HAL
open science

# On an algorithm for optimal control using Pontryagin's maximum principle

J. Frederic Bonnans

## To cite this version:

J. Frederic Bonnans. On an algorithm for optimal control using Pontryagin's maximum principle. [Research Report] RR-0254, Inria. 1983. inria-00076304

**HAL Id: inria-00076304**

**https://inria.hal.science/inria-00076304**

Submitted on 24 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IRIA

CENTRE DE ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél. (3) 954 90 20

Rapports de Recherche

N° 254

# ON AN ALGORITHM FOR OPTIMAL CONTROL USING PONTRYAGIN'S MAXIMUM PRINCIPLE

Joseph Frédéric BONNANS

Décembre 1983

# ON AN ALGORITHM FOR OPTIMAL CONTROL USING PONTRYAGIN'S MAXIMUM PRINCIPLE

Joseph Frédéric BONNANS
INRIA
Domaine de Voluceau
BP 105 - Rocquencourt
78153 LE CHESNAY Cedex (France)

RESUME

Y. Sakawa et Y. Shindo ont récemment proposé un algorithme de résolution des problèmes de contrôle en boucle ouverte, qui utilise de façon directe le principe du maximum de Pontryagin. Nous établissons ici que, sous certaines hypothèses, l'algorithme est bien défini et converge globalement, dans un sens faible. Lorsque le pas de l'algorithme tend vers zéro, les déplacements calculés sont équivalents à ceux d'une méthode de gradient avec projection. Enfin, si le contrôle intervient de façon linéaire dans l'équation d'état et quadratique dans le critère, l'algorithme s'interprète comme une méthode de gradient avec projection dans une certaine métrique.

ABSTRACT

Y. Sakawa and Y. Shindo recently proposed an algorithm to solve open-loop optimal control problems, using Pontryagin's maximum principle. It is established here that, under some hypothesis, the algorithm is well-defined and globally converges in some weak sense. When the stepsize of the algorithm tends to zero, the displacements are equivalent to those of a gradient with projection method. If the control enters in a linear way in the state equation and in a quadratic way in the criterion, the algorithm can be interpreted as a gradient plus projection method in some new metric.

PAPIER RECUPERE ET RECYCLE

## I - INTRODUCTION

We are concerned with the open-loop optimal control of a system governed by an ordinary differential equation. We assume that there are no state constraints and that the control constraints are local with respect to the time. Y. Sakawa and Y. Shindo [3] proposed the following iterative algorithm : a control and its state being given, compute the associated costate, then compute simultaneously a new control and a new state such that the state equation holds and that, at each time, the new control minimizes some function. This function is the sum of the Hamiltonian and of a quadratic term penalizing the difference between the new and the old controls. It is proved in [3] that, under some hypothesis, the sequence of the criterion decreases and that, if a subsequence of the controls converges a.e. towards some control $\bar{u}$, then $\bar{u}$ checks the first-order necessary optimality conditions.

We obtain here the following results :

- Under some regularity hypothesis, the algorithm is well-defined.

- The norm of the projected gradients of the controls computed by the algorithm tends towards zero. If the problem is convex, we deduce from it the weak convergence towards some optimal control.

- If the step size tends towards zero, the controls computed by the algorithm are equivalent in some way to those computed by a gradient plus projection method.

- If the control enters in a linear way in the state equation and in a quadratic way in the criterion, the algorithm reduces to some gradient plus projection method in some new metric.

## II - DEFINITION AND WELL-POSEDNESS OF THE ALGORITHM

We consider the following system

$$(2.1)\quad \begin{cases} \frac{dy}{dt}(t) = f(y(t),\ u(t),\ t), \text{a.e. } t \in (0,T), \\ y(0) = 0\ ; \end{cases}$$

with $T > 0$ given, $y(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^p$. To this system is associated the following criterion

$$(2.2)\qquad J(u) = \int_0^T l(y(t),\ u(t),\ t)dt + g(y(T)).$$

Let $\mathcal{U}_{ad}$ be a closed, bounded, convex subset of $\mathbb{R}^p$. The control problem is

$$(2.3)\quad \begin{cases} \text{Min } J(u) \\ u(t) \in \mathcal{U}_{ad},\ \text{a.e. } t \in (0,T). \end{cases}$$

A control u is said to be admissible if $u(t) \in \mathcal{U}_{ad}$, a.e. $t \in (0,T)$ and $J(u)$ makes sense.

The Hamiltonian of the problem is

$$(2.4)\quad \begin{cases} H : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}\ , \\ (y,u,p,t) \to l(y,u,t) - p^t f(y,u,t). \end{cases}$$

The costate equation can be written as

$$(2.5)\quad \begin{cases} \dot{p}(t) = \frac{\partial H}{\partial y}(y(t),\ u(t),\ p(t),\ t),\ \text{a.e. } t \in (0,T)\ , \\ p(T) = \frac{\partial g}{\partial y}(y(T)). \end{cases}$$

We denote by $||.||_p$ the euclidean norm in $\mathbb{R}^p$. Define the augmented hamiltonian

$$(2.6)\quad \begin{aligned} &K_\varepsilon : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \\ &(y,u,v,p,t) \to H(y,u,p,t) + \frac{1}{2\varepsilon}\ ||u-v||_p^2. \end{aligned}$$

Here is the algorithm proposed by Y. Sakawa and Y. Shindo [3] :

*Algorithm 1*

0) Let some admissible control $u^0(t)$ and a sequence $\{\varepsilon^k\}$ of positive numbers be given. Set $k = 0$. Compute $y^0$, the state associated to $u^0$.

1) Compute the costate associated to $u^k(t)$.

2) $k = k+1$. Compute $u^k$, $y^k$ such that $y^k$ is the state associated to $u^k$ and

$$K_{\varepsilon^k}(y^k, u^k, u^{k-1}, p^{k-1}, t) \le K_{\varepsilon^k}(y^k, u, u^{k-1}, p^{k-1}, t), \ \forall u \in \mathcal{U}_{ad}, \quad \text{a.e. } t \in (0,T).$$

3) Stop if some convergence test is checked. Otherwise, go to 1. □

We now state the hypothesis made on f, l, g :

(2.7) f is twice continuously differentiable.

(2.8) $\begin{cases} \text{There exists } C_1 > 0 \text{ such that} \\ u \text{ admissible} \Rightarrow \ ||y(t,u)||_{\mathbb{R}^n} < C_1, \ \forall t \in [0,T]. \end{cases}$

(2.9) $l, \frac{\partial l}{\partial y}, \frac{\partial l}{\partial u}, \frac{\partial^2 l}{\partial y^2}, \frac{\partial^2 l}{\partial y \partial u}$ are continuous on $\mathbb{R}^n \times \mathbb{R}^p \times [0,T]$.

(2.10) $\begin{cases} \text{There exists } \mu > 0 \text{ such that, } C_1 \text{ being the constant in (2.8) :} \\ ||z||_n < C_1 \\ \text{implies that the application} \\ u \to l(z,u,t) + \mu ||u||_p^2 \\ \text{is convex.} \end{cases}$

(2.11) g is twice continuously differentiable from $\mathbb{R}^n$ onto $\mathbb{R}$.

*Remark 2.1.*

These hypothesis are less restrictive that those of [3] in which g is null and $\frac{\partial^2 H}{\partial u^2}$ is supposed to exist and to be a positive matrix. □

We note that (2.8) holds if

$$(2.12)\quad \begin{cases} \exists C_2 > 0 \text{ such that} \\ |f(y,u,t)| \leq C_2(1 + ||y||_p),\ \forall (y,u) \in \mathbb{R}^n \times \mathcal{U}_{ad}. \end{cases}$$

We now give a mathematical sense to (2.1)-(2.5). As $\mathcal{U}_{ad}$ is bounded, (2.6)(2.7) (2.8) imply that y is in

$$Y = \{y \in L^\infty(0,T,\mathbb{R}^n)\ ;\ \frac{dy}{dt} \in L^\infty(0,T,\mathbb{R}^n)\}.$$

As $Y \subset \mathcal{C}(0,T,\mathbb{R}^n)$, the initial condition of (2.1) makes sense. So does also the criterion. We easily check that if y(u) exists, (2.5) has a unique solution in Y and that there exists $\delta > 0$ such that

$$||y(t)||_n < \delta,\ ||p(t)||_n < \delta.$$

We denote

$$E = \{(z,q) \in \mathbb{R}^n \times \mathbb{R}^n\ ;\ ||z||_n < \delta \text{ and } ||q||_n < \delta\}.$$

*Theorem 2.1.*

Under hypothesis (2.7) to (2.11), there exists $\varepsilon_0 > 0$ such that, if $\varepsilon^k < \varepsilon_0$, $\forall k = 1, 2, \ldots$, algorithm 1 defines a sequence $\{u^i\}$ (non necessarily unique) of admissible controls. □

Before proving the theorem, we first give the :

*Lemma 2.1.*

Let $t \in [0,T]$ be given. Define the (a priori multivoque) application

$$E \times \mathcal{U}_{ad} \to \mathbb{R}^p ,$$

$$(z,q,v) \to u_\varepsilon(z,q,v),$$

defined by

$$(2.13) \quad \begin{cases} u_\varepsilon \in \mathcal{U}_{ad}, \\ K_\varepsilon(z, u_\varepsilon, v, q, t) \le K_\varepsilon(z, u, v, q, t), \ \forall u \in \mathcal{U}_{ad}. \end{cases}$$

Denote $w = (z,q)$. Under hypothesis (2.7) to (2.10), for any $\alpha > 0$, there exists $\beta > 0$ and $\varepsilon_0 > 0$, independant of t, such that, if $\varepsilon < \varepsilon_0$, problem (2.13) has a unique solution $u_\varepsilon$ such that :

$$(2.14) \quad ||u_\varepsilon(w_1,v) - u_\varepsilon(w_2,v)||_p^2 \le \alpha ||w_1-w_2||_{2n}, \ \forall w_1,w_2 \in E, \ v \in \mathcal{U}_{ad},$$

$$(2.15) \quad ||u_\varepsilon(w,v_2) - u_\varepsilon(w,v_1)||_p \le (1+\alpha)||v_2-v_1||_p, \ \forall w \in E, \ v_1,v_2 \in \mathcal{U}_{ad},$$

$$(2.16) \quad ||u_\varepsilon(w,v) - v||_p \le \beta \varepsilon, \ \forall w \in E, \ v \in \mathcal{U}_{ad}. \quad \square$$

*Proof of Lemma 2.1.*

From (2.7)(2.9)(2.10) we deduce the existence of $C_1 > 0$ such that, for any $(z,q) \in E$, $t \in [0,T]$, $v,u_1,u_2 \in \mathcal{U}_{ad}$ :

$$(2.17) \quad \begin{aligned} K_\varepsilon(z,u_2,v,q,t) \ge K\,(z,u_1,u,q,t) + \frac{\partial K_\varepsilon}{\partial u}\,(z,u_1,v,q,t)^t(u_2-u_1) + \\ + (\frac{1}{2\varepsilon} - c_2)||u_2-u_1||_p^2. \end{aligned}$$

Consequently, if $\varepsilon < 1/2C_2$, $K_\varepsilon$ is strictly convex with respect to u : this implies that (2.13) then has a unique solution. From now, on we suppose that $\varepsilon < \varepsilon_0$, with $\varepsilon_0 < 1/2C_2$.

We remark that the first order necessary optimality condition of (2.13) is

$$\left(\frac{\partial K_\varepsilon}{\partial u}(z,\ u_\varepsilon,\ v,\ q,\ t),\ u-u_\varepsilon\right)_{\mathbb{R}^p} \geq 0,\ \forall u \in \mathcal{U}_{ad},$$

so that (2.19) implies

$$K_\varepsilon(z,u,v,q,t) \geq K_\varepsilon(z,\ u_\varepsilon,\ v,\ q,\ t,) + \left(\frac{1}{2\varepsilon} - C_1\right)||u-u_\varepsilon||_p^2 .$$

We fixe $v \in \mathcal{U}_{ad}$, $t \in [0,T]$ and denote $w = (z,p)$ and

$$\phi_\varepsilon(w,u) = K_\varepsilon(z,u,v,q,t).$$

The above inequality can be written as

$$\phi_\varepsilon(w,u) \geq \phi_\varepsilon(w,u_\varepsilon) + \left(\frac{1}{2\varepsilon} - C_1\right)||u-u_\varepsilon||_p^2 .$$

Let $w_1,w_2$ be in E and denote $u_i = u(w_i,v)$, $i = 1, 2$. Then

$$\phi_\varepsilon(w_1,u_2) \geq \phi_\varepsilon(w_1,u_1) + \left(\frac{1}{2\varepsilon} - C_1\right)||u_2-u_1||_p^2,$$

$$\phi_\varepsilon(w_2,u_1) \geq \phi_\varepsilon(w_2,u_2) + \left(\frac{1}{2\varepsilon} - C_1\right)||u_2-u_1||_p^2 ,$$

and so, by addition

$$\phi_\varepsilon(w_1,u_2) - \phi_\varepsilon(w_2,u_2) + \phi_\varepsilon(w_2,u_1) - \phi_\varepsilon(w_1,u_1) \geq \left(\frac{1}{\varepsilon} - 2C_1\right)||u_2-u_1||_p^2.$$

This implies

$$2 \sup_{w\in E}\left|\left|\frac{\partial\phi_\varepsilon}{\partial w}\right|\right|_{2n} ||w_2-w_1||_{2n} \geq \left(\frac{1}{\varepsilon} - 2C_2\right)||u_2-u_1||_p^2 .$$

Note that $\frac{\partial\phi_\varepsilon}{\partial w}$ does not depend on $\varepsilon$ and is bounded on E, uniformly on $t \in [0,T]$ and $v \in \mathcal{U}_{ad}$, by some constant $C_3 > 0$, and so

$$||u_2-u_1||_p^2 \leq 2C_3\left(\frac{1}{\varepsilon} - 2C_2\right)^{-1}||w_2-w_1||_{2n} .$$

This implies (2.14).

We now prove (2.15). Fix $t \in [0,T]$ and denote $\phi(w,u) = K(z,u,q,t)$.

Then

$$K_\varepsilon(z,u,v,q,t) = \phi(w,u) + \frac{1}{2\varepsilon}||u-v||_p^2 .$$

Let $v_1, v_2$ in $\mathcal{U}_{ad}$ be given and $u_1$ (resp. $u_2$) be the solution of (2.13) with $v=v_1$ (resp. $v=v_2$). The necessary optimality condition of (2.13) is

$$(\frac{\partial\phi}{\partial u}(w,u_i) + \frac{1}{\varepsilon}(u_i-v_i),\ u-u_i) \geq 0,\ \forall u \in \mathcal{U}_{ad},\ i = 1,2.$$

We fix u to $u_{3-i}$, $i = 1,2$, and sum the inequalities : it follows that

$$\varepsilon(\frac{\partial\phi}{\partial u}(w,u_1) - \frac{\partial\phi}{\partial u}(w,u_2),\ u_2-u_1) + (v_2-v_1,\ u_2-u_1) \geq ||u_2-u_1||_p^2 .$$

Using the Cauchy-Schwartz inequality and dividing by $||u_2-u_1||_p$ we deduce that

$$||u_2-u_1||_p \leq ||v_2-v_1||_p + 2\varepsilon \sup_{\substack{w\in E \\ u\in \mathcal{U}_{ad}}} ||\frac{\partial\phi}{\partial u}(w,u)||.$$

This with (2.7)(2.9) implies (2.15).

We now write the optimality condition of (2.13) in the form

$$(\frac{\partial\phi}{\partial u}(w,u_\varepsilon) + \frac{1}{\varepsilon}(u_\varepsilon-v),\ u-u_\varepsilon) \geq 0,\ \forall u \in \mathcal{U}_{ad} .$$

With $u = v$ this implies

$$||u-v||_p \leq \varepsilon \sup_{\substack{w\in E \\ u\in \mathcal{U}_{ad}}} ||\frac{\partial\phi}{\partial u}(w,u)||_p$$

from which (2.16) is easily deduced. □

*Proof of Theorem 2.1.*

The non trivial point is to prove that the differential equation (2.1) coupled to the minimization problem (2.13) has at least one solution. It is sufficient to prove that given $u^o$ admissible, there exists some $u^1$ for $\varepsilon < \varepsilon_o$.

Let $n > 0$ be given. Define $\Delta t = \frac{T}{n}$ and $u^o_{\Delta t} \in L^\infty(0,T,\mathbb{R}^p)$ by

$$u^o_n(t) = \Delta t \int_{(k-1)\Delta t}^{k\Delta t} u^o(t)\, dt, \ \forall t \in [(k-1)\Delta t, k\Delta t[\ , \ k = 1, \ldots, n.$$

Note that $u^o_n(t) \in \mathcal{U}_{ad}$, $\forall t \in [0,T]$.

Let $y^1_n(t)$, $u^1_n(t)$ be the solution of

$$(2.18) \quad \begin{cases} \frac{d}{dt} y^1_n(t) = f(y^1_n(t), u^1_n(t), t), \\ y^1_n(0) = 0, \end{cases}$$

$$(2.19) \quad \begin{cases} u^1_n(k\Delta t) = u_\varepsilon(y^1_n(k\Delta t), u^o_n(k\Delta t), p^o(k\Delta t), k\Delta t), \ \forall k \in [1, n-1], \\ u^1_n(t) = u^1_n(k\Delta t), \ \forall t \in [(k-1)\Delta t, k\Delta t[\ , \ k \in [1, n-1]. \end{cases}$$

Note that (2.18)(2.19) has a unique solution for $\varepsilon < \varepsilon_o$. We study the limit of $y^1_n$, $u^1_n$ when $n \to \infty$. As $u^1_n(t) \in \mathcal{U}_{ad}$, $\forall t \in [0,T]$, and because of (2.9) we have for some $C_4 > 0$

$$||y^1_n||_Y < C_4.$$

As $\mathcal{U}_{ad}$ is bounded this implies the existence of $y^1$, $u^1 \in Y \times L^\infty(0,T,\mathbb{R}^p)$ such that

$$y_n^1 \to y^1 \ Y \ \text{weak} *,$$

$$u_n^1 \to u^1 \ L^\infty(0,T,\mathbb{R}^p) \ \text{weak} *,$$

$$u^1(t) \in \mathcal{U}_{ad} \quad \text{a.e. } t \in (0,T).$$

From the Ascoli-Arzela theorem we deduce that

$$y_n^1 \to y^1 \quad \mathcal{C}(0,T,\mathbb{R}^n),$$

and so

$$y^1(0) = 0.$$

As $u_n^o \to u^o \quad L^2(0,T,\mathbb{R}^p)$, we have

$$u_n^o\ (t) \to u^o(t) \text{ a.e. } t \in (0,T).$$

As, by lemma 2.1, $u_n^1(t)$ is a continuous function of its arguments if $\varepsilon < \varepsilon^o$, this implies

$$u_n^1(t) \to \hat{u}^1(t) = u_\varepsilon(y^1(t),\ u^o(t),\ p^o(t),\ t), \text{ a.e. } t \in (0,T).$$

The continuity of f implies

$$f(y_n^1(t),\ u_n^1(t),\ t) \to f(y^1(t),\ \hat{u}^1(t),\ t) \ , \text{ a.e. } t \in (0,T).$$

Using Lebesgue's theorem we find that

$$u^1(t) = \hat{u}^1(t); \text{ a.e. } t \in (0,T),$$

and

$$\frac{dy^1(t)}{dt} = f(y^1(t),\ \hat{u}^1(t),\ t), \text{ a.e. } t \in [0,T].$$

Thus $(y^1,u^1)$ check the conditions required by the algorithm. □

*Remark 2.2.*

If f and l are twice continuously differentiable, it can be showed that, if $\varepsilon$ is small enough, $u_\varepsilon$ is a lipschitzian function of its arguments. Then the existence of a solution to

$$\begin{cases} y' = f(y, u_\varepsilon(y, u_0, p), t), \ t \in (0,T), \\ y(0) = 0, \end{cases}$$

is a direct consequence of the general results on ordinary differential equations, and the sequence $\{u^i\}$ generated by the algorithm is unique. □

## III - CONVERGENCE OF THE ALGORITHM

This paragraph studies the asymptotic behaviour of the sequences generated by the algorithm. We first state a result similar to one of [3].

*Theorem 3.1*

Under hypothesis (2.7) to (2.11), there exists $\alpha > 0$ such that any sequence generated by the algorithm checks

$$(3.1) \qquad J(u^k) - J(u^{k-1}) \leq -\left(\frac{2}{\varepsilon_i} - \alpha\right) ||u^k - u^{k-1}||^2_{L^\varepsilon(0,T,\mathbb{R}^p)}. \quad \square$$

*Proof*

We drop the variable t when there is no ambiguity. The decrease of the criterion can be written as

$$J(u^k) - J(u^{k-1}) = \int_0^T [H(y^k, u^k, p^{k-1}, t) - H(y^{k-1}, u^{k-1}, p^{k-1}, t) +$$
$$+ (p^{k-1})^t (f(y^k, u^k, t) - f(y^{k-1}, u^{k-1}, t))]dt + g(y^k(T)) - g(y^{k-1}(T)).$$

Define

$$\delta y^k = y^k - y^{k-1},$$

$$\delta u^k = u^k - u^{k-1}.$$

Then, with (2.1) :

$$(3.2)\quad \begin{cases} J(u^k) - J(u^{k-1}) = \int_0^T [H(y^k,u^k,p^k,t) - H(y^{k-1},u^{k-1},p^{k-1},t) + \\ \qquad + p^{k-1}\frac{d}{dt}(\delta y^k)]dt + g(y^k(T)) - g(y^{k-1}(T)). \end{cases}$$

We obviously have

$$H(y^k,u^k,p^{k-1},t) - H(y^{k-1},u^{k-1},p^{k-1},t) = H(y^k,u^k,p^{k-1},t) -$$

$$- H(y^k,u^k-\delta u^k,p^{k-1},t) + H(y^{k-1}+\delta y^k,u^{k-1},p^{k-1},t) -$$

$$- H(y^{k-1},u^{k-1},p^{k-1},t).$$

It is easy to check that the hypothesis made imply the existence of $C_5 > 0$, such that, $\forall t \in [0,T]$ :

$$H(y^k,u^k-\delta u^k,p^{k-1},t) \geq H(y^k,u^k,p^{k-1},t) - (\frac{\partial H}{\partial u}(y^k,u^k,p^{k-1},t),\delta u^k) -$$

$$- C_5||\delta u^k||_p^2.$$

$$H(y^{k-1}+\delta y^k,u^{k-1},p^{k-1},t) \leq H(y^{k-1},u^{k-1},p^{k-1},t) +$$

$$+ (\frac{\partial H}{\partial y}(y^{k-1},u^{k-1},p^{k-1},t),\delta y^k) + C_5||\delta y^k||_n^2.$$

These inequalities with (3.2) imply

$$(3.3)\quad \begin{cases} J(u^k) - J(u^{k-1}) \leq \int_0^T [(\frac{\partial H}{\partial u}(y^k,u^k,p^{k-1},t),\delta u^k) + C_5||\delta u^k||_n^2 + \\ \qquad + (\frac{\partial H}{\partial y}(y^{k-1},u^{k-1},p^{k-1},t),\delta y^k) + C_5||\delta y^k||_n^2 + p^{k-1}\frac{d}{dt}(\delta y^k)]dt + \\ \qquad + g(y^k(T)) - g(y^{k-1}(T)). \end{cases}$$

We note that

$$(\frac{\partial H}{\partial u}(y^k,u^k,p^{k-1},t),\delta u^k) = (\frac{\partial K_\varepsilon}{\partial u}(y^k,u^k,u^{k-1},p^{k-1},t),\delta u^k) - \frac{1}{\varepsilon}||\delta u^k||^2$$

and, by definition of $u^k$

$$(\frac{\partial K_\varepsilon}{\partial u}(y^k,u^k,u^{k-1},p^{k-1},t),\delta u^k) \le 0, \text{ a.e. } t \in (0,T).$$

On the other hand, using (2.5) and integrating by parts

$$\int_0^T [(\frac{\partial H}{\partial y}(y^{k-1},u^{k-1},p^{k-1},t),\delta y^k) + p^{k-1}\frac{d}{dt}(\delta y^k)]dt =$$

$$= \int_0^T [(\frac{d}{dt}p^{k-1}, \delta y^k) + (p^{k-1}, \frac{d}{dt}(\delta y^k))]dt =$$

$$= [p^{k-1}(t), \delta y^k(t)]_0^T = (\frac{\partial g}{\partial y}(y^{k-1}(T)),\delta y^k(T)).$$

As $y(T)$ is uniformly bounded in $\mathbb{R}^n$ we deduce from (2.11) that for some $C_6 > 0$ :

$$g(y^k(T)) - g(y^{k-1}(T)) + (\frac{\partial g}{\partial y}(y^{k-1}(T)), \delta y^k(T)) \le C_6||\delta y^k(T)||_n^2.$$

Using (3.3) we deduce that

$$J(u^k) - J(u^{k-1}) \le \int_0^T [(C_5 - \frac{1}{\varepsilon})||\delta u^k(t)||_p^2 + C_5||\delta y^k(t)||_n^2]dt + \tag{3.4}$$

$$+ C_6||\delta y^k(T)||_n^2.$$

We end the proof as in [3] : we have

$$\frac{d}{dt}||\delta y^k(t)||_n \le ||\frac{d}{dt}\delta y^k(t)||_n = ||f(y^k,u^k,t) - f(y^{k-1},u^{k-1},t)||_n.$$

Hypothesis (2.7) and the fact that y and u are uniformly bounded imply the existence of $C_7 > 0$ such that

$$\frac{d}{dt}||\delta y^k(t)||_n \le C_7(||\delta y^k(t)||_n + ||\delta u^k(t)||_p).$$

Using Gronwall's inequality we deduce that for some $c_8 > 0$

$$||\delta y^k(t)||_n \le C_8 \int_0^T ||\delta u^k(t)||_p \, dt,$$

and with the Cauchy-Schwartz inequality :

$$||\delta y^k(t)||_n^2 \le C_8 \, T \int_0^T ||\delta u^k(t)||_p^2 \, dt.$$

This implies

$$\int_0^T ||\delta y^k(t)||_n^2 \, dt \le C_8(T)^2 \int_0^T ||\delta u^k(t)||_p^2 \, dt.$$

The two last inequalities with (3.4) yields (3.1). □

We denote by $P_{\mathcal{U}_{ad}}$ the application from $L^2(0,T,\mathbb{R}^p)$ onto itself such that $P_{\mathcal{U}_{ad}}(u)(t)$ is a.e. in t the projection of u(t) on $\mathcal{U}_{ad}$. The gradient of J is denoted $\nabla J$.

We state a result of global convergence :

*Theorem 3.2.*

We suppose that J is lower bounded and that hypothesis (2.7) to (2.11) hold. Then there exists $\varepsilon_0 > 0$ such that, if $\varepsilon_k < \varepsilon_0$, any sequence generated by the algorithm checks :

(3.5) $\quad J(u^k)$ decreases,

(3.6) $\quad ||u^k - u^{k-1}||_{L^2(0,T,\mathbb{R}^p)} \to 0,$

(3.7) $\quad ||u^k - P_{\mathcal{U}_{ad}}(u^k - \varepsilon_k \nabla J(u^k))||_{L^2(0,T,\mathbb{R}^p)} \to 0.$

*Proof*

Relations (3.5) and (3.6) are an easy consequence of theorem 3.1. Let us check (3.7). From the definition of $u^k$ we deduce that

$$(u^k - u^{k-1} + \varepsilon_k \frac{\partial H}{\partial u}(y^k, u^k, p^{k-1}, t),\ v - u^k) \geq 0,\ \forall v \in \mathcal{U}_{ad},\ \text{a.e. } t \in (0,T).$$

This can be written as

$$u^k = P_{\mathcal{U}_{ad}}(u^{k-1} - \varepsilon_k \frac{\partial H}{\partial u}(y^k, u^k, p^{k-1}, t)).$$

We note that

$$\nabla J(u^{k-1}) = \frac{\partial H}{\partial u}(y^{k-1}, u^{k-1}, p^{k-1}, t),$$

so that

$$u^{k-1} - P_{\mathcal{U}_{ad}}(u^{k-1} - \varepsilon_k \nabla J(u^{k-1}) = u^{k-1} - u^k + P_{\mathcal{U}_{ad}}(u^{k-1} -$$
$$- \varepsilon_k \frac{\partial H}{\partial u}(y^k, u^k, p^{k-1}, t)) - P_{\mathcal{U}_{ad}}(u^{k-1} - \varepsilon_k \frac{\partial H}{\partial u}(y^{k-1}, u^{k-1}, p^{k-1}, t)).$$

As $P_{\mathcal{U}_{ad}}$ is a contraction we obtain

$$\|u^k - P_{\mathcal{U}_{ad}}(u^{k-1} - \varepsilon_k \nabla(u^{k-1}))\|_{L^2(0,T,\mathbb{R}^p)} \leq \|u^{k-1} - u^k\|_{L^2(0,T,\mathbb{R}^p)} +$$
$$(3.8) \qquad + \varepsilon \left\|\frac{\partial H}{\partial u}(y^k, u^k, p^{k-1}, t) - \frac{\partial H}{\partial u}(y^{k-1}, u^{k-1}, p^{k-1}, t)\right\|_{L^2(0,T,\mathbb{R}^p)}.$$

From (3.6) we deduce that

$$\begin{cases} u^k(t) - u^{k-1}(t) \to 0 \\ y^k(t) - y^{k-1}(t) \to 0 \\ p^k(t) - p^{k-1}(t) \to 0 \end{cases} \qquad \text{a.e. } t \in (0,T).$$

As H is continuous this implies that

$$\frac{\partial H}{\partial u}(y^k,u^k,p^{k-1},t) - \frac{\partial H}{\partial u}(y^{k-1},p^{k-1},u^{k-1},t) \to 0, \text{ a.e. } t \in [0,T].$$

Using Lebesgue theorem we see that the right side of (3.8) tends towards zero. □

The hypothesis we made do not allow to prove the convergence of $\{u^k\}$ ; in fact they do not even insure the existence of an optimal control. However, stronger results are easily deduced from (3.8) associated to some additional hypothesis : here is an example :

*Theorem 3.3.*

We suppose that $J(u)$ is a l.s.c. convex function, defined on a convex domain, and lower bounded, that hypothesis (2.7) to (2.11) hold, that $\varepsilon$ is small enough so that (3.5)(3.6)(3.7) hold and that $\liminf \varepsilon_k > 0$. Then some subsequence of $\{u^k\}$ (still denoted $\{u^k\}$ ) checks

$$u^k \to \bar{u} \qquad L^2(0,T,\mathbb{R}^p) \text{ weak},$$

where $\bar{u}$ is an optimal control. □

The proof of the theorem is a direct consequence of (3.7) and of the following lemma :

*Lemma 3.1.*

Let U be an Hilbert space and K a closed convex subset of U. Let $J : U \to \mathbb{R}$ be a l.s.c. proper convex functional. Let $\{u^k\}$ be a bounded sequence such that J is Gâteaux differentiable on $\{u^k\}$ and $\{\nabla J(u^k)\}$ is bounded. We assume that there exists $\rho > 0$ such that

$$||u^k - P_K(u^k - \rho\nabla J(u^k))||_U \to 0.$$

Then some subsequence of $\{u^k\}$, still denoted $\{u^i\}$, checks

$$u^k \to \bar{u} \text{ in } U \text{ weak},$$

and

$$(3.8 \quad \begin{cases} \bar{u} \in K, \\ J(\bar{u}) \leq J(u), \ \forall u \in K. \end{cases} \quad \square$$

*Proof of Lemma 3.1.*

Denote $g^k = \nabla J(u^k)$. We have for some subsequence

$$u^k \to \bar{u} \quad U \text{ weak},$$

$$g^k \to \bar{g} \quad U \text{ weak},$$

and $\bar{u} \in K$. Denote

$$v^k = P_K(u^k - \rho g^k).$$

We have

$$(v^k - u^k + \rho g^k), w - v^k) \geq 0, \forall w \in K,$$

i.e.

$$(v^k-u^k,w-v^k) + \rho(g^k,w-u^k) + \rho(g^k,u^k-v^k) \geq 0, \forall v \in K.$$

We made the hypothesis $||v^k-u^k|| \to 0$. Then

$$\liminf_{i\to+\infty} (g^k,w-u^k) \geq 0, \forall w \in K,$$

and consequently

$$(\bar{g},w) \geq \limsup_{i\to+\infty} (g^k,u^k), \forall w \in K.$$

This holds in particular for $w = \bar{u}$. From this, we deduce (see [1] p. 27) that $\bar{g} \in \partial J(\bar{u})$, where $\partial J(\bar{u})$ is the subdifferential of J, and that $(g^k,u^k) \to (\bar{g},\bar{u})$. Consequently

$$(\bar{g},w-\bar{u}) \geq 0, \forall w \in K,$$

and this implies (3.8). □

*Remark 3.1.*

Lemma 3.1. still holds if we only suppose that $g^k$ is an element of the subdifferential of J at $u^k$. □

*Remark 3.2.*

When applied to a discretized version of (2.5) a formula of type (3.7) is sufficient to prove that for every limit point $\bar{u}$ of $\{u^k\}$, the following necessary optimality condition holds :

$$\bar{u} = P_K (\bar{u} - \varepsilon\nabla J(\bar{u})), \forall \varepsilon > 0. \quad \square$$

## IV - RELATIONS WITH THE GRADIENT PLUS PROJECTION METHOD

We have seen in the proof of theorem 3.2. that the algorithm can be written as

$$(4.1) \qquad u^k = P_{\mathcal{U}_{ad}} \left(u^{k-1} - \varepsilon \frac{\partial H}{\partial u} (y^k, u^k, p^{k-1}, t)\right).$$

The gradient plus projection method for a control problem can be written as

$$(4.2) \qquad \hat{u}^k = P_{\mathcal{U}_{ad}} \left(u^{k-1} - \varepsilon \frac{\partial H}{\partial u} (y^{k-1}, u^{k-1}, p^{k-1}, t)\right),$$

where $\varepsilon$ is the step size. We fix $u^{k-1}$ and show that the two points obtained by (4.1) and (4.2) are very near when $\varepsilon$ is small.

*Proposition 4.1.*

Let $u^{k-1} \in L^2(0,T,\mathbb{R}^p)$ be an admissible control. We denote $u^k_\varepsilon$ and $\hat{u}^k_\varepsilon$ the new controls obtained by (4.1) and (4.2). Then, if $\hat{u}^k \neq u^{k-1}$ for some $\varepsilon > 0$ :

$$\lim_{\varepsilon \to 0} \frac{\|u^k_\varepsilon - \hat{u}^k_\varepsilon\|_{L^2(0,T,\mathbb{R}^p)}}{\|\hat{u}^k_\varepsilon - u^{k-1}\|_{L^2(0,T,\mathbb{R}^p)}} \to 0.$$

*Proof*

As $P_{\mathcal{U}_{ad}}$ is a contraction, (4.1), (4.2) imply that (dropping the subscripts $\varepsilon$):

$$\|u^k - \hat{u}^k\|_{L^2(0,T,\mathbb{R}^p)} \leq \varepsilon \left\|\frac{\partial H}{\partial u} (y^k, u^k, p^k, t) - \right.$$

$$\left. - \frac{\partial H}{\partial u} (y^{k-1}, u^{k-1}, p^{k-1}, t)\right\|_{L^2(0,T,\mathbb{R}^p)} .$$

The continuous function $\frac{\partial H}{\partial u}$ is uniformly continuous on the bounded set of admissible states, controls and costates, for $t \in [0,T]$, so that $\forall \alpha > 0$, there exists $\varepsilon_1$ such that

$$(4.3) \qquad \varepsilon < \varepsilon_1 \Rightarrow \|u^k - \hat{u}^k\|_{L^2(0,T,\mathbb{R}^p)} \leq \alpha\varepsilon .$$

The lemma 1 of the annex implies that

$$\frac{1}{\varepsilon}\|\hat{u}_\varepsilon^k - u^{k-1}\|_{L^2(0,T,\mathbb{R}^p)} \geq \frac{1}{\varepsilon'}\|\hat{u}_\varepsilon^k - u^{k-1}\|_{L^2(0,T,\mathbb{R}^p)}, \ \forall\varepsilon' > \varepsilon > 0,$$

i.e.

$$\|\hat{u}^k - u^{k-1}\|_{L^2(0,T,\mathbb{R}^p)} \geq C_1\varepsilon$$

for some $C_1 > 0$. This with (4.3) proves the proposition. □

We now see that in an important particular case, the algorithm is equivalent to a gradient plus projection method in some metric.

*Theorem 4.1.*

We assume that the control enters in a linear way into the state equation and in a quadratic way into the criterion, i.e.

$$f(y,u,t) = f_1(y,t) + B(t)u,$$

$$l(y,u,t) = \frac{1}{2}\, u^t\, N(t)u + l_1(y,t),$$

where $B(t)$ and $N(t)$ are time-dependant matrices of convenient dimension, and $N(t)$ is symmetric. We denote

$$M_\varepsilon(t) = I + \varepsilon\, N(t).$$

Then the algorithm reduces to the gradient plus projection method in $L^2(0,T,\mathbb{R}^p)$ endowed with the new metric

$$(4.4) \qquad ((u,v))_{L^2(0,T,\mathbb{R}^p)} = \int_0^T u(t)^t \, M_\varepsilon(t) \, v(t) \, dt. \quad \square$$

*Proof*

From (4.1) written with i = 1 we see that $u^1$ is characterized by

$$(u^1(t) - u^0(t) + \varepsilon(N(t)\, u^1(t) + B^*(t)\, p^0(t)),\ v-u^1(t))_{\mathbb{R}^p} \geq 0,$$

$$\forall v \in \mathcal{U}_{ad}\ , \quad \text{a.e. } t \in (0,T).$$

This is equivalent to

$$(I + \varepsilon N(t))u^1(t) - (I + \varepsilon N(t))u^0(t) + \varepsilon(N(t)\, u^0(t) + B^*(t)\, p^0(t)),$$

$$v-u^1(t))_{\mathbb{R}^p} \geq 0,\ \forall v \in \mathcal{U}_{ad},\ \text{a.e. } t \in [0,T].$$

We note that $\nabla J(u^0)(t) = N(t)u^0(t) + B^*(t)p^0(t)$, so that

$$(M_\varepsilon(t)(u^1(t) - u^0(t) - \varepsilon\, (M_\varepsilon)^{-1}\nabla J(u^0)(t)),\ v-u^1(t))_{\mathbb{R}^p} \geq 0,$$

$$\forall v \in \mathcal{U}_{ad},\ \text{a.e. } t \in (0,T).$$

This is the characterization of the point obtained by the gradient plus projection method associated to the metric (4.4). $\square$

## V - CONCLUSION

In this study of an algorithm of resolution of optimal control problems, we establish some results of well-posedness and of convergence. In addition, we show that this algorithm is strongly related to the gradient plus projection method. For numerical experiments using this algorithm, see [3], [4], [5].

The following technical lemma, which can be interesting in itself, has been established. Let K be a closed convex subset of an Hilbert space H, $u \in K$, $d \in H$. $P_K$ denotes the projection onto K. Then $\frac{1}{\varepsilon} ||u-P_K(u+\varepsilon d)||$ is a decreasing function of $\varepsilon$.

Acknowledgment

I thank Professor Y. Sakawa, from Osaka University, for an interesting discussion concerning the algorithm studied in this paper.

REFERENCES

[1] H. BREZIS
Operateurs maximaux monotones et semi-groupes de contraction dans les espaces de Hilbert.
North-Holland, Amsterdam, 1973.

[2] E. LEE - L. MARKUS
Foundations of optimal control theory.
J. Wiley, New York, 1967.

[3] Y. SAKAWA - Y. SHINDO
On global convergence of an algorithm for optimal control
IEEE Trans. AC- 25, pp. 1149-1153, 1980.

[4] Y. SAKAWA - Y. SHINDO
Optimal control of container cranes.
Automatica 18, pp. 257-266, 1982.

[5] Y. SAKAWA - Y. SHINDO - Y. HASHIMOTO
Optimal control of a rotary crane.
J.O.T.A. 35, pp. 535-557, 1981

ANNEX

Let $K$ be a closed convex subset of an Hilbert space $H$. Let $u$ be in $K$ and $d \in H$. Call

$$u(\varepsilon) = P_K(u + \varepsilon d)$$

where $\varepsilon \in \mathbb{R}^+$ and $P_K$ is the projection onto $K$. Then

*Lemma 1*

The function

$$\phi(\varepsilon) = \frac{1}{\varepsilon}||u - u(\varepsilon)||$$

is decreasing with respect to $\varepsilon$ . □

*Proof*

If

$$u + \varepsilon d \in K, \ \forall \ \varepsilon \geq 0,$$

the result is trivial. If this is not true, there exists $\bar{\varepsilon} \geq 0$ such that

$$u + \varepsilon d \in K \iff \varepsilon \leq \bar{\varepsilon}.$$

As

$$\phi(\varepsilon) = ||d|| \quad \text{if } \varepsilon \leq \bar{\varepsilon},$$

it is sufficient to prove that

$$\frac{1}{\varepsilon_2}||u - u(\varepsilon_2)|| \leq \frac{1}{\varepsilon_1}||u - u(\varepsilon_1)||, \quad \varepsilon_2 > \varepsilon_1 \geq \bar{\varepsilon}.$$

We call $u_i = u(\varepsilon_i)$, $i = 1, 2$ ; let $\Pi$ be the hyperplane

$$\Pi = \{v \in H \ ; \ (v - u_1, u_1 - (u + \varepsilon_1 d)) = 0\}.$$

Let $P_\Pi$ be the projection onto $\Pi$ and denote

$$\tilde{u}(\varepsilon) = P_\Pi(u + \varepsilon d).$$

Call $\tilde{u}_2 = \tilde{u}(\varepsilon_2)$. We first check

$$(1) \qquad \frac{1}{\varepsilon_2}||u - \tilde{u}_2|| \leq \frac{1}{\varepsilon_1}||u - u_1||.$$

We decompose d in

$$d = d_n + d_t,$$

with $d_n$ (resp. $d_t$) normal (resp. parallel) to $\Pi$. It is not difficult to see that

$$(2) \qquad \tilde{u}(\varepsilon) = u_1 + (\varepsilon - \varepsilon_1)d_t.$$

With $\varepsilon = 0$ this gives :

$$\tilde{u}(0) = u_1 - \varepsilon_1 d_t.$$

As $\tilde{u}(0)$ is the projection of u onto $\Pi$, we get

$$||u - u_1||^2 = ||u - \tilde{u}(0)||^2 + ||\tilde{u}(0) - u_1||^2$$

$$= ||u - \tilde{u}(0)||^2 + \varepsilon_1^2||d_t||^2,$$

and so :

$$(3) \qquad ||u - u_1|| \geq \varepsilon_1||d_t||.$$

On the other hand, we have

$$||u - \tilde{u}_2|| \leq ||u - u_1|| + ||u_1 - \tilde{u}_2||$$

and with (2) :

$$||u - \tilde{u}_2|| \leq ||u - u_1|| + (\varepsilon_2 - \varepsilon_1)||d_t||.$$

We devide by $\varepsilon_2$ and rearrange it as

$$\frac{1}{\varepsilon}||u - \tilde{u}_2|| \leq \frac{1}{\varepsilon_1}||u - u_1|| + (\frac{1}{\varepsilon_2} - \frac{1}{\varepsilon_1})(||u - u_1|| - \varepsilon_1||d_t||).$$

This with (3) implies (1).

The lemma will be proved if the following inequality holds :

(4) $$||u - u_2|| \leq ||u - \tilde{u}_2||.$$

We have

$$||u - u_2||^2 = ||u - u_1||^2 + 2(u - u_1, u_1 - u_2) + ||u_1 - u_2||^2,$$

$$||u - \tilde{u}_2||^2 = ||u - u_1||^2 + 2(u - u_1, u_1 - \tilde{u}_2) + ||u_1 - \tilde{u}_2||^2.$$

We will prove that

(5) $$||u_1 - u_2|| \leq ||u_1 - \tilde{u}_2||$$

and then

(6) $$(u - u_1, u_1 - u_2) \leq (u - u_1, u_1 - \tilde{u}_2).$$

This implies (4). Let us first prove (5). We may suppose that $u_1 \neq u_2$. Let D be the straight line

$$D = \{v \in H \; ; \; v = u_1 + t(u_2 - u_1), \; t \in \mathbb{R}\}.$$

Let w be the projection of $u + \varepsilon_2 d$ onto D. Then

$$w = u_1 + \alpha(u_2 - u_1)$$

for some $\alpha > 0$, and by definition of w :

$$(u_2 - u_1, \; u_1 + \alpha(u_2 - u_1) - (u + \varepsilon_2 d) = 0,$$

or

$$(u_2 - u_1, \; u_1 - (u + \varepsilon_2 d)) = - \alpha ||u_2 - u_1||^2,$$

i.e.

$$(7) \qquad \alpha = \frac{(u_2 - u_1, \; u + \varepsilon_2 d - u_1)}{||u_2 - u_1||^2} \quad .$$

As $u + \varepsilon_1 d$ is outside K, $u + \varepsilon_2 d$ is also outside K. The hyperplan $\Pi$ separates K from $u + \varepsilon_1 d$, and a-fortiori K from $u + \varepsilon_2 d$. As $\tilde{u}_2$ is the projection of $u + \varepsilon_2 d$ on $\Pi$, we have

$$(u_2 - \tilde{u}_2, \; \tilde{u}_2 - (u + \varepsilon_2 d)) \geq 0.$$

As $u_1$ and $\tilde{u}_2$ belong to $\Pi$ :

$$(u_1 - \tilde{u}_2, \; \tilde{u}_2 - (u + \varepsilon_2 d)) = 0,$$

and so, by soustraction

$$(u_2 - u_1, \; \tilde{u}_2 - (u + \varepsilon_2 d)) \geq 0,$$

or

$$(u_2 - u_1, \; \tilde{u}_2 - u_1) + (u_2 - u_1, \; u_1 - (u + \varepsilon_2 d)) \geq 0.$$

With (7) we deduce that

$$(u_2 - u_1, \tilde{u}_2 - u_1) \geq \alpha ||u_2 - u_1||^2 ,$$

and so

$$(8) \qquad ||\tilde{u}_2 - u_1|| \geq \alpha ||u_2 - u_1||.$$

Now, $u_2$ is the projection of $u + \varepsilon_2 d$ on K, and so a-fortiori on $[u_1, u_2]$. As $w$ is the projection of $u + \varepsilon_2 d$ on D, $\alpha \geq 1$. This with (8) implies (5).

Let us prove (6). Let $\hat{u}$ be the projection of $u_2$ onto $\Pi$. We have

$$(9) \qquad \begin{cases} (u - u_1, u_1 - \tilde{u}_2) = (\tilde{u}(0) - u_1, u_1 - \tilde{u}_2) \\ \qquad\qquad\qquad\qquad = ||\tilde{u}(0) - u_1|| \; ||u_1 - \tilde{u}_2||, \end{cases}$$

and

$$(u - u_1, u_1 - u_2) = (u - \tilde{u}(0) + \tilde{u}(0) - u_1, u_1 - \hat{u} + \hat{u} - u_2)$$

$$= (u - \tilde{u}(0), \hat{u} - u_2) + (\tilde{u}(0) - u_1, u_1 - \hat{u}).$$

As $u$ and $u_2$ are on the same half-space limited by $\Pi$, we have

$$(u - \tilde{u}(0), \hat{u} - u_2) \leq 0,$$

so that

$$(u - u_1, u_1 - u_2) \leq ||\tilde{u}(0) - u_1|| \; ||u_1 - \hat{u}||.$$

As the projection onto $\Pi$ is a contraction :

$$(u - u_1, u_1 - u_2) \leq ||\tilde{u}(0) - u_1|| \; ||u_1 - u_2||.$$

This with (9) and (5) implies (6), and proves the lemma. □

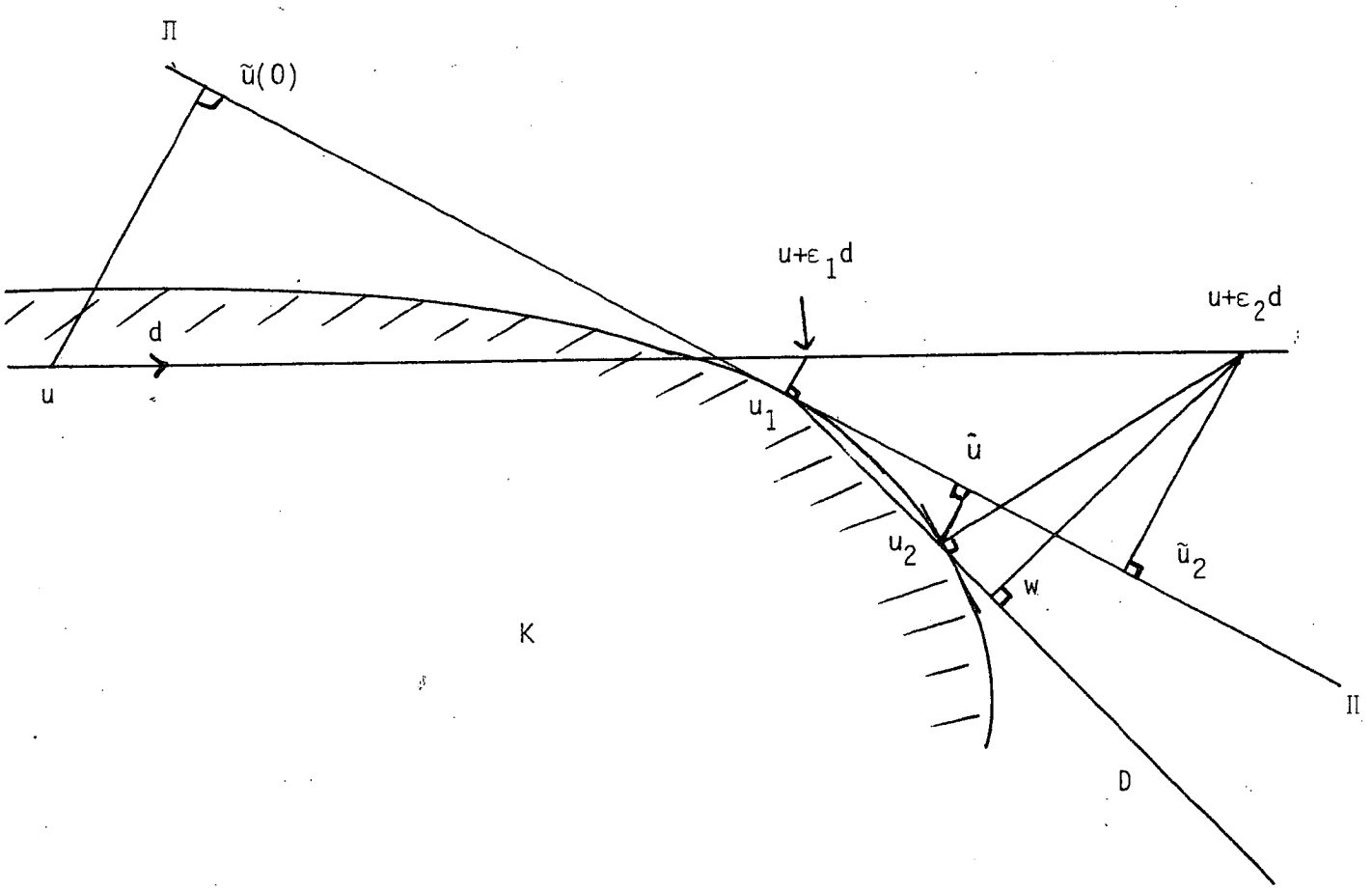


FIGURE 1

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique