# Optimal Control and Geometric Algebra.

## Introduction and some Mathematics.

Students in economics and microeconomics are not mathematicians, with a few exceptions. Neither are engineers. That simple truth is forgotten by most of the authors dealing with optimization and specially with optimal control. They try to be rigourous, complete, maybe brilliant, and so are discouraging for « ordinary people » whose ambition is not to understand the finest details of mathematical demonstrations, or to know the most intricate particular cases where the theorem is not valid.

I am myself not sure at all to have mastered in depth all the subtleties of the theory, but I nevertheless used it, I think with some success. I could do that because I found in some books, or constructed myself rough demonstrations and economical interpretations that allowed me to safely use the theory and to understand what I was doing.

I will try here to transmit some of that knowledge.

Somewhat in contradiction with what I told above, I will try to use coordinate free mathematics, both old (Clifford) and new, that is « Geometric Algebra » and « Geometric Calculus » as developed by David Hestenes and his followers (see Internet). It seems possible to me « to hunt two rabbits at the same time » as we say in French, because I will use a light version of GA and GC, involving only the well known inner product between vectors, that is $(a.b)$ . The reader will use a new language, almost without noticing it ...

First we must define the directional derivative for scalar or vectorial valued functions of vectors. If $f(x)$ is a scalar valued function and $a$ some vector, we define :

(1) $$a.\nabla_x f(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}[f(x+\varepsilon a) - f(x)]$$

The expression $a.\nabla_x$ is the interior product of two vectors and thus a scalar differential operator which in tensorial notation would be written $a^i \frac{\partial}{\partial x^i}$ . But we will not need coordinates.

In the same way we can define the directional derivative of a vector $u(x)$ , possibly located in another vector space than $x$ :

(2) $$a.\nabla_x u(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}[u(x+\varepsilon a) - u(x)]$$

On both sides of (2) we get vectors. Try to convince yourself by some simple examples that such a definition is much clearer than defining in the traditional way $\nabla u$ as a matrix and multiplying it (from the left ?) by $a$ .

We must also keep in memory that, by definition in GC, as a differential operator, $\nabla_x$ acts only on the scalar or vector immediately at its right, unless other indications are given. Thus we could not write carelessly $f(x)a.\nabla_x$ without indicating by some superscript that $\nabla$ acts on $f(x)$ on its left.

In the second place, we need reciprocal vectors of a set $e^j$ of independent vectors, which determine some vector space $\mathcal{A}^m$ . Calling them $e_i$ , we define them by the relations :

(3) $$e_i . e^j = \delta_i^j$$

It is not a difficult task with standard linear algebra to show that they exist in number $m$, are linearly independent and located in $\mathcal{A}^m$ (supposed to be a metric space). But with GA that is much easier[1] done.

1. Interested readers will find the necessary explanations in the Cambridge GA course on Internet.

Finally we must discuss briefly a subject which is strangely omitted in many more or less old books on mathematical economics. I mean that economic models are normally defined in affine spaces, not in metric ones. A lot of books dont mention that and produce drawings where the direct space and the dual space are without explanation superposed, and where a metric is used in some demonstrations. Why does that work ? Because without mentioning it, or perhaps without knowing it .... , the authors establish an isomorphism between the affine economic space and an euclidean representation space.

There is no problem as long as you dont have the strange idea, - for an economist - , to do some change of basis vectors. For example, as I did, you could try tensorial methods and find a nice proof of the Kuhn and Tucker conditions, discovering that they simply express a change of curvilinear coordinates ! In such a proof you must work without metrics and use the dual space.

But GA gives us an even simpler proof. So we will stay with the above mentioned isomorphism, but knowingly.

## Non-linear Optimization.

We consider the following well-known constrained optimization problem :

(4) $$\begin{cases} \text{Max } f(x) \\ g^k(x) \geqslant 0 \\ k \in (1, 2, \ldots\ldots, p) \qquad x \in E^n \end{cases}$$

We suppose the $f$ and $g$ functions continuous, with first and second continuous derivatives. Let $x_0$ be a local maximum, where $g^s(x)$ are $m$ saturated constraints. We name $\mathcal{D}^s$ the admissible domain defined by the $g^s$ , and $\mathcal{A}^m$ the m-dimensional vector space defined by the gradients $\nabla_x g^s(x_0)$ . $\mathcal{A}^m$ is a subspace of $E^n$ . Its orthogonal complement is written $\mathcal{C}\mathcal{A}^m$ . Of course we suppose $x_0$ to be a « normal » point, that is with linearly independent constraint gradients.

Let $\Omega$ be a neighbourhood of $x_0$ in $\mathcal{D}^s$ . It is easy to see that we can define admissible vectors $a \in E^n$, associated with some $\varepsilon_0$, such that :

(5) $$\begin{cases} g^s(x_0) = 0 \\ x = (x_0 + \varepsilon a) \in \Omega \subset \mathcal{D}^s \qquad a \in E^n \qquad \varepsilon_0 \geqslant \varepsilon \geqslant 0 \\ \delta g^s(x_0) = g^s(x_0 + \varepsilon a) = \varepsilon a.\nabla_x g^s(x_0) + \varepsilon\, o(\varepsilon) \geqslant 0 \end{cases}$$

Thus with $\varepsilon$ small enough we see that (5) implies :

(6) $$a.\nabla_x g^s(x_0) \geqslant 0$$

The existence of admissible vectors is obviously related to the fact that the gradients of the saturated constraints are linearly independent. Then every $a$ satisfying (6) restricted to inegality is an admissible direction. The $a$'s can be split in $(a_{||} + a_{\perp})$ with $a_{||} \in \mathcal{A}^m$ and $a_{\perp} \in \mathcal{C}\mathcal{A}^m$ . Thus only the parallel component is constrained.

Considering now $f(x)$ we can write :

(7) $$\delta f(x_0) = f(x_0 + \varepsilon a) - f(x_0) = \varepsilon a.\nabla f(x_0) + \varepsilon\, o(\varepsilon)$$

As $x_0$ should be a local optimum, **we must have** :

(8) $$\delta f(x_0) \leqslant 0$$

thus :

(9) $$a.\nabla f(x_0) \leqslant 0$$

Let us consider now the reciprocal vectors in $\mathcal{A}^m$ . They are defined by :

(10) $$a_i.\nabla_x g^s(x_0) = \delta_i^s \qquad \forall s \in (1,2,\ldots.,m)$$

Those relations mean that $a_i$ is tangent to all saturated constraint surfaces, the $i$th excepted. Note that $g^i(x_0+\varepsilon\, a_i) = \varepsilon + \varepsilon o(\varepsilon)$ .

The $a_i$'s may not be admissible vectors (it depends on the concavity/convexity of the $g^s$ 's, as you may see with simple two-dimensional examples), but they can always be defined as limits of sequences of admissible $a$'s .

We can also split $\nabla_x f(x_0)$ between $\mathcal{A}^m$ and $\mathcal{CA}^m$ :

(11) $$\nabla_x f(x_0) = \nabla_x f(x_0)_{\|} + \nabla_x f(x_0)_{\perp} = -\lambda_s \nabla_x g^s(x_0) + \nabla_x f(x_0)_{\perp}$$

Multiplying by an admissible $a$ , we get :

(12) $$a.\nabla_x f(x_0) = -\lambda_s\, a.\nabla_x g^s(x_0) + a_{\perp}.\nabla_x f(x_0)_{\perp}$$

As $a_{\perp}$ is free, we must have :

(13) $$\nabla_x f(x_0)_{\perp} = 0 \qquad \nabla_x f(x_0) = \nabla_x f(x_0)_{\|} \in \mathcal{A}_m$$

Then we get :

(14) $$a.\nabla_x f(x_0) = -\lambda_s\, a.\nabla_x g^s(x_0)$$

Considering a sequence of $a$'s whose limit is a reciprocal vector $a_i$ , we can write :

(15) $$-\lambda_s\, a_i.\nabla_x g^s(x_0) = -\lambda_s\, \delta_i^s = -\lambda_i = a_i\,.\nabla_x f(x_0) = \lim_{a \to a_i} a.\nabla_x f(x_0) \leqslant 0$$

Thus all the $\lambda$'s are positive or zero.

That limiting process is the one we find in more traditional presentations where an admissible trajectory is defined in $\Omega$, possibly tangent to the tangent cone in $x_0$ (a notion which we did not use in our proof).

Of course we demonstrate also the well-known economical interpretation of the $\lambda$'s, as we can write, using (15) and (7) :

(16) $$\delta f(x_0) = f(x_0+\varepsilon a_s) - f(x_0) \simeq -\varepsilon\,\lambda_s \simeq -\lambda_s\, g^s(x_0+\varepsilon\, a_s) = -\lambda_s\, \delta g^s(x_0)$$

Thus we have proved **the necessary Kuhn-Tucker conditions**. We only have to incorporate the non-saturated $g^{\bar{s}}$ constraints, with $\lambda_{\bar{s}} = 0$ , to get :

(17) $$\left\{ \begin{array}{ll} \nabla_x f(x_0) + \lambda_k \nabla_x g^k(x_0) = 0 & \\ \qquad\qquad \lambda_k\, g^k(x_0) = 0 & \text{without summation} \\ g^k(x_0) \geqslant 0 \qquad \lambda_k \geqslant 0 & \end{array} \right.$$

We see that these relations have affine character, as expected. So the metrics we employed play no role in the final results.

The restriction of our demonstration to « normal » points may seem too severe, as the K-T theorem has a more general validity. We can easily expand it to a case where more than $n$ constraints are saturated. Suppose the $g^s$ are in number $n$ , and we define everything as above. Let us then add another saturated constraint $g^{n+1}$. We can write :

$$\nabla_x f(x_0) + \lambda_s \nabla_x g^s(x_0) = 0$$

$$\alpha \nabla_x\, g^{n+1}(x_0) + \alpha\mu_s \nabla_x g^s(x_0) = 0 \qquad \alpha > 0$$

Adding these two relations we get :

$$\nabla_x f(x_0) + (\lambda_s + \alpha\,\mu_s) \nabla_x g^s(x_0) + \alpha \nabla_x\, g^{n+1}(x_0) = 0$$

We can choose $\alpha$ small enough to have all coefficients positive. So step by step we can extend the theorem to an arbitrary number of constraints. Of course we must verify that there remains an admissible neighbourhood for $x_0$ . It is very easy to do that by using the reciprocate vectors. For example when adding $g^{n+1} \geqslant 0$ we find that there still exist admissible vectors if and only if there is at least one negative $\mu_s$ .

We will now establish **sufficient conditions**.

It is enough to write down the lagrangian function and to develop it :

(18) $$\mathcal{L}(x) = f(x) + \lambda_k g^k(x) = f(x) + \lambda_s g^s(x)$$

For the local maximum $x_0$ we have :

(19) $$\mathcal{L}(x_0) = f(x_0) \qquad a.\nabla_x \mathcal{L}(x_0) = 0 \qquad g^s(x_0) = 0$$

Then we get :

(20) $$\mathcal{L}(x_0 + \varepsilon a) - \mathcal{L}(x_0) = f(x_0 + \varepsilon a) - f(x_0) + \lambda_s g^s(x_0 + \varepsilon a) = \frac{1}{2}\varepsilon^2 (a.\nabla_x)^2 \mathcal{L}(x_0) + \varepsilon^2 o(\varepsilon)$$

which we can write[2] :

(21) $$f(x_0 + \varepsilon a) - f(x_0) = -\lambda_s g^s(x_0 + \varepsilon a) + \frac{1}{2}\varepsilon^2 (a.\nabla_x)^2 \mathcal{L}(x_0) + \varepsilon^2 o(\varepsilon)$$

where $a$ is of course an admissible vector, and $(x_0 + \varepsilon a) \in \Omega$ .

Let us call $\lambda_\sigma$ the $\lambda_s$ which are strictly positive[3]. Then we get :

(22) $$f(x_0 + \varepsilon a) - f(x_0) = -\lambda_\sigma g^\sigma(x_0 + \varepsilon a) + \frac{1}{2}\varepsilon^2 (a.\nabla_x)^2 \mathcal{L}(x_0) + \varepsilon^2 o(\varepsilon)$$

As $-\lambda_\sigma g^\sigma(x_0 + \varepsilon a)$ is a first order $\varepsilon$ , strictly negative quantity[4], except when $(x_0 + \varepsilon a)$ is a boundary point where **all** $g^\sigma(x_0 + \varepsilon a) = 0$ , we see immediately that the sufficiency problem is related to those boundary-points, and that **a sufficient condition is** :

(23) $$(a.\nabla_x)^2 \mathcal{L}(x_0) < 0$$

To resume, we have established necessary and sufficient conditions for the optimization problem as formulated in (4) for « good » functions, with a minimum of mathematical preliminaries. Once that work done and understood in depth the study of more general or particular cases should be easier for the student.

2. I remember one of my math teachers who sternly warned us against writing formulas like (21), mixing developed and non-developed terms. Here that breaking of the rules seems efficient ....

3. It is interesting to note that $\lambda_s = 0$ means that the associated reciprocal vector $a_s$ is perpendicular to $\nabla_x f(x_0)$, that is tangent to the $f(x)$ surface. (In the tensorial method $a_s$ appears to be the basis vector associated to the new curvilinear coordinate line with index $s$ ).

4. We may note here the particular case where $\boldsymbol{n}$ constraints are saturated at $x_0$ . Then the first term remains strictly negative and the K-T conditions are both necessary and sufficient.

**Optimal Control .... , almost painless.**

That subtitle looks like a joke. But I am serious. What I mean is that a lot of people who step back or at least hesitate to invest in the difficult mathematical reasoning involved by optimal control techniques, are in fact able to use them. But they lack an « Ariadne's thread » which could help them to procede methodically in practical problems to be solved and to stay in control of what they are doing. In my experience, economic interpretation of the principal relations could be that guideline and even furnish some theoretical demonstrations.

As before we will try to work coordinate-free with GA and GC.

Let us consider first a basic problem :

$$(24)\qquad \begin{cases} \text{Max} \int_0^T \mathcal{L}(x,u,t)dt & & & \text{associated multipliers} \\ \dot{x} = f(x,u,t) & x \in E^n & u \in U^m & \psi \text{ (vector)} \\ g^k(x,u,t) \geqslant 0 & \text{for} & k \in (1,2,....,p) & \lambda_k \\ x(0) = x_0 & & & \end{cases}$$

$x$ is the state-variable vector, $u$ the control-variable vector. We write $\bar{x}$ , $\bar{u}$ their optimal values at time $t$. In a first case we suppose that every constraint function contains explicitly some control variable.

Let us form an hamiltonian function $\mathcal{H}$ :

$$(25)\qquad \mathcal{H}(x,u,t) = \mathcal{L}(x,u,t) + \psi(t).\, f(x,u,t)$$

The detailed theory tells us that for an optimal arc we must write :

$$(26)\qquad \begin{cases} -\,\dot{\psi} = \nabla_x \bar{\mathcal{H}} + \lambda_k \nabla_x \bar{g}^k \\ \quad 0 = \nabla_u \bar{\mathcal{H}} + \lambda_k \nabla_u \bar{g}^k \\ \lambda_k(t)\, g^k(\bar{x},\bar{u},t) = 0 \qquad \text{without summation} \\ \lambda_k(t) \geqslant 0 \end{cases}$$

Theory tells us also important information on the continuity and derivability of the multipliers (see annex), which we will not study here, as we limit us to a rough outline. Often the most difficult problem is how to put together the optimal arcs of a complete trajectory.

How can we explain those apparently mysterious relations ? What do they mean ?

Let us start with the famous Bellman Principle, which tells us that any part of an optimal trajectory must be optimal. Thus if we call $\mathcal{J}(\bar{x}(t),t)$ the economic function which we have to optimize between $t$ and $T$ , we write :

$$(27)\qquad \mathcal{J}[\bar{x}(t),t] = \text{Max}_{u,\tau} \int_t^T \mathcal{L}(x,u,\tau)\, d\tau = \text{Max}_{u,\tau} \left[ \int_t^{t+dt} \mathcal{L}\, d\tau + \int_{t+dt}^T \mathcal{L}\, d\tau \right]$$

$$= \text{Max}_{u,\tau} \left[ \int_t^{t+dt} \mathcal{L}\, d\tau \right] + \mathcal{J}[\bar{x}(t+dt), t+dt]$$

If we develop $\bar{\mathcal{J}}$ we get :

$$(28)\qquad \mathcal{J}[\bar{x}(t+dt), t+dt] \simeq \mathcal{J}[\bar{x}(t),t] + (\nabla_x \bar{\mathcal{J}}).\, \dot{x}\, dt + \frac{\partial \bar{\mathcal{J}}}{\partial t}\, dt$$

On the optimal trajectory the vector $\nabla_x \bar{\mathcal{J}}$ is, through $\bar{x}(t)$ and $t$ , a function of $t$ which we call $\psi(t)$.

From (24), (27), (28), and dividing by $dt$ , we get :

$$(29)\qquad \text{Max}_{u(t)}[\mathcal{L}(\bar{x},u,t) + \psi(t).\, f(\bar{x},u,t)] = \text{Max}_{u(t)} \mathcal{H}(\bar{x},u,t) = -\frac{\partial \bar{\mathcal{J}}}{\partial t}$$

Thus we have to do a static optimization, for a given $\bar{x}$ at the instant $t$ , under constraints on the control variables. We obviously have to write the necessary K-T conditions :

$$(30)\qquad 0 = \nabla_u \bar{\mathcal{H}} + \lambda_k \nabla_u \bar{g}^k \qquad \lambda_k \bar{g}^k = 0 \qquad \lambda_k \geqslant 0$$

Before we go further it might be interesting to give some **economic interpretation** of these relations. We may look at the expression

(31) $$\boldsymbol{\mathcal{H}(\bar{x},\bar{u},t)=\mathcal{L}(\bar{x},\bar{u},t)+\psi(t).\dot{\bar{x}}=-\frac{\partial\bar{\mathcal{J}}}{\partial t}}$$

as the current cash flow of an industrial activity between $t$ and $t+dt$ , where $x$ represents the stock of products, $u$ the production factors, $\dot{\bar{x}}$ the physical variation of stocks, $\psi(t)$ their unit price in the period. With relations (30) the manager respects the constraints and optimizes the result, **price variation excluded**, of the period.

The significance of equation (31) is clear ; the direct contribution to the global result (27) by the elementary period $(t,t+dt)$ is equal to the diminution of $\bar{\mathcal{J}}(t)$ during that time. Isn't it remarkable that such sophisticated mathematics lead to such an obvious result !

Let us take now total derivatives relative to $x(t)$ (dont forget that $t$ is here not the time variable, but the initial value of it). We first must note that of course a variation $\delta x$ of the initial product stock has an incidence $\delta u$ on the control variable between $t$ and $(t+dt)$ :

(32) $$\delta\bar{u}=\delta\bar{x}\,.\nabla_x\,u(x_0)$$

As the saturated constraints must be respected we must write :

(33) $$\begin{aligned}\delta g^s(x_0)=0&=\delta\bar{x}\,.\nabla_x\,g^s(x_0)+\delta\bar{u}\,.\nabla_u\,g^s(x_0)\\&=\delta\bar{x}\,.\nabla_x\,g^s(x_0)+\delta\bar{x}\,.\nabla_x\bar{u}\,.\nabla_u\,g^s(x_0)\\&=\delta\bar{x}\,.[\nabla_x\,g^s(x_0)+\nabla_x\bar{u}\,.\nabla_u\,g^s(x_0)]\end{aligned}$$

As $\delta\bar{x}$ is a free vectorial variable in $E^n$ , (33) implies :

(34) $$\nabla_x\,g^s(x_0)+\nabla_x\bar{u}\,.\nabla_u\,g^s(x_0)=0 \qquad \forall s$$

And for a non saturated $g^{\bar{s}}$ we write of course $\lambda_{\bar{s}}=0$ .

Now by deriving (31) and with (30) , (32) , (34) , we find :

(35) $$\begin{aligned}&\delta\bar{x}\,.\nabla_x\bar{\mathcal{H}}+\delta\bar{u}\,.\nabla_u\bar{\mathcal{H}}=-\,\delta\bar{x}\,.\nabla_x(\frac{\partial\bar{\mathcal{J}}}{\partial t})=-\,\frac{\partial}{\partial t}(\delta\bar{x}\,.\nabla_x\bar{\mathcal{J}})=-\,\delta\bar{x}\,.\dot{\psi}\\&\delta\bar{x}\,.\nabla_x\bar{\mathcal{H}}-\lambda_s\,\delta\bar{u}\,.\nabla_u\,\bar{g}^s=-\,\delta\bar{x}\,.\dot{\psi}\\&\delta\bar{x}\,.\nabla_x\bar{\mathcal{H}}-\lambda_s\,\delta\bar{x}\,.\nabla_x\,\bar{u}\,.\nabla_u\,\bar{g}^s=-\,\delta\bar{x}\,.\dot{\psi}\\&\delta\bar{x}\,.\nabla_x\bar{\mathcal{H}}+\lambda_s\,\delta\bar{x}\,.\nabla_x\,\bar{g}^s=-\,\delta\bar{x}\,.\dot{\psi}\end{aligned}$$

And finally we have :

(36) $$\nabla_x\bar{\mathcal{H}}+\lambda_s\nabla_x\,\bar{g}^s=-\,\dot{\psi}$$

which is the so-called coordination equation (between successive elementary periods).

If we now multiply (30) by $(\delta\bar{u}.)$ , (36) by $(\delta\bar{x}.)$ , and add the two expressions, we get :

(37) $$\delta\bar{\mathcal{H}}+\lambda_s\,\delta\bar{g}^s=-\,\delta\bar{x}.\dot{\psi}$$

Let us define :

(38) $$\boldsymbol{\mathcal{V}(\bar{x},\bar{u},t)=\mathcal{L}(\bar{x},\bar{u},t)+\psi(t).\dot{\bar{x}}+\dot{\psi}(t).\bar{x}=\mathcal{H}(\bar{x},\bar{u},t)+\dot{\psi}(t).\bar{x}}$$

Here we have the complete flow of financial result of our industrial activity, between $t$ and $t+dt$ , **with price variations included**. That is what the accountants would calculate, with the important exception that our unit stock price vector has a precise economic justification (the future oriented value of the products we have in stock at time $t$ ).

Now we can write (37) as :

(39) $$\delta\bar{\mathcal{V}}=\delta\bar{\mathcal{H}}+\dot{\psi}.\delta\bar{x}=-\,\lambda_s\,\delta\bar{g}^s$$

which gives us an economical interpretation of the $\lambda_s$ (here we suppose that we replace the constraints $g^s \geqslant 0$ by $g^s \geqslant \delta\bar{g}^s$ , where the $\delta\bar{g}^s$ are given small scalars).

It occurs very often that some or all constraints depend only on state variables. This is a source of additional complications. Let us take a look at such a problem :

$$(40) \qquad \begin{cases} \text{Max} \int_0^T \mathcal{L}(x,u,t)dt & & & \text{associated multipliers} \\ \dot{x} = f(x,u,t) & x \in E^n & u \in U^m & \psi \text{ (vector)} \\ g^k(x,t) \geqslant 0 & \text{for} & k \in (1,2,....,p) & \lambda_k \\ x(0) = x_0 \end{cases}$$

Then we must also consider the derived functions :

$$(41) \qquad \frac{\mathrm{d}g^k}{\mathrm{dt}} = h^k(x,u,t) = f(x,u,t).\nabla_x g^k(x,t) + \frac{\partial g^k}{\partial t}$$

The hamiltonian remains :

$$(42) \qquad \mathcal{H}(x,u,t) = \mathcal{L}(x,u,t) + \psi(t).f(x,u,t)$$

With the already defined (27) function $\mathcal{J}[\bar{x}(t),t]$ we find again :

$$(43) \qquad \text{Max}_{u,\tau}\left[\int_t^{t+dt} \mathcal{L}d\tau\right] + (\nabla_x\bar{\mathcal{J}}).\,\dot{\bar{x}}\,dt = -\frac{\partial\bar{\mathcal{J}}}{\partial t}dt$$

But here we define :

$$(44) \qquad \nabla_x\bar{\mathcal{J}} = \psi + \lambda_s\nabla_x\bar{g}^s$$

Thus we get :

$$(45) \qquad \text{Max}_{u(t)}[(\mathcal{L} + \psi.\dot{x}) + \lambda_s h^s] = -\frac{\partial\bar{\mathcal{J}}}{\partial t} + \lambda_s\frac{\partial\bar{g}^s}{\partial t}$$

$$\text{Max}_{u(t)}[\mathcal{H} + \lambda_s h^s] = \bar{\mathcal{H}} + \lambda_s\bar{h}^s = -\frac{\partial\bar{\mathcal{J}}}{\partial t} + \lambda_s\frac{\partial\bar{g}^s}{\partial t}$$

The static optimization gives :

$$(46) \qquad \nabla_u\bar{\mathcal{H}} + \lambda_k\nabla_u\bar{h}^k = 0 \qquad \lambda_k\bar{g}^k = 0 \qquad \lambda_{\bar{s}} = 0 \text{ (see annex)}$$

If we now differentiate (45) with respect to $x$ and consider the derived constraints, we get :

$$(47) \qquad \nabla_x\bar{\mathcal{H}} + \lambda_s\nabla_x\bar{h}^s = -\nabla_x\left(\frac{\partial\bar{\mathcal{J}}}{\partial t} - \lambda_s\frac{\partial\bar{g}^s}{\partial t}\right) = -\frac{\partial}{\partial t}(\nabla_x\bar{\mathcal{J}} - \lambda_s\nabla_x\bar{g}^s) = -\dot{\psi}$$

There are other formulations which do not explicitly use derived constraints , but then we lose continuity of the multipliers when passing from one arc to another (jump conditions).

Of course there is a lot to learn before being able to solve practical cases, but the goal is not out of reach. I hope having shown that coordinate-free mathematical techniques facilitate it. GA offers more flexibility than standard matricial calculus.

G.Ringeisen

November 2005

**Annex.**

For complete information I resume here the properties of multipliers, without demonstrations.

For the basic problem (24) we have :

– The $\lambda$ are piecewise continuous between 0 and $T$ , and continuous at each point of continuity of $\bar{u}(t)$. $\lambda_s \geqslant 0$ and $\lambda_{\bar{s}} = 0$ if $\bar{g}^{\bar{s}} > 0$ .

– The $\psi(t)$ are continuous and have piecewise continuous derivatives.

– The $\mathcal{H}$ function completed with the constraints is continuous, and on each continuity interval of $\bar{u}$ we have $\dfrac{d\mathcal{H}}{dt} = \dfrac{\partial\mathcal{H}}{\partial t}$ .

For the problem (40) we have :

– The $\lambda$ are piecewise continuous, non increasing between 0 and $T$, constant $(\lambda_s)$ on each interval where $\bar{g}^s > 0$ .

– They are continuous when $\bar{u}$ is continuous, and $(\lambda_k)$ on each point of discontinuity of $\bar{h}^k$ .

– Given $p$ constants $c_k$ , we find the optimum equations and the marginal prices $\nabla_x \bar{\mathcal{J}}$ are not modified by the following transformation

$$\psi \rightarrow \psi + c_k \nabla_x \bar{g}^k \qquad \lambda_k \rightarrow \lambda_k - c_k$$

Thus we can fix $\lambda_k = 0$ at some point of the trajectory (and perhaps some interval).

**Bibliography.**

Calculus of Variations and Optimal Control Theory - Magnus R. Hestenes - Wiley.

Clifford Algebra to Geometric Calculus - David Hestenes , Garret Sobczyk - Reidel.

Geometric Algebra for Physicists - Chris Doran , Anthony Lasenby - Cambridge

David Hestenes and Cambridge on Internet