

F. MESLIER

P. SANDRIN

## **Résolution de problèmes d'optimisation dynamique par programmation dynamique marginaliste**

*RAIRO. Recherche opérationnelle*, tome 15, n° 4 (1981),  
p. 363-381

[http://www.numdam.org/item?id=RO\\_1981\\_\\_15\\_4\\_363\\_0](http://www.numdam.org/item?id=RO_1981__15_4_363_0)

© AFCET, 1981, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## RÉSOLUTION DE PROBLÈMES D'OPTIMISATION DYNAMIQUE PAR PROGRAMMATION DYNAMIQUE MARGINALISTE (\*)

par F. MESLIER et P. SANDRIN <sup>(1)</sup>

---

*Résumé. — La programmation dynamique marginaliste, dont les principes ont été définis par Pierre Masse dès 1946, est une technique numérique qui semble avoir assez peu retenu l'attention, par rapport à d'autres comme la programmation dynamique, la commande optimale ou encore les méthodes générales d'optimisation. Nous utilisons cette approche, avec satisfaction, depuis de nombreuses années, pour traiter divers problèmes d'optimisation dynamique — le plus souvent stochastique — qui se posent en particulier en matière de gestion des réserves hydrauliques.*

*Nous exposons ici les principes de cette méthode et nous en montrons l'optimalité.*

Mots clés : Programmation dynamique, gestion des stocks, décision en avenir aléatoire.

*Abstract. — Differential dynamic programming, firstly investigated by P. Masse as soon as in 1946, is a numerical technique which has not been much used, compared with dynamic programming, optimal control or general optimization methods. However it has helped us to solve several dynamic optimization problems, especially in random future, met in the management of hydraulic reserves.*

*In this paper, we state the principles of that method and prove its optimality.*

Keywords: Dynamic programming, inventory control, stochastic decision.

### INTRODUCTION

Les activités d'une entreprise productrice d'électricité posent de nombreux problèmes d'optimisation où le caractère dynamique est bien souvent essentiel (la gestion des réserves hydrauliques est sans doute l'exemple le plus connu). Nous utilisons depuis plusieurs années, pour divers modèles conçus pour des études de gestion prévisionnelle et de planification, la programmation dynamique marginaliste, dont les principes ont été exposés par Pierre Masse dès 1946 [1]. Cette technique moins usitée que la classique programmation dynamique, ou la commande optimale, ou encore les méthodes générales de programmation mathématique, nous donne toute satisfaction pour certains types d'études [2], [3].

En effet, bien que le domaine d'utilisation de cette méthode soit, ainsi que nous le verrons par la suite, relativement restreint, son emploi à bon escient présente

---

(\*) Reçu septembre 1979.

(1) E.D.F., Direction des Études et Recherches, 1, avenue du Général-de-Gaulle, 92141 Clamart, France.

d'excellentes propriétés dans diverses directions : le traitement de la prise de décision en avenir aléatoire, la robustesse bien connue de ce type de technique, de bonnes performances sur le plan du temps de calcul et une grande clarté de l'interprétation des résultats.

Cet article présente les fondements théoriques de la programmation dynamique marginaliste, dont l'optimalité est démontrée par application des conditions de Kuhn et Tucker aux équations de la programmation dynamique classique.

Nous aborderons successivement les points suivants :

1. position du problème;
2. les conditions de Kuhn et Tucker;
3. calcul par récurrence des valeurs marginales de l'état;
4. conclusion.

## 1. POSITION DU PROBLÈME

**1.1.** Considérons le problème d'optimisation dynamique de gestion d'un stock en avenir aléatoire :

Soient :

$x_t$ ,  $t = 1, T$ , variable d'état;

$u_t$ ,  $t = 1, T$ , variable de commande efficace;

$v_t$ ,  $t = 1, T$ , variable de commande inefficace <sup>(2)</sup>;

$x_1$ , état initial donné;

$e_t$ ,  $t = 1, T$ , entrées, aléatoires, dont on connaît une loi de probabilité et telles que  $e_t \geq 0$ ;

$G_t(u_t, e_t)$ ,  $t = 1, T$ , fonctions de gain instantané, « aléatoires » en ce sens qu'elles sont fonctions des entrées  $e_t$ , et que nous supposerons concaves tout au long de cet article;

$Z_{T+1}(x_{T+1})$ , fonction de  $x_{T+1}$ , décrivant la valeur que l'on attribue à l'état final, que nous supposerons également concave dans cet article.

---

<sup>(2)</sup> *Remarque* : La commande inefficace  $v_t$ , qui représente par exemple un déversement à un réservoir hydraulique, permet la prise en compte des cas où, dans l'évolution de l'état :

$$x_{t+1} = x_t + e_t - u_t - v_t,$$

une grande valeur de l'entrée  $e_t$  conduit à la butée simultanée sur les contraintes :

$$x_{t+1} = x_{\max} \quad \text{et} \quad u_t = u_{\max}.$$

Nous définissons le critère à maximiser comme l'espérance, sur les aléas, de la somme des gains dus aux différentes commandes, et de la valeur de l'état final. Le problème d'optimisation est alors résoudre :

$$\text{maximiser } E_{\text{aléas}} \sum_{t=1}^T G_t(u_t, e_t) + Z_{T+1}(x_{T+1}),$$

sous les contraintes :

$$x_{t+1} = x_t + e_t - u_t - v_t, \quad t = 1, T,$$

$$x_{\min} \leq x_t \leq x_{\max}, \quad t = 1, T+1,$$

$$u_{\min} \leq u_t \leq u_{\max}, \quad t = 1, T,$$

$$0 \leq v_t, \quad t = 1, T.$$

1.2. Le problème déterministe qui correspond au problème précédent, s'en déduit en considérant les entrées  $e_t$  comme fixes et non aléatoires. Ce problème s'écrit :

$$\text{maximiser } \sum_{t=1}^T G_t(u_t) + Z_{T+1}(x_{T+1}),$$

sous les contraintes :

$$x_{t+1} = x_t + e_t - u_t - v_t, \quad t = 1, T,$$

$$x_{\min} \leq x_t \leq x_{\max}, \quad t = 1, T+1,$$

$$u_{\min} \leq u_t \leq u_{\max}, \quad t = 1, T,$$

$$0 \leq v_t.$$

(a) Sous les seules hypothèses précédentes, ce problème peut être résolu à l'aide de la technique classique de programmation dynamique. Reprenons la formulation due à Bellmann [4], qui définit une fonction valeur de l'état  $Z_t(x_t)$  comme la somme des gains que l'on peut réaliser en parcourant la trajectoire optimale de  $x_t$  à  $x_{T+1}$ , pour les semaines  $t, t+1, \dots, T$  :

$$Z_t(x_t) = \text{Max}_{\substack{u_t, u_{t+1}, \dots, u_T \\ v_t, v_{t+1}, \dots, v_T}} \left[ \sum_{\theta=t}^T G_\theta(u_\theta) \right] + Z_{T+1}(x_{T+1}),$$

Une relation de récurrence s'en déduit :

$$Z_t(x_t) = \text{Max}_{u_t, v_t} [G_t(u_t) + Z_{t+1}(x_{t+1})], \quad (1)$$

avec :

$$x_{t+1} = x_t + e_t - u_t - v_t,$$

$$x_t \in [x_{\min}, x_{\max}],$$

$$u_t \in [u_{\min}, u_{\max}],$$

$$v_t \geq 0.$$

L'algorithme de programmation dynamique procède en deux étapes :

— détermination des fonctions valeurs de l'état, par récurrence, en « remontant » la variable de temps  $t$  de  $t = T$  à  $t = 1$  (en avenir certain on peut en fait remonter ou descendre le temps; par contre, en avenir aléatoire, on ne peut que le remonter);

— détermination des commandes optimales, en « redescendant » la variable de temps  $t$ , à partir d'un niveau connu de l'état  $x_1$ .

(b) L'introduction d'hypothèses plus fortes permet d'avoir recours à des méthodes générales d'optimisation qui compte tenu de la structure particulièrement simple des contraintes, sont extrêmement performantes.

Si les fonctions  $G_t(u_t)$  et  $Z_{T+1}(x_{T+1})$  sont linéaires en leurs arguments, le programme linéaire correspondant possède alors une structure de flot pour laquelle des algorithmes sur mesure [5] donnent d'excellentes performances.

Si les fonctions  $G_t(u_t)$  et  $Z_{T+1}(x_{T+1})$  sont continûment différentiables, on peut alors faire appel à des méthodes générales d'optimisation non linéaire, telles que le gradient réduit généralisé pour lequel des variantes spécifiques [6, 7] permettent de traiter l'exemple proposé dans de très bonnes conditions.

Le cas des fonctions  $G_t(u_t)$  et  $Z_{T+1}(x_{T+1})$  presque partout continûment différentiables (mais non linéaires) peut faire penser à une approche de type sous-gradient déjà développée en matière d'optimisation sans contraintes [8], mais nous n'en connaissons pas d'extension à des problèmes contenant des contraintes autres que de simples contraintes de borne.

(c) Le problème peut aussi se résoudre par programmation dynamique marginaliste. Il suppose des hypothèses complémentaires : les démonstrations que nous donnons ci-après supposent les  $G_t(u_t)$  continûment différentiables et  $Z_{T+1}(x_{T+1})$  presque partout continûment différentiable.

En fait, ainsi que nous l'indiquerons rapidement à la fin de ce document, cette méthode convient également au cas où les fonctions  $G_t(u_t)$  sont seulement presque partout continûment différentiables.

1.3. Pour résoudre les problèmes stochastiques, nous introduisons en général l'hypothèse de stratégies markoviennes de type « hasard-décision » [9], c'est-à-dire que les décisions optimales sont supposées être de la forme :

$$u_t = f_t(x_t, e_t),$$

où les  $f_t$  sont les inconnues.

Il est alors équivalent de résoudre le problème initial et de résoudre le problème de programmation dynamique stochastique défini par la relation de récurrence :

$$Z_t(x_t) = \frac{E}{e_t, u_t, v_t} \text{Max} \{ G_t(u_t, e_t) + Z_{t+1}(x_{t+1}) \},$$

sous les mêmes contraintes que précédemment.

Dans les deux configurations (stochastique et déterministe), les traitements sont extrêmement similaires (à une espérance mathématique près). Nous limiterons donc notre présentation de la programmation dynamique marginaliste au seul cas déterministe.

## 2. CONDITIONS DE KUHN ET TUCKER [11]

2.1. Reprenons l'expression de la fonction objectif (1) avec les contraintes qui lui sont associées :

$$\text{Max}_{u_t, v_t} [G_t(u_t) + Z_{t+1}(x_{t+1})],$$

$$x_{t+1} + u_t + v_t = x_t + e_t, \quad (\lambda)$$

$$u_t \geq u_{\min}, \quad (\mu_1)$$

$$-u_t \geq -u_{\max}, \quad (\mu_2)$$

$$x_{t+1} \geq x_{\min}, \quad (\mu_3)$$

$$-x_{t+1} \geq -x_{\max}, \quad (\mu_4)$$

$$v_t \geq 0, \quad (\mu_5)$$

De manière triviale, les contraintes sont concaves et différentiables. Par ailleurs, la concavité de la fonction objectif s'établirait par récurrence [12]. Reste la question de sa différentiabilité.

Si la fonction  $G_t(u_t)$  est différentiable par hypothèse, il n'en est pas de même de la fonction  $Z_{t+1}(x_{t+1})$ , qui n'est que dérivable par morceaux : sa dérivée admet

des points de discontinuité en nombre fini comme nous le démontrerons plus loin, par récurrence.

En ce qui concerne le cas où  $G_t(u_t)$  est presque partout continûment différentiable, nous renvoyons à la remarque de la fin du paragraphe 3.3.

Nous convenons d'appeler  $x_t^i$ ,  $i = 2, I - 1$ , les points de discontinuité, en leur adjoignant :

$$x_t^1 = x_{\min},$$

$$x_t^I = x_{\max}.$$

Par ailleurs, la notation  $\partial Z_t / \partial x_t$  signifie implicitement, dans l'intervalle  $[x_t^i, x_t^{i+1}]$  :

$$(\alpha) \quad \forall i, \forall x_t \in ]x_t^i, x_t^{i+1}[ : \quad \frac{\partial Z_t}{\partial x_t},$$

$$(\beta) \quad \frac{\partial Z_t}{\partial x_t}(x_t^i) = \lim_{h \rightarrow 0^+} \frac{Z(x_t^i + h) - Z(x_t^i)}{h},$$

$$(\gamma) \quad \frac{\partial Z_t}{\partial x_t}(x_t^{i+1}) = \lim_{h \rightarrow 0^+} \frac{Z(x_t^{i+1} - h) - Z(x_t^{i+1})}{-h}$$

Dans chacun des intervalles de différentiabilité, du type  $[x_t^i, x_t^{i+1}]$ , les contraintes s'écrivent :

$$x_{t+1} + u_t + v_t = x_t + e_t, \quad (\lambda^i)$$

$$u_t \geq u_{\min}, \quad (\mu_1^i)$$

$$-u_t \geq -u_{\max}, \quad (\mu_2^i)$$

$$x_{t+1} \geq x_{t+1}^i, \quad (\mu_3^i)$$

$$-x_{t+1} \geq -x_{t+1}^{i+1}, \quad (\mu_4^i)$$

$$v_t \geq 0, \quad (\mu_5^i)$$

Dans chacun de ces intervalles de différentiabilité, les conditions de Kuhn et Tucker sont donc des conditions nécessaires et suffisantes d'existence d'un optimum :

$$\frac{\partial G_t}{\partial u_t} + \lambda^i + \mu_1^i - \mu_2^i = 0,$$

$$\frac{\partial Z_{t+1}}{\partial x_{t+1}} + \lambda^i + \mu_3^i - \mu_4^i = 0,$$

$$\lambda^i + \mu_5^i = 0.$$

En changeant de notations, après avoir défini les fonctions :  
gain marginal :

$$g_t = \frac{\partial G_t}{\partial u_t},$$

valeur marginale de l'état :

$$z_{t+1} = \frac{\partial Z_{t+1}}{\partial x_{t+1}}$$

Les conditions de Kuhn et Tucker s'écrivent :

$$\begin{array}{l} g_t + \lambda^i + \mu_1^i - \mu_2^i = 0, \\ z_{t+1} + \lambda^i + \mu_3^i - \mu_4^i = 0, \\ \lambda^i + \mu_5^i = 0, \end{array} \quad \mu_1^i, \mu_2^i, \mu_3^i, \mu_4^i, \mu_5^i \geq 0,$$

avec :

$$\begin{cases} \mu_1^i (u_t - u_{\min}) = 0, \\ \mu_2^i (-u_t + u_{\max}) = 0, \\ \mu_3^i (x_{t+1} - x_{t+1}^i) = 0, \\ \mu_4^i (-x_{t+1} + x_{t+1}^{i+1}) = 0, \\ \mu_5^i v_t = 0. \end{cases}$$

Si l'on suppose que les fonctions valeurs marginales de l'état  $z_{t+1}(x_{t+1})$  sont connues, ces premiers résultats se prêtent bien à une interprétation graphique, conduisant à la détermination de l'optimum (cette interprétation que nous présentons ici dans le cadre d'un intervalle de différentiabilité, s'étend, ainsi que nous l'indiquerons par la suite, à l'ensemble du domaine de définition et l'on peut donc envisager le développement qui suit sous ces deux aspects).

On vérifie simplement que les différents cas possibles sont ceux du tableau 1 ci-dessous :

TABLEAU 1

Énumération des cas possibles

|                                               | $x_{t+1}^i < x_{t+1} < x_{t+1}^{i+1}$ | $x_{t+1} = x_{t+1}^i$ | $x_{t+1} = x_{t+1}^{i+1}$ |
|-----------------------------------------------|---------------------------------------|-----------------------|---------------------------|
| $u_{\min} < u_t < u_{\max} \dots \dots \dots$ | Cas 1                                 | Cas 5                 | Cas 4                     |
| $u_t = u_{\min} \dots \dots \dots$            | Cas 2                                 | Cas 6                 | Cas 8                     |
| $u_t = u_{\max}; v_t = 0 \dots \dots \dots$   | Cas 3                                 | Cas 7                 | Cas 9                     |
| $u_t = u_{\max}; v_t > 0 \dots \dots \dots$   |                                       |                       | Cas 10                    |



La représentation graphique des différents cas possibles est directement inspirée des travaux de P. Masse [1] : dans un système d'axes valeur marginale en abscisses, niveau de l'état en ordonnées, on porte la fonction valeur marginale  $z_{t+1}(x_{t+1})$ ; par ailleurs, étant donné un état  $x_t$ , on peut lui associer les états « limites »  $x_{t+1}^1$  et  $x_{t+1}^2$  définis de la manière suivante :

$$x_{t+1}^1 = x_t + e_t - u_{\min},$$

$$x_{t+1}^2 = x_t + e_t - u_{\max}.$$

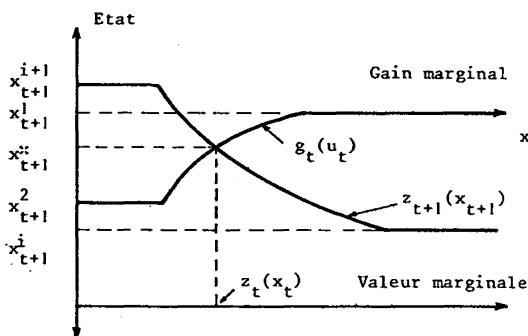
A tout point  $x_{t+1} \in [x_{t+1}^2, x_{t+1}^1]$ , on peut associer une valeur de la commande  $u_t$  :

$$u_t = x_t + e_t - x_{t+1} \quad (x_t \text{ donné})$$

et donc une valeur du gain marginal  $g_t(u_t)$ , qui se trouve ainsi portée dans le système d'axes gain marginal-commande efficace.

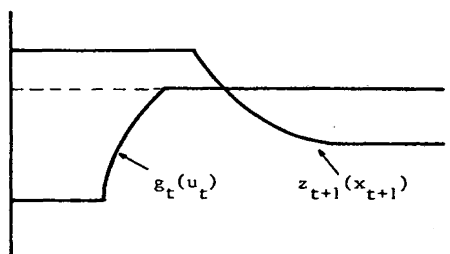
Dans le cas 1, où les contraintes ne sont pas actives, l'optimum  $x_{t+1}^*$  (d'où  $u_t^* = x_{t+1}^1 - x_{t+1}^* + u_{\min}$ ) s'obtient à l'intersection de deux courbes de valeur marginale de l'état et de gain marginal; par ailleurs, ainsi que nous le verrons dans le troisième paragraphe, l'abscisse du point d'intersection fournit la valeur marginale de l'état  $x_t$  au début de l'intervalle  $t$ , soit  $z_t(x_t)$ .

Dans le cas de contraintes actives, un prolongement judicieux des courbes permet de retrouver l'optimum à l'intersection, comme le montrent les différents graphes ci-après représentant les principaux cas possibles :

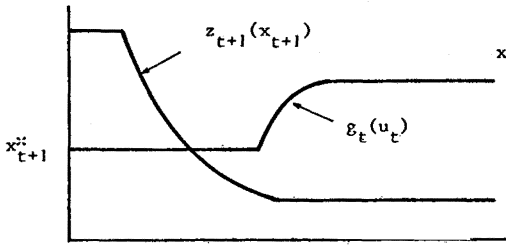


Commande efficace

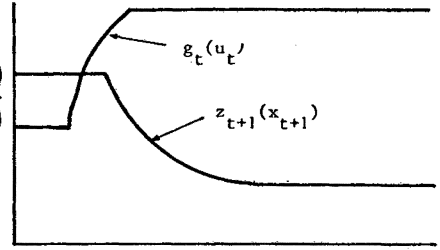
CAS 1.



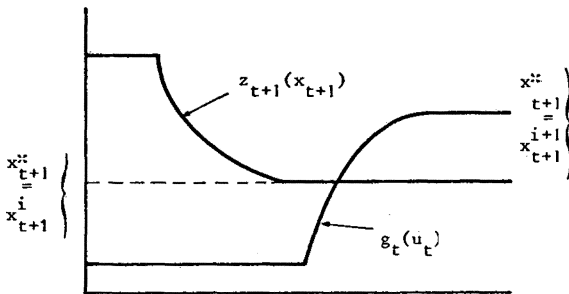
CAS 2.



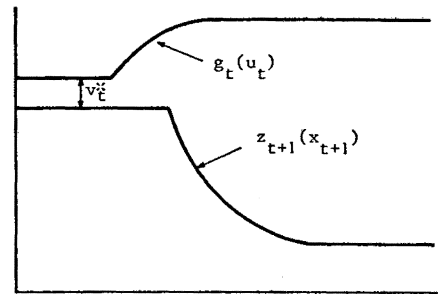
CAS 3.



CAS 4.



CAS 5.



CAS 10.

2.2. Les fonctions valeurs marginales de l'état n'étant pas *a priori* connues, les conditions de Kuhn et Tucker permettent leur détermination par la mise en œuvre d'une relation de récurrence.

Au préalable, il est nécessaire de déterminer le signe de l'expression [gain marginal-valeur marginale de l'état] dans les différents cas possibles. Cette détermination s'effectue simplement à partir des conditions de Kuhn et Tucker, aussi nous limiterons-nous à un seul cas, puis présenterons directement l'ensemble des résultats, sur le tableau II.

Cas 1

$$\begin{aligned} u_t &\in ]u_{\min}, u_{\max}[ , \\ x_{t+1} &\in ]x_{t+1}^i, x_{t+1}^{i+1}[ , \\ v_t &\geq 0, \end{aligned}$$

donc :

$$\begin{aligned} \mu_1^i &= \mu_2^i = 0 \quad \text{et} \quad \mu_3^i = \mu_4^i = 0, \quad \mu_5^i \neq 0, \\ g^t + \lambda^i &= 0, \\ z_{t+1} + \lambda^i &= 0, \end{aligned}$$

et donc :

$$g_t - z_{t+1} = 0.$$

TABLEAU II

Signe de la différence : Gain marginal-valeur marginale de l'état ( $g_t - z_{t+1}$ )

|                                               | $x_{t+1}^i < x_{t+1} < x_{t+1}^{i+1}$ | $x_{t+1} = x_{t+1}^i$ | $x_{t+1} = x_{t+1}^{i+1}$ |
|-----------------------------------------------|---------------------------------------|-----------------------|---------------------------|
| $u_{\min} < u_t < u_{\max} \dots \dots \dots$ | $= 0$                                 | $> 0$                 | $< 0$                     |
| $u_t = u_{\min} \dots \dots \dots$            | $< 0$                                 | Indéterminé           | $< 0$                     |
| $u_t = u_{\max}; v_t = 0 \dots \dots \dots$   | $> 0$                                 | $> 0$                 | Indéterminé               |
| $u_t = u_{\max}; v_t > 0 \dots \dots \dots$   |                                       |                       | Indéterminé               |

2.3. Nous avons vu, jusqu'ici, qu'étant donné un état  $x_t$  à une date  $t$ , il existait  $I$  intervalles où l'optimisation était possible, et que l'on obtenait ainsi  $I$  optima locaux.

On démontre simplement qu'il n'existe qu'un seul optimum global, auquel on peut associer un intervalle de différentiabilité, à l'aide des propositions suivantes :

(a) S'il existe un intervalle  $[x_{t+1}^i, x_{t+1}^{i+1}]$  où l'optimum est libre (par rapport aux contraintes  $x_{t+1}^i \leq x_{t+1}$  et  $x_{t+1} \leq x_{t+1}^{i+1}$ ), alors nécessairement, dans tous les autres intervalles, l'optimum est contraint et moins bon que l'optimum libre : ceci se démontrerait facilement en utilisant les résultats précédents. Dans ce cas, on élimine tous les intervalles à optimum bloqué, pour ne garder que celui où il est libre.

(b) S'il n'existe pas d'intervalle  $[x_{t+1}^i, x_{t+1}^{i+1}]$  où l'optimum est libre, il est alors possible d'effectuer une partition de l'ensemble  $I$  :

- de  $i=1$  à  $i_0-1$  : les intervalles  $[x_{t+1}^i, x_{t+1}^{i+1}]$  admettent un optimum contraint à la borne supérieure  $x_{t+1}^{i+1}$ ;
- de  $i=i_0$  à  $I$  : les intervalles  $[x_{t+1}^i, x_{t+1}^{i+1}]$  admettent un optimum contraint à la borne inférieure  $x_{t+1}^i$ .

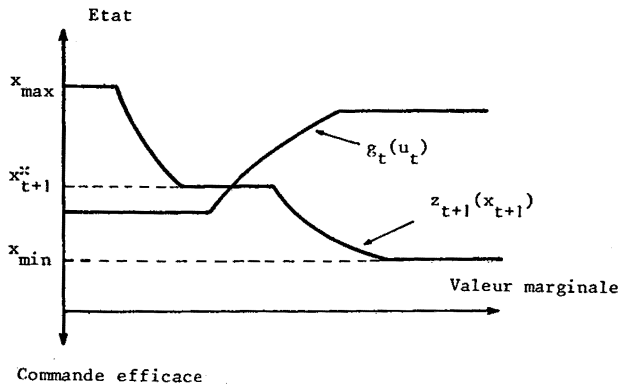
De ce fait, les intervalles  $[x_{t+1}^{i_0-1}, x_{t+1}^{i_0}]$  et  $[x_{t+1}^{i_0}, x_{t+1}^{i_0+1}]$  admettent le même optimum en  $x_{t+1}^{i_0}$ , et cet optimum est supérieur à tous ceux des autres intervalles, qu'il convient donc d'éliminer.

(De même que ci-dessus, ces résultats se démontreraient à l'aide des résultats précédents, et en particulier le signe de l'expression [gain marginal-valeur marginale], ainsi que la concavité des fonctions  $G_t$  et  $Z_t$ .)

Il reste alors à éliminer arbitrairement l'un des deux intervalles  $[x_{t+1}^{i_0-1}, x_{t+1}^{i_0}]$  et  $[x_{t+1}^{i_0}, x_{t+1}^{i_0+1}]$ , qui conduisent au même optimum. Convenons de ne garder que l'intervalle  $[x_{t+1}^{i_0}, x_{t+1}^{i_0+1}]$ .

**2.4.** Ces dernières propositions permettent de conclure à l'existence d'un optimum absolu unique sur l'ensemble des intervalles, optimum associé à un intervalle de différentiabilité particulier. Dans les paragraphes suivants, nous nous limiterons donc à cet intervalle particulier.

Ces résultats montrent également que la construction graphique présentée dans le paragraphe 2.1 permet bien de trouver l'optimum en recherchant directement l'intersection des fonctions  $z_{t+1}(x_{t+1})$  et  $g_t(u_t)$ , ainsi que l'illustre le schéma ci-après :



### 3. CALCUL PAR RÉCURRENCE DES VALEURS MARGINALES DE L'ÉTAT

#### 3.1. Établissement d'une relation de récurrence

Nous substituons ici, à la relation de Bellmann, une relation de récurrence sur les valeurs marginales de l'état.

La relation de Bellmann s'écrit :

$$Z_t(x_t) = \text{Max}_{u_t, v_t} [G_t(u_t) + Z_{t+1}(x_{t+1})].$$

Étant donné l'état  $x_t$ , nous cherchons à déterminer sa valeur marginale  $z_t(x_t)$ , connaissant  $g_t$  et  $z_{t+1}$ .

Posons :

$u_t^*, v_t^*, x_{t+1}^*$  = valeurs de  $u_t, v_t, x_{t+1}$  à l'optimum;

$u_t^*, v_t^*$  et  $x_{t+1}^*$  sont donc fonctions de  $x_t$  et sont liées par la relation :

$$x_{t+1}^* = x_t + e_t - u_t^* - v_t^*,$$

une variation  $dx_t$  entraîne donc des variations  $du_t^*$ ,  $dv_t^*$ ,  $dx_{t+1}^*$  liées par la relation :

$$dx_{t+1}^* = dx_t - du_t^* - dv_t^*.$$

Par définition :

$$z_t(x_t) = \lim_{dx_t \rightarrow 0} \left[ \frac{Z_t(x_t + dx_t) - Z_t(x_t)}{dx_t} \right].$$

Étudions l'expression  $z_t(x_t) dx_t$ ,  $dx_t$  étant infiniment petit :

$$\begin{aligned} z_t(x_t) dx_t &= \text{Max}_{du_t^*, dv_t^*} [G_t(u_t^* + du_t^*) \\ &\quad + Z_{t+1}(x_{t+1}^* + dx_{t+1}^*) - G_t(u_t^*) - Z_{t+1}(x_{t+1}^*)], \\ z_t(x_t) dx_t &= \text{Max}_{du_t^*, dv_t^*} [g_t du_t^* + z_{t+1} dx_{t+1}^*] \end{aligned}$$

or :

$$dx_{t+1}^* = dx_t - du_t^* - dv_t^*,$$

d'où :

$$z_t(x_t) dx_t = z_{t+1}(x_{t+1}) dx_t + \text{Max}_{du_t^*, dv_t^*} [(g_t - z_{t+1}) du_t^* - z_{t+1} dv_t^*],$$

d'où la relation de récurrence qui lie  $z_t(x_t)$  à  $z_{t+1}(x_{t+1})$  :

$$z_t = z_{t+1} + \lim_{dx_t \rightarrow 0} \text{Max}_{du_t^*, dv_t^*} \left[ \frac{(g_t - z_{t+1}) du_t^* - z_{t+1} dv_t^*}{dx_t} \right]. \quad (2)$$

### 3.2. Mise en œuvre de cette relation de récurrence

Les conditions de Kuhn et Tucker, analysées au chapitre 2, auxquelles nous joignons la relation (2) établie ci-dessus, permettent de définir la relation de récurrence cherchée.

*Remarque préalable* nécessaire pour la suite de l'étude :

Une variation  $dx_t$  entraîne, par l'intermédiaire des fonctions  $u_t^*(x_t)$  et  $v_t^*(x_t)$ , une variation  $dx_{t+1}^*$  :

$$dx_{t+1}^* = dx_t - du_t^* - dv_t^*.$$

Or, il découle du principe d'optimalité que deux trajectoires optimales ne peuvent se couper, et donc  $dx_t$  et  $dx_{t+1}^*$  sont de même signe.

En particulier, dans les cas où  $dv_t^* = 0$ :

$$du_t^* = dx_t - dx_{t+1}^*.$$

Donc :

$$\begin{aligned} \text{si } dx_t > 0, & \quad \text{alors } du_t^* \leq dx_t, \\ \text{si } dx_t < 0, & \quad \text{alors } du_t^* \geq dx_t. \end{aligned}$$

Reprenons, pour la mise en œuvre de la relation de récurrence (2), le même découpage des cas possibles qu'au chapitre 2. Dans chaque cas, le maximum de l'expression  $(g_t - z_{t+1}) du_t^*$  sera atteint, suivant le signe de  $(g_t - z_{t+1})$  pour l'une ou l'autre des deux bornes de l'intervalle de variation possible en  $du_t^*$ .

De même que précédemment, nous limiterons l'exposé analytique à quelques cas représentatifs, puis présenterons un tableau récapitulant l'ensemble des résultats (tableau III).

Cas 1

$$\begin{aligned} u_t &\in ]u_{\min}, u_{\max}[ , \\ x_{t+1} &\in ]x_{\min}, x_{\max}[ , \\ v_t &= 0, \end{aligned}$$

$g_t = z_{t+1}$ , et donc :

$$z_t = z_{t+1} = g_t$$

Cas 4

$$\begin{aligned} u_t &\in ]u_{\min}, u_{\max}[ , \\ x_{t+1} &= x_{\max}, \\ v_t &= 0, \end{aligned}$$

alors  $dx_{t+1} \leq 0$ , d'où  $dx_t \leq 0$ , d'où  $du_t \geq dx_t$ , et puisque  $(g_t - z_{t+1}) < 0$  le maximum est atteint pour  $du_t^* = dx_t$  et donc :

$$z_t = g_t$$

Cas 6

$$u_t = u_{\min},$$

$$x_{t+1} = x_{\min},$$

$$v_t = 0,$$

$$du_t \geq 0, \quad dx_{t+1} \geq 0$$

or :

$$du_t = dx_t - dx_{t+1}.$$

Donc nécessairement  $dx_t \geq 0$ , d'où  $0 \leq du_t^* \leq dx_t$ .

On ne sait rien du signe de l'expression  $(g_t - z_{t+1})$  mais, quoiqu'il en soit :

$$z_t = \sup(g_t, z_{t+1}).$$

Cas 8

$$u_t = u_{\min},$$

$$x_{t+1} = x_{\max},$$

$$v_t = 0.$$

On sait que, dans ce cas :

$$(g_t - z_{t+1}) \leq 0.$$

Il faut là encore, distinguer la dérivée à droite et la dérivée à gauche :

— si  $dx_t \geq 0$ , d'où  $dx_{t+1} = 0$  nécessairement, car il doit être de même signe que  $dx_t$ , et vérifier  $dx_{t+1} \leq 0$ , et donc:  $du_t = dx_t$ .

D'où :

$$(z_t)_d = g_t ;$$

— si  $dx_t < 0$ , d'où  $dx_t \leq 0 \leq du_t$ .

Le maximum est atteint pour  $du_t = 0$  et donc :

$$(z_t)_g = z_{t+1} ,$$

dans ce cas, la fonction  $z_t(x_t)$  est discontinue.

TABLEAU III

*Relation de récurrence sur les valeurs marginales de l'état,  
suivant les différents cas possibles  
Valeurs de  $z_t(x_t)$*

|                                           | $x_{\min} < x_{t+1} < x_{\max}$ | $x_{t+1} = x_{\min}$                     | $x_{t+1} = x_{\max}$                            |
|-------------------------------------------|---------------------------------|------------------------------------------|-------------------------------------------------|
| $u_{\min} < u_t < u_{\max} \dots$         | $z_{t+1} = g_t$                 | $g_t$                                    | $g_t$                                           |
| $u_t = u_{\min} \dots$                    | $z_{t+1}$                       | $\sup(g_t, z_{t+1})$                     | à droite : $g_t$<br>à gauche : $z_{t+1}$        |
| $u_t = u_{\max} \dots$<br>$v_t = 0 \dots$ | $z_{t+1}$                       | à droite : $z_{t+1}$<br>à gauche : $g_t$ | à droite : 0<br>à gauche : $\inf(g_t, z_{t+1})$ |
| $u_t = u_{\max} \dots$<br>$v_t > 0 \dots$ |                                 |                                          | 0                                               |

Remarquons donc qu'à chaque pas de la variable de temps  $t$ , il ne peut apparaître plus de trois discontinuités nouvelles pour la fonction  $z_t(x_t)$ . De la sorte, après  $N$  pas de la récurrence, le nombre de discontinuités de la fonction  $z_{T-N+1}(x_{T-N+1})$  est majoré par  $3 \times N$ .

Ceci établit donc que la fonction  $Z_t$  du paragraphe 2 est bien presque partout différentiable (les problèmes traités en pratique ont en effet un horizon fini).

### 3.3. Interprétation graphique de ces résultats

Reprenant l'interprétation graphique du chapitre 2, il est possible de retrouver les cas de discontinuité que l'analyse vient de mettre en évidence :

Une variation de l'état  $x_t$  entraîne un déplacement vers le haut ou le bas de la courbe de gain marginal. On étudie donc la variation de la valeur marginale  $z_t(x_t)$  en fonction de  $x_t$ , en observant l'évolution de l'abscisse du point d'intersection des deux courbes, quand la courbe de gain marginal se déplace parallèlement à l'axe des ordonnées.

*Cas 7*

$$u_t = u_{\max},$$

$$x_{t+1} = x_{\min},$$

$$v_t = 0.$$

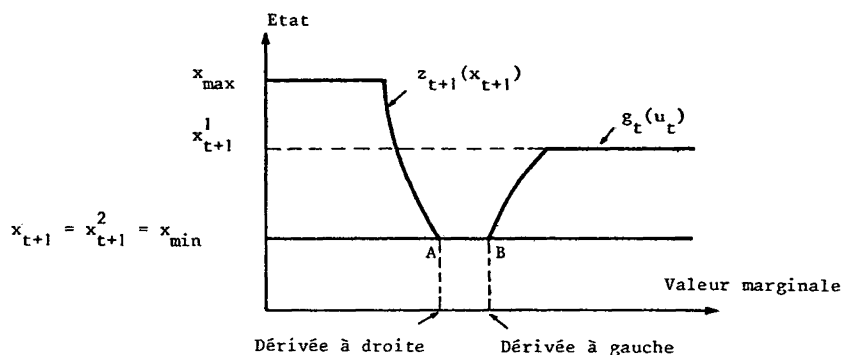
Dans ce cas, comme le montre la figure suivante, l'intersection des deux courbes est indéterminée sur le segment  $AB$  :

— soit une variation positive de  $x_t$  : la courbe de gain marginal se décale vers le haut et l'intersection vient sur la courbe de valeur marginale (point  $A$ ) : on retrouve bien  $(z_t)_d = z_{t+1}$ ;



soit une variation négative de  $x_t$  : la courbe de gain marginal se décale vers le bas, et l'intersection vient sur la courbe de gain marginal (point  $B$ ) donc :  $(z_t)_g = g_t$ .

Le cas 8 conduirait à des observations identiques.



Cas 9

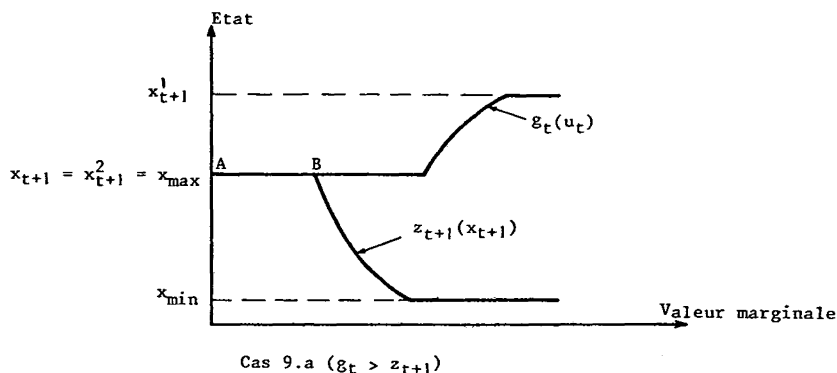
$$u_t = u_{\max},$$

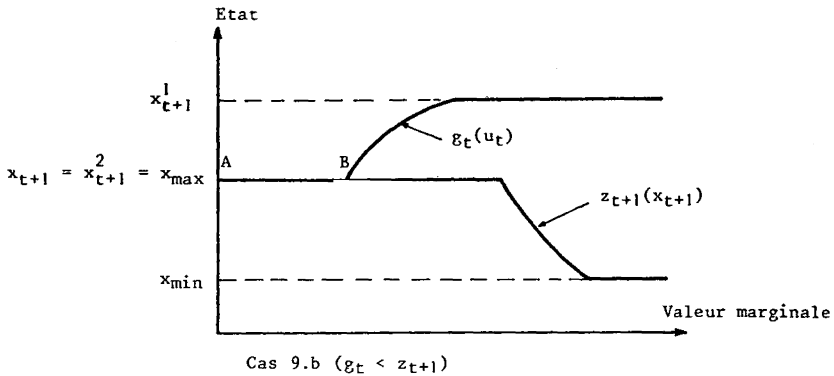
$$x_{t+1} = x_{\max},$$

$$v_t = 0.$$

Deux situations sont alors à considérer comme le montrent les deux schémas suivants.

L'intersection est, là encore, indéterminée sur le segment  $AB$ . Dans les deux cas, pour une variation positive de  $x_t$ , la courbe de gain marginal se décale vers le haut, et l'intersection vient sur l'axe des ordonnées :  $(z_t)_d = 0$ .





Pour une variation négative de  $x_t$  :

Cas 9 (a) : l'intersection vient sur la courbe de valeur marginale :

$$(z_t)_g = z_{t+1};$$

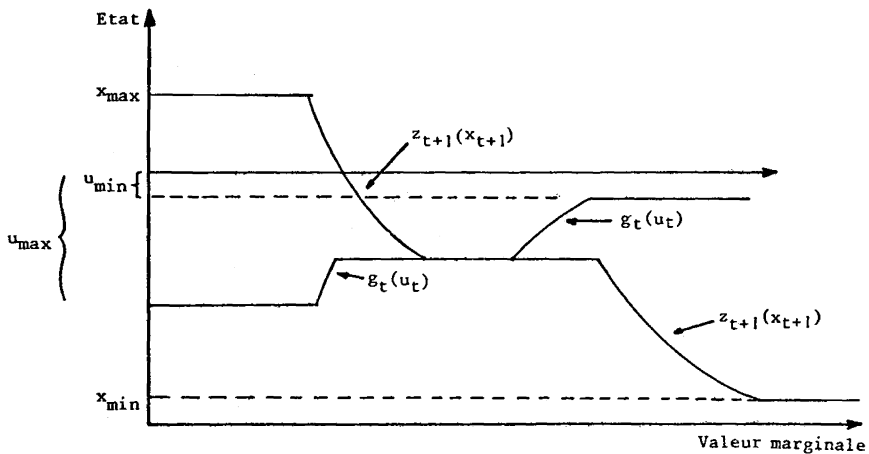
Cas 9 (b) : l'intersection vient sur la courbe de gain marginal :

$$(z_t)_g = g_t;$$

Dans les deux cas l'interprétation graphique rejoint la théorie :

$$(z_t)_g = \inf(g_t, z_{t+1}).$$

REMARQUE : Nous avons développé le cas où  $G_t(u_t)$  est de classe C1. Les différents résultats montrent que le raisonnement par intervalle de différentiabilité s'étend aisément au cas où  $G_t(u_t)$  est seulement presque partout continûment différentiable. Nous illustrerons simplement ce point à l'aide du cas suivant :



Il n'y a toujours pas d'ambiguïté quant au choix du  $u_t$  optimal : il est unique; par contre il apparaît l'existence d'une dérivée à gauche et d'une dérivée à droite au niveau de la relation de récurrence sur les valeurs marginales (tous ces cas sont d'ailleurs répertoriés sur le tableau III), à savoir pour cette configuration :

- dérivée à droite :  $\sup(z_{t+1}, g_t)$ , correspondant au cas :

$$u_t = u_{\min}, \quad x_{t+1} = x_{\min}$$

du tableau III;

- dérivée à gauche :  $\inf(g_t, z_{t+1})$ , correspondant au cas :

$$u_t = u_{\max}, \quad x_{t+1} = x_{\max}, \quad v_t = 0$$

du tableau III.

#### 4. CONCLUSION

(a) Les méthodes de programmation dynamique voient leur utilisation limitée par le nombre de variables d'état admissible : 2 à 3 au maximum pour la procédure classique, et 1 pour la procédure marginaliste. En revanche, pour la résolution de problèmes stochastiques, ces méthodes apportent un avantage important par leur aptitude à représenter le caractère auto-adaptatif des systèmes : la stratégie définie par  $u_t = f_t(x_t, e_t)$  est en boucle fermée, c'est-à-dire que la décision optimale à un instant  $t$  dépend non seulement de l'instant, mais aussi de l'état du système effectivement réalisé au début de l'instant  $t$ .

De plus, les méthodes de programmation dynamique sont simples et robustes, ce qui, dans certains cas, peut constituer un avantage par rapport à des techniques d'optimisation beaucoup plus raffinées, d'autant que les performances sont quand même tout à fait raisonnables.

(b) La programmation dynamique marginaliste supplante la programmation dynamique classique pour certains types de problèmes, par ses qualités propres :

1. Lorsque  $T$  est assez grand ( $T$  supérieur à 100 par exemple), la programmation dynamique classique nous a donné des résultats décevants : en effet, après un certain nombre d'itérations sur la relation de récurrence, les fonctions valeurs de l'état, devenues très grandes, ne sont plus commensurables avec les gains instantanés. L'optimisation se heurte à un problème de précision. L'avantage de la programmation dynamique marginaliste réside en l'utilisation de quantités qui restent constamment du même ordre de grandeur : les dérivées des fonctions valeurs du stock et gains instantanés.

2. Enfin, la programmation dynamique marginaliste dispose de l'avantage qualitatif non négligeable que constitue la plus grande facilité d'interprétation des résultats marginaux.

En contrepartie de ces avantages, le champ d'application de cette méthode est réduit : alors que la programmation dynamique classique s'applique sous des conditions très larges, la programmation dynamique marginaliste exige que les fonctions soient concaves et presque partout continûment différentiables.

(c) Cette étude devrait donc constituer une contribution non négligeable à la connaissance de la programmation dynamique marginaliste, dans la mesure où elle éclaire certains aspects restés, semble-t-il, assez obscurs jusqu'à présent, ne serait-ce que sur le simple plan de l'optimalité de cette technique.

L'extension de cette méthode aux problèmes en dimension 2 (état à 2 dimensions), pour intéressante qu'elle soit, se heurte à de délicates questions de différentiabilité, qu'il ne nous semble pas possible de résoudre par simple généralisation de la théorie ici présentée.

#### BIBLIOGRAPHIE

1. P. MASSE, *Les réserves et la régulation de l'avenir dans la vie économique*, tomes I et II, Hermann et Cie, 1946.
2. F. MESLIER, *Simulation of the Optimal Control of a Power System : Latest Developments of the Greta Model*, Sixth Power Systems Computation Conference, vol. 1, p. 558-567, Darmstadt, August 1978.
3. F. MESLIER et P. SANDRIN, *Le modèle Niagara. Détermination des hypothèses de production hydraulique nécessaire aux études de réseau*, note interne E.D.F. n° HR 31.0490.
4. R. E. BELLMANN et S. E. DREYFUS, *La programmation dynamique et ses applications*, Dunod, paris, 1965.
5. J. F. MAURRAS, *Optimisation of the Flow Through Networks With Gains*, Mathematical programming, vol. 3, n° 2, octobre 1972.
6. J. ABADIE et M. BICHARA, *Résolution de certains problèmes de commande optimale*, R.A.I.R.O., vol. V2, mai 1973, p. 77-109.
7. J. ABADIE, *Application of the GRG Algorithm to Optimal Control Problems*, chap. 8 In *Integer and Non Linear Programming*, J. ABADIE, éd., North-Holland Publishing Company, Amsterdam, 1970.
8. LEMARECHAL, *An Algorithm for Minimizing Convex Functions*, *Informations Processing*, 1974, North-Holland Publishing Company.
9. A. KAUFMANN et R. CRUON, *La programmation dynamique. Gestion scientifique séquentielle*, Dunod, Paris, 1965.
10. F. FALGARONE, P. LEDERER et BUCHER, *Optimal Operation of the Large French Hydropower System*. Sixth Power Systems Computation Conference, vol. 1, p. 568-575, Darmstadt, august 1978.
11. J. ABADIE, *On the Kuhn Tucker Theorem in Non Linear Programming*, North-Holland Publishing Company, Amsterdam.
12. G. HADLEY, *Non Linear and Dynamic Programming*, Addison, Wesley Publishing Company.