + 2\,b_2)/3\,,$$
$$\forall x \in [a'_2, b'_2] : \; f_2(x) = 1 - 1/2^3\,,$$
$$f_2((\alpha_2 + \beta_2)/2) = 1 - 1/2^4\,,$$
$$\forall t \in]-\infty, a_2] \cup [b_2, \alpha_2] \cup [\beta_2 + \infty[: \; f_2(t) = f_1(t)\,,$$
$$f_2 \text{ linéaire sur } [a_2, a'_2], [b'_2, b_2], [\alpha_2, (\alpha_2 + \beta_2)/2]\,,$$
$$[(\alpha_2 + \beta_2)/2, \beta_2]\,.$$

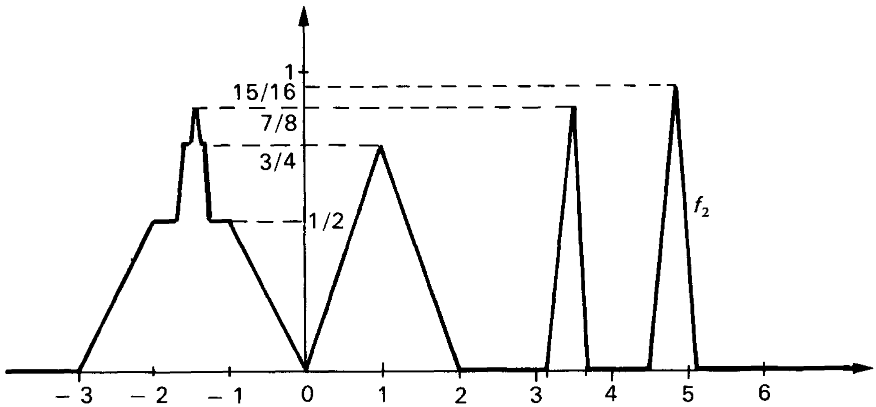

Figure 7.

On continue de la même façon en définissant ainsi des fonctions $f_i \in \mathcal{M}\mathcal{U}_c(\mathbb{R})$, des intervalles $[a_i, b_i]$, $[\alpha_i, \beta_i]$ des indices n_i tels que pour tout $i \in \mathbb{N}^*$:

$$a'_1 = (2\,a_i + b_i)/3\,, \quad b'_i = (a_i + 2\,b_i)/3\,,$$
$$[a_{i+1}\, b_{i+1}] \subset [a'_i, b'_i]\,, \qquad (A_i)$$
$$[\alpha_i, \beta_i] \subset [2\,i, 2\,i + 2]\,.$$
$$\forall t \in [a'_i, b'_i] : \; f_i(t) = 1 - 1/2^{i+1}\,,$$
$$f_i((\alpha_i + \beta_i)/2) = 1 - 1/2^{i+2}\,. \qquad (B_i)$$
$$\forall t \in]-\infty, a_i] \cup [b_i, \alpha_i] \cup [\beta_i, +\infty[: \; f_i(t) = f_{i-1}(t)$$
$$f_i \text{ linéaire sur } [a_i, a'_i], [b'_i, b_i], [\alpha_i, (\alpha_i + \beta_i)/2], [(\alpha_i + \beta_i)/2, \beta_i]\,, \qquad (C_i)$$
$$n_{i+1} > n_i\,,$$
$$[a_{i+1}, b_{i+1}] \cap X_{n_i}(f_i) = \varnothing\,, \qquad (D_i)$$
$$[\alpha_{i+1}, \beta_{i+1}] \cap X_{n_i}(f_i) = \varnothing\,,$$
$$x_{n_i}(f_i) \subset [2\,i, 2\,i + 2]\,. \qquad (E_i)$$

Les formules (A_i), (B_i), (C_i) montrent que la suite de fonctions (f_i) est uniformément convergente sur tout compact de $\mathbb{R}$ et donc admet pour limite une fonction f qui est continue. D'après (A_i), (B_i) cette fonction f possède un maximum qui est 1 et est atteint en un point unique γ $\{\gamma\} = \bigcap_{i \in \mathbb{N}} [a_i, b_i]$.

D'après (D_i) et (E_i) la suite des réponses $(x_n(f))$ donnée par x pour f va vérifier que :

$$\forall i \in \mathbb{N} : x_{n_i}(f) \subset [2\,i, 2\,i + 2]\,.$$

Contrairement aux hypothèses, x n'est donc pas convergent pour f.

Remarque : La partie (ii) du théorème s'étend à tout sous-ensemble E de $\mathbb{R}^n$ d'intérieur non vide et non compact.

III. QUELQUES RÉSULTATS COMPLÉMENTAIRES

Selon la famille de fonctions à laquelle on a affaire, on peut demander plus que la convergence ou moins que la convergence. Nous introduisons ici les notions de convergence uniforme d'un générateur et de semi-convergence d'un générateur qui expriment l'une plus, l'autre moins que la convergence. Nous donnons quelques résultats sans démonstration.

Dire que le générateur M (resp. x) pour le problème (A) (resp. (B)) est convergent sur $\mathcal{S} \subset \mathcal{M}(E)$ signifie par définition que :

$$\forall f \in \mathcal{S}, \quad \forall \varepsilon \in \mathbb{R}^{+*}, \quad \exists n \in \mathbb{N}, \quad \forall m \geqslant n : |\, M_m(f) - M(f)\,| \leqslant \varepsilon$$

$$(\text{resp. } \forall f \in \mathcal{S}, \quad \exists \hat{x} \in \operatorname{opt}(f), \quad \forall \varepsilon \in \mathbb{R}^{+*}, \quad \exists n \in \mathbb{N}, \quad \forall m \geqslant n : |\, x_m(f) - \hat{x}\,| \leqslant \varepsilon)\,.$$

On dira que le générateur M (resp. x) pour le problème (A) (resp. (B)) est *uniformément convergent* sur $\mathcal{S}$ ssi :

$$\forall \varepsilon \in \mathbb{R}^{+*}, \quad \exists n \in \mathbb{N}, \quad \forall f \in \mathcal{S}, \quad \forall m \geqslant n : |\, M_m(f) - M(f)\,| \leqslant \varepsilon$$

$$(\text{resp. } \forall \varepsilon \in \mathbb{R}^{+*}, \quad \exists n \in \mathbb{N}, \quad \forall f \in \mathcal{S}, \quad \exists \hat{x} \in \operatorname{opt}(f), \quad \forall m \geqslant n : |\, x_m(f) - \hat{x}\,| \leqslant \varepsilon)\,.$$

Lorsque (x_n) est une suite non convergente, on notera $\mathcal{A}(x_n)$ l'ensemble de ses valeurs d'adhérence (ou points d'accumulation).

On dira que le générateur x pour le problème (B) est *semi-convergent* sur $\mathcal{S}$ ssi :

$$\forall f \in \mathcal{S} : \mathcal{A}(x_n(f)) \neq \varnothing \quad \text{et} \quad \mathcal{A}(x_n(f)) \subset \operatorname{opt}(f) .$$

Par définition, pour tout générateur, on a :

uniformément convergent $\Rightarrow$ convergent $\Rightarrow$ semi-convergent.

a) Générateur pour $\mathcal{M}(E)$

Soit E un espace métrique. Rappelons que $\mathcal{M}(E)$ désigne l'ensemble de toutes les applications de E dans $\mathbb{R}$ qui atteignent leur maximum.

On a les résultats suivants :

(i) E fini
Il existe des générateurs à information finie *uniformément convergents* pour (A) (resp. (B)) sur $\mathcal{M}(E)$;

(ii) E infini dénombrable
Il existe des générateurs à information finie *convergents* pour (A) (resp. (B)) sur $\mathcal{M}(E)$;
Il n'existe pas de générateurs à information finie *uniformément convergents* pour (A) (resp. (B)) sur $\mathcal{M}(E)$.

(iii) E infini non dénombrable
Il n'existe pas de générateurs à information finie *convergents* pour (A) sur $\mathcal{M}(E)$;
Il n'existe pas de générateurs à information finie *semi-convergents* pour (B) sur $\mathcal{M}(E)$.

b) Générateurs pour $\mathcal{M}_c$, $\mathcal{M}\mathcal{U}_c$, $\mathcal{M}\mathcal{D}_c$, $\mathcal{M}\mathcal{D}\mathcal{U}_c$

Appelons fonction à maximum dominants toute fonction $f \in \mathcal{M}(E)$ telle que :

$$\forall (x_n) \in E_{\mathbb{N}} : \lim_{n \to \infty} f(x_n) = \max_{t \in E} f(t) \Rightarrow \mathcal{A}(x_n) \neq \varnothing .$$

Nous noterons :

$\mathcal{M}\mathcal{D}(E)$ l'ensemble des fonctions $f \in \mathcal{M}(E)$ à maximum dominants

$$\begin{aligned} \mathcal{M}\mathcal{D}\mathcal{U}(E) &= \mathcal{M}\mathcal{D}(E) \cap \mathcal{M}\mathcal{U}(E) \\ \mathcal{M}\mathcal{D}_c(E) &= \mathcal{M}\mathcal{D}(E) \cap \mathcal{M}_c(E) \\ \mathcal{M}\mathcal{D}\mathcal{U}_c(E) &= \mathcal{M}\mathcal{D}\mathcal{U}(E) \cap \mathcal{M}_c(E) . \end{aligned}$$

On établit facilement que :

- Si E est compact $\mathcal{M}_c(E) = \mathcal{M}\mathcal{D}_c(E)$
  et $\mathcal{M}\mathcal{U}_c(E) = \mathcal{M}\mathcal{D}\mathcal{U}_c(E)$
- $f \in \mathcal{M}\mathcal{D}(E) \Rightarrow \mathrm{opt}(f)$ compact
- $f \in \mathcal{M}\mathcal{D}(E) \Rightarrow (\forall (x_n) \in E : \lim\limits_{n \to \infty} f(x) = \max\limits_{t \in E} f(t) \Rightarrow (x_n)$ bornée).

(i) *Convergence uniforme*

Bien que la famille $\mathcal{M}\mathcal{D}\mathcal{U}_c(E)$ soit la plus petite des familles envisagées dans cette section, la condition de convergence uniforme est encore trop exigeante en général :

si E est infini, il n'existe pas de générateurs à information finie uniformément convergents pour (A) (resp. pour (B)) sur $\mathcal{M}\mathcal{D}\mathcal{U}_c(E)$.

(ii) *Convergence*

Puisque lorsque E est compact, $\mathcal{M}\mathcal{U}_c(E) = \mathcal{M}\mathcal{D}\mathcal{U}_c(E)$, le résultat suivant améliore le théorème 5 (i) :

Si E est séparable, il existe des générateurs convergents à information finie pour le problème (B) sur $\mathcal{M}\mathcal{D}\mathcal{U}_c(E)$.

Puisque lorsque E est compact $\mathcal{M}_c(E) = \mathcal{M}\mathcal{D}_c(E)$ le résultat suivant améliore le théorème 1 (i) :

Si E est sans point isolé, il n'existe pas de générateurs convergents à information finie pour le problème (B) sur $\mathcal{M}\mathcal{D}_c(E)$.

(iii) *Semi-convergence*

Si E est séparable, il existe des générateurs semi-convergents à information finie pour (B) sur $\mathcal{M}\mathcal{D}_c(E)$.

Ce qui donne, en particulier :

Si E est compact, il existe des générateurs semi-convergents à information finie pour (B) sur $\mathcal{M}_c(E)$.

Par contre, le problème simple suivant reste non résolu jusqu'à présent :
Existe-t-il des générateurs semi-convergents à information finie pour (B) sur $\mathcal{M}_c(\mathbb{R}^m)$, $\mathcal{M}\mathcal{U}_c(\mathbb{R}^m)$?

BIBLIOGRAPHIE

[1] O. ABERTH, *Analysis in computable number field.* J.A.C.M., vol. 15, n° 2, April 1968.

[2] M. BOUHIER, *Quelques questions d'analyse numérique traitées du point de vue de la calculabilité.* Thèse de 3e Cycle, Université Scientifique et Médicale de Grenoble, 1974.

[3] J.-P. DELAHAYE, *Théorie des transformations de suites en analyse numérique. Applications.* Thèse d'État, Université de Lille, 1982.

[4] J.-P. DELAHAYE, *Optimalité du procédé Δ^2 d'Aitken pour l'accélération de la convergence linéaire.* R.A.I.R.O., Analyse Numérique, 15, 1981, pp. 321-330.

[5] J.-P. DELAHAYE, *Sur quelques limitations des algorithmes dans le traitement des suites.* Pub. A.N.O., n° 116. Université de Lille, juin 1983.

[6] J.-P. DELAHAYE et B. GERMAIN-BONNE, *Résultats négatifs en accélération de la convergence.* Num. Math. 35, 1980, pp. 443-457.

[7] J.-P. DELAHAYE et B. GERMAIN-BONNE, *The set of logarithmically convergent sequences cannot be accelerated.* S.I.A.M., J. Num. Anal., 1982, pp. 840-844.

[8] J. DENEL, *Contribution à la synthèse des algorithmes d'optimisation.* Thèse d'État, Université de Lille, 1979.

[9] N. GASTINEL, *Introduction à l'analyse calculable.* Polycopié d'un cours de D.E.A., 1972. Université Scientifique et Médicale de Grenoble.

[10] P. HUARD, *Point-to-set maps and mathematical programming.* Mathematical Programming Study, n° 10, 1979 (Avec A. Auslender, J. M. Borwein, J.-P. Delahaye, J. Denel, J. Ch. Fiorot, E. G. Gol'Shtein, P. Huard, D. Klatte, R. Klessig, B. Kummer, G. G. L. Meyer, E. Polak, S. M. Robinson, A. Ruszczynski, R. Saigal, J. Szymanowski, S. Tishydhigama, N. V. Tret'Yakow).

[11] E. POLACK, *Computation methods in optimization.* Academic Press, New York, 1971.

[12] J. F. TRAUB and H. WOZNIAKOWSKI, *A general theory of optimal algorithms.* Academic Press, New York, 1980.

[13] W. I. ZANGWILL, *Nonlinear programming a unified approach.* Prentice Hall, inc. Englewood Cliffs, N.J., 1969.