The Euler Lagrange equation and the Pontriagin Maximum Principle

Arrigo Cellina

Dipartimento di Matematica e Applicazioni

Università di Milano Bicocca

3rd December 2004

1 Introduction

We are interested in minimizing a functional of the type

$$\int_a^b L(t, x(t), x'(t))\, dt$$

or

$$\int_\Omega L(x, u(x), \nabla u(x))\, dx$$

under suitable boundary conditions.
More precisely, assuming that the minimum problem admits a solution, $\tilde{x}(\cdot)$ or $\tilde{u}(\cdot)$,our goal is to discuss appropriate necessary conditions. a basic principle of analysis is that, given a minimum point ξ belonging to the interior of the domain of a differentiable function $F(\cdot)$, we obtain necessary condition exploring a neighborhood of ξ, and we obtain the condition $\langle \nabla F(\xi), \delta\rangle = 0$, yielding

$$\nabla F(\xi) = 0$$

In the same order of ideas, one considers an admissible variation, i.e., a smooth function $\eta(\cdot)$, equal to zero at the boundary, multiplies this functions by a scalar ε and considers the function $x + \varepsilon\eta$. In principle, by deriving with respect to the parameter ε and passing to the limit under the integral sign (this is the difficult step), one obtains the Euler Lagrange equations (E-L):

$$\int_a^b \langle \nabla_{x'} L(t, \tilde{x}(t), \tilde{x}'(t)), \eta'(t)\rangle + \langle \nabla_x L(t, \tilde{x}(t), \tilde{x}'(t)), \eta(t)\rangle\, dt = 0$$

or

$$\int_\Omega [\langle \nabla_{\nabla u} L(x, \tilde{u}(x), \nabla \tilde{u}(x)), \nabla \eta(x)\rangle + L_u(x, \tilde{u}(x), \nabla \tilde{u}(x))\eta(x)]\, dx = 0$$

for every variation η such that η equals zero at the boundary. In the variational notation, the Euler lagrange equations are written as

$$\frac{d}{dt}\nabla_{x'}L(t,\tilde{x}(t),\tilde{x}'(t)) = \nabla_x L(t,\tilde{x}(t),\tilde{x}'(t))$$

or

$$\text{div } \nabla_{\nabla u}L(x,\tilde{u}(x),\nabla\tilde{u}(x)) = L_u(x,\tilde{u}(x),\nabla\tilde{u}(x))$$

Let $\tilde{x}$ be a solution to the problem of minimizing

$$\int_a^b L(t,x(t),x'(t))\,dt\,x(a)=\alpha,\ x(b)=\beta.$$

Let λ be a positive scalar, and η be an admissible variation. Then we have that

$$0 \le \frac{1}{\lambda}[\int_a^b L(t,\tilde{x}(t)+\lambda\eta(t),\tilde{x}'(t)+\lambda\eta'(t))\,dt - \int_a^b L(t,\tilde{x}(t),\tilde{x}'(t))\,dt] =$$

$$\int_a^b \frac{1}{\lambda}[\text{Ł}(t,\tilde{x}(t)+\lambda\eta(t),\tilde{x}'(t)+\lambda\eta'(t)) - L(t,\tilde{x}(t))+\lambda\eta(t),\tilde{x}'(t))]\,dt$$

$$+\int_a^b \frac{1}{\lambda}[\text{Ł}(t,\tilde{x}(t)+\lambda\eta(t),x'(t)) - L(t,\tilde{x}(t),\tilde{x}'(t))]\,dt.$$

Pointwise, the integrands converge to $\langle\nabla_{x'}L(t,\tilde{x}(t),\tilde{x}'(t)),\eta'(t)\rangle$ and to $\langle\nabla_x L(t,\tilde{x}(t),\tilde{x}'(t)),\eta(t)\rangle$ respectively. Establishing the validity of the Euler Lagrange equation consists simply in proving that one can pass to the limit under the integral sign. If one assumes that both $\tilde{x}(\cdot)$ and $\tilde{x}'(\cdot)$ are continuous on the compact interval $[a,b]$, and also that the gradients of the Lagrangean L are continuous in their arguments, there is no problem for passing to the limit: in fact, in this case, there is some number K that bounds from above both $\|\nabla_x L(t,\tilde{x}(t),\tilde{x}'(t))\|$ and $\|\nabla_{x'}L(t,\tilde{x}(t),\tilde{x}'(t))\|$, hence $|\langle\nabla_x L(t,\tilde{x}(t),\tilde{x}'(t)),\eta(t)\rangle|$ and $|\langle\nabla_{x'}L(t,\tilde{x}(t),\tilde{x}'(t)),\eta'(t)\rangle|$. By continuity and compactness, there is a $\delta > 0$ such that $\|\nabla_x L(t,y,\xi)\| \le (K+1)$ and $\|\nabla_{x'}L(t,y,\xi)\| \le (K+1)$ whenever $\|y-\tilde{x}(t)\| \le \delta$ and $\|\xi-\tilde{x}'(t)\| \le \delta$. By the mean value Theorem, the integrands are the scalar products $\langle\nabla_x L,\eta\rangle$ and $\langle\nabla_\xi L,\nabla\eta\rangle$ computed nearby the solution: hence, by the previous remark on continuity, they are dominated by some scalar and one can pass to the limit. The very same reasoning applies to the case where one minimizes

$$\int_\Omega L(x,u(x),\nabla u(x))\,dx$$

under additional boundary conditions: the assumption that both $\nabla\tilde{u}$ and $\tilde{u}$ are continuous on the closure of Ω (at least for bounded Ω) allows one to pass to the limit by dominated convergence. However, these assumptions, that both $\tilde{x}$ and $\tilde{x}'$ are continuous, are not satisfied even in very simple cases: consider, for instance, the problem of minimizing

$$\int_0^1 (x(t)x'(t))^2\,dt$$

under the conditions $x(0)=0, x(1)=1$. Then, the solution is $x(t)=\sqrt{t}$, whose derivative is unbounded on $[0,1]$.

2 The proof of the validity of the Euler Lagrange equations for integrals defined on an interval of R^1, and considerations on the general problem of its validity

Is it true that the validity of the Euler Lagrange equations depends only on suitable assumptions of regularity of the integrand (hence on conditions that can be checked a priori, i.e., without the knowledge of the minimizer $\tilde{x}$ and of its properties, in addition, possibly, to the natural assumption that the integral $\int_a^b L(t, \tilde{x}(t), \tilde{x}'(t))\, dt$ be finite? In [1], Ball and Mizel built a variational problem possessing a minimum $\tilde{x}$, such that $\int_a^b L(t, \tilde{x}(t), \tilde{x}'(t))\, dt$ is finite, but not satisfying the Euler Lagrange equations. What happens in the case of this example is that the function $t \to \nabla_x L(t, \tilde{x}(t), \tilde{x}'(t))$ is not in L^1. Hence the requirement of the integrability of the map $t \to \nabla_x L(t, \tilde{x}(t), \tilde{x}'(t))$ is essential to the proof of the validity of (E-L). So far, theorems on the validity of E-L require a stronger condition. When the Lagrangean L, as well as its gradients w.r.t x and x', satisfy Carathéodory conditions, i.e. they are measurable in t for fixed (x, x') and continuous in (x, x') for a. e. t, a suitable condition to assume is the existence of an integral bound for $\|\nabla_x L(t, y, \tilde{x}'(t))\|$ for y in a neighborhood of the solution $\tilde{x}(t)$. More precisely, one assumes the existence of a scalar δ and of a map $S \in L^1(I)$, such that, for a.e. t in I, we have that $\|y - \tilde{x}(t)\| \leq \delta$ implies $\|\nabla_\xi L(t, y, \tilde{x}'(t))\| \leq S(t)$. This condition, in particular, implies that the map $x \to L(t, x, \tilde{x}'(t))$ is locally Lipschitzean, in a neighborhood of $\tilde{x}(t)$, with Lipschitz constant $S(t)$. The validity of (E-L) under this condition of local Lipschitzianity w.r.t. x was first proved by Clarke [12] in the context of Differential Inclusions. A weaker condition, that does not imply this Lipschitzianity, has been recently presented by Ferriero and Marchini [13] for the standard case of the Calculus of Variations. Their method of proof consists in proving first the integrability of the term $\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|$ and then in deriving further regularity and the validity of (E-L). Here we sketch the main idea in the proof of the integrability of $\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|$. The idea is to write $\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|$ as

$$\langle \nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t)), \frac{\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))}{\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|} \rangle,$$

and this formula suggests naturally to consider using a variation η whose gradient is

$$\nabla \eta(t) = \frac{\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))}{\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|}.$$

Applying this idea, by Lusin's and Scorza Dragoni's Theorems, one infers the existence of C_n, closed subsets of I, such that $C_n \subset C_{n+1}$, $m(I \setminus C_n) \to 0$ and that both the restriction of x' to C_n is continuous and so is the restriction of $\nabla_\xi L$ to $C_n \times R^n \times R^n$. In particular, there are reals k_n that bound $\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|$ on C_n; by continuity and compactness, there is a $\delta > 0$ such that $\|\nabla_{x'} L(t, y, \tilde{x}'(t))\| \leq (k_n + 1)$ whenever $\|y - \tilde{x}(t)\| \leq \delta$ and $t \in C_n$. Consider

$$\theta(t) = \frac{\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))}{\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|}.$$

($\theta = 0$ whenever $\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t)) = 0$. For every n, call A_n the sets $C_n \setminus C_{n-1}$. One would like to use the restriction of θ to A_n as the gradient of a variation; since, in general,

$\int_{A_n} \theta(t)dt \neq 0$, so that the integral of θ would not be an admissible variation on I, fix subsets $B_n \subset C1$ such that $m(B_n) = \int_{A_n} \theta(t)dt$ and set

$$\theta_n'(t) = -\theta(t)\chi_{A_n}(t) + \frac{v_n}{\|v_n\|}\chi_{B_n}(t).$$

With this definition, θ_n is an admissible variation and one also infers that $\|\theta_n\|_{L^\infty} \leq 2m(A_n)$. Since both terms appearing in deriving (E-L) are (locally) bounded, the term with the derivatives w.r.t. x' by the choice of C_n and the one with the derivatives w.r.t. x by the assumption of local Lipschitzianity on that variable, one can pass to the limit and obtain that, since $\tilde{x}$ is a minimum,

$$\int_a^b \left[\langle \nabla_{x'} L(t, \tilde{x}(t), \tilde{x}'(t)), \theta_n'(t)\rangle + \langle \nabla_x L(t, \tilde{x}(t), \tilde{x}'(t)), \theta_n(t)\rangle\right] dt \geq 0,$$

i.e., that

$$\int_{A_n} \|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\| \leq \int_{B_n} \left[\langle \nabla_{x'} L(t, \tilde{x}(t), \tilde{x}'(t)), \theta_n'(t)\rangle + \langle \nabla_x L(t, \tilde{x}(t), \tilde{x}'(t)), \theta_n(t)\rangle\right] dt.$$

By the assumptions on the dependence of L on x and the choice of B_n (a subset of C_1, where $\|\nabla_{x'} L(t, \tilde{x}(t), \tilde{x}'(t))\|$ is bounded by k_1), there exists a constant C, independent of n, such that

$$\int_{A_n} \|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\| \leq Cm(A_n).$$

At this moment the proof is essentially completed, since the sequence of functions

$$\left(\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|\chi_{\cup_2^m A_n}(t)\right)_m$$

converges monotonically to $\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|\chi_{\cup_2^\infty A_n}(t)$. In addition, we have the bound k_1 on C_1. Hence we have established that the map $\|\nabla_\xi L(t, \tilde{x}(t), \tilde{x}'(t))\|$ is integrable on (a, b).

In order for the Euler Lagrange equation to make sense, we must have that both $\nabla_\xi L$ and $\nabla_x L$ be integrable along the solution. The integrability of $\nabla_\xi L$ and of $\nabla_x L$ along a given function does *not* follow from the fact that the integral of L exists finite when computed along that function. For example, when $L(\xi)$ is e^{ξ^2}, the integrability of $e^{\|x'(\cdot)\|^2}$ does not imply the integrability of $2\|x'(\cdot)\|e^{\|x'(\cdot)\|^2}$. The meaning of the previous reasoning is as follows:

to establish the validity of (E-L) we must impose *some* conditions in order to have the integrability of the terms $\nabla_{x'} L$ and $\nabla_x L$ along the solution. By imposing a condition on $\nabla_x L$ and building a suitable variation, one has obtained an integrable bound for $\nabla_{x'} L$ along the solution, and this has been accomplished without assuming that the solution and its derivatives are bounded, an assumption that would make the result inapplicable, and without imposing growth assumptions on the dependence of L on x', an assumption that would greatly limit the interest of the result.

Consider now a minimization problem for integrals defined on a multidimensional set, for example the problem of minimizing

$$\int_\Omega [f(\|\nabla u(x)\|) + u(x)]\, dx$$

where $f(\xi) = e^{\xi^2}$,under suitable boundary conditions. Here the condition of Lipschitzianity of L w.r.t. u along the solution $\tilde{u}(\cdot)$ is obviously satisfied, since

$$L(x, u, \nabla\tilde{u}(x)) - L(x, v, \nabla\tilde{u}(x)) = u - v.$$

Still, this author does not know of a result that would ensure that along the solution to this minimum problem the Euler Lagrange equation holds. The proof we have sketched for the case of integrals defined on an interval is based on using special variations that were obtained by defining in an appropriate way their derivatives on special given sets. Unfortunately, this cannot be accomplished on a multi-dimensional setting: we cannot define functions by (arbitrarily) prescribing their gradients on given sets! The validity of the Euler Lagrange equation for integrands having fast growth in ∇u is an open problem; some partial results can be found in [8]

3 The Maximum Principle

Around 1950, Pontriagin worked on an innovative minimum problem, that of minimizing

$$\int_a^b L(t, x(t), u(t))\, dt$$

with the additional conditions

$$x'(t) = f(t, x(t), u(t))$$

and $u(t) \in U$ The functions $u(\cdot)$ are called *controls*; the corresponding theory, the theory of *optimal control*. In the special case when the differential equation that connects the control u with the state x, i.e., the equation $x'(t) = f(t, x(t), u(t))$, becomes $x'(t) = u(t)$, so that $u(t)$ is simply a name given to $x'(t)$, we have again the problem of minimizing

$$\int_a^b L(t, x(t), x'(t))\, dt.$$

In this formulation, however, there appears the new condition that $x'(t)$ in this set "of controls" U. Hence, two are the novelties of the problem: the appearance of a dynamics (in general, non-linear) linking the variable in the Lagrangean with the state x and that of a constraint on the controls u, or, in the case of the Calculus of Variations, on the set of $x'(t)$ that are allowed.

According to the account given by Boltianski , Pontriagin proposed the Maximum Principle around as a *sufficient* condition (that it is not!); the proof of the validity of the Maximum

Principle as a *necessary* condition is due to Boltianski. The name of Maximum Principle derives from the basic condition proposed by Pontriagin : there exists a non-trivial vector function $(p_0(\cdot), p(\cdot))$, a solution to

$$p_0'(t) = 0; p'(t) = p_0 \nabla_x L(t, \tilde{x}(t), \tilde{u}(t)) - p(t) D_x f(t, \tilde{x}(t), \tilde{u}(t))$$

such that, for a.e. t in (a, b)

$$p_0 L(t, \tilde{x}(t), \tilde{u}(t)) - \langle p(t), f(t, \tilde{x}(t), \tilde{u}(t)) \rangle = max_{\omega \in U} p_0 L(t, \tilde{x}(t), w) - \langle p(t), f(t, \tilde{x}(t), w) \rangle$$

In the case of the Calculus of Variations, this condition reduces to: there exists $p_0(\cdot), p(\cdot)$, a solution to

$$p_0'(t) = 0; p'(t) = \nabla_x L(t, \tilde{x}(t), \tilde{u}(t))$$

such that, a.e. in (a, b),

$$-p_0 L(t, \tilde{x}(t), \tilde{u}(t)) + p(t)\tilde{u}(t) = \max_{w \in U} \{-p_0 L(t, \tilde{x}(t), w) + p(t)w\}.$$

The conditions proposed by Pontriagin differ from the conditions one usually meets even when the control differential equations reduces to $f(t, x, u) = u$, i.e., in the Calculus of Variations case. The main feature of these conditions are the lack of any differentiability condition w.r.t. x' and the role of the control set U. About this set, one would expect conditions of regularity: what is surprising is that, for the validity of the Maximum Principle, *there are no conditions* on U; U is *any* set. This fact has two important consequences: first, since U is not assumed to be an open set, the classical approach, consisting in "exploring" a neighborhood of the solution through variations, is not applicable. Second, since U can be a closed set, constrained problem with constraints $x' \in U$ are included. This formulation of the control problem appeared in the classical monograph [14]

Consider the following Example.(Example 1)

We wish to minimize

$$\int_{-a}^{+a} [F(x'(t)) + x(t)]\, dt;\ x(-a) = x(+a) = 0$$

where U is the set consisting of the four points $\{-2, -1, 1, 2\}$ and F equals 0 when $|x'| = 1$ and equals 1 when $|x'| = 2$. Extending F as $+\infty$ to the values that are forbidden, we wish to minimize

$\int_{-a}^{+a} [F(x'(t)) + x(t)]\, dt$ with $x(-a) = x(+a) = 0$ where

$$F(x') = \begin{cases} 0 & \text{when } |x'| = 1 \\ 1 & \text{when } |x'| = 2, \\ +\infty & \text{otherwise} \end{cases}$$

The solution to this problem (that depends on a), is derived by solving the equations obtained by the Maximum Principle. Because of the symmetry w.r.t. the origin, this

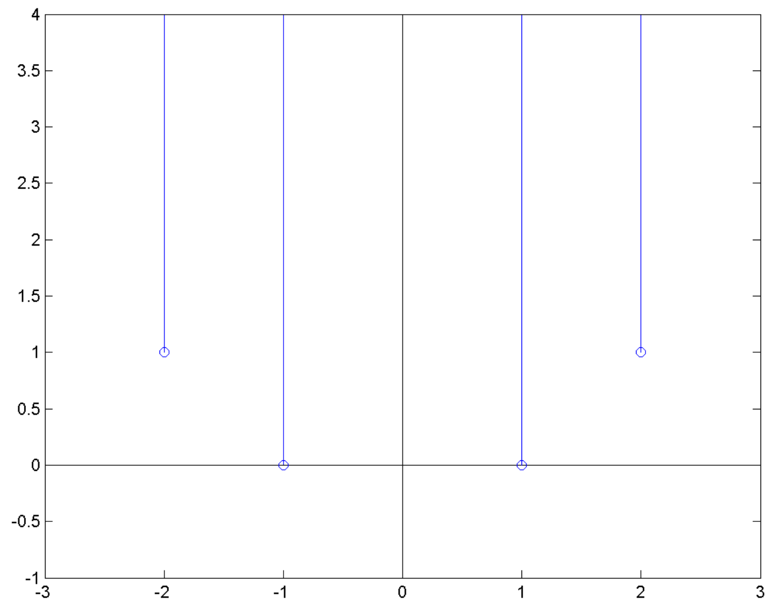


Figure 1: Epigraph of the function F

problem can conveniently be seen as a problem from $-a$ to 0 with free end point conditions at 0 (and the extends by symmetry to the interval $[0, a]$: in this way one obtains the "final" (at time $T = 0$) condition $p(0) = 0$. The differential equation for the variable p is simply

$$p'(t) = 1,$$

a very smooth differential equation (even for a non smooth problem as the one we consider!), so we obtain $p(t) = t$; the maximality condition connects the values of p and of x', in a way that will be clarified in the next section. The solution one obtains is described in the following picture.

By looking at this solution, one notices that the solution has a derivative in absolute value $= 1$ in the central part of the interval, while using larger values for the derivative at the extremes. Hence, from this remark, it follows that when the interval is small ($a \leq 1$), the solution will be such as to have derivative in norm equal 1 only.

4 Convex functions: the Euler Lagrange equation in semi-classical form

Let the Lagrangian be the sum of $F(x')$ and of $G(x)$, where G is a generic differentiable function while F is a convex function, possibly *extended valued*, i.e. possibly taking the value $+\infty$.

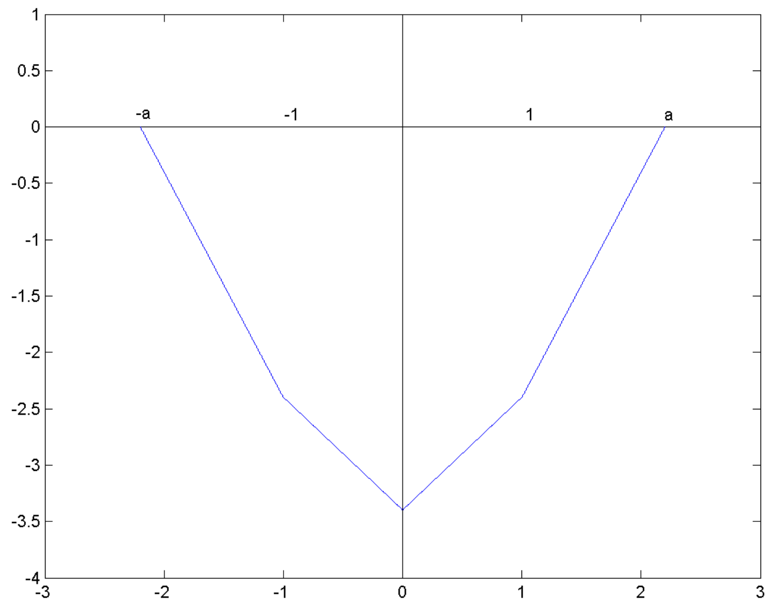


Figure 2: The solution $\tilde{x}$

For the problem we have just considered it makes no difference to have, under the integral sign, the function F finite only on the four points $\{-2, -1, 1, 2\}$ or its convexified, whose epigraph is the convex envelope of the epigraph of the original function. The solution we have found is as well solution to the problem where F is the convex function

$$F(x') = \begin{cases} 0 & \text{se } |x'| \leq 1 \\ |x'| - 2 & \text{se } 1 \leq |x'| \leq 2, \\ +\infty & \text{altrove} \end{cases}$$

The effective domain of a convex function (the set of points where the function takes a finite value) has convex closure and the convex function admits a (non empty) subdifferential at every interior point of its effective domain. If we assume that the subdifferential ∂F is non-empty at every point of the effective domain, the necessary conditions provided by the Maximum Principle for the minimization of

$$\int_a^b [F(x'(t)) + G(x(t))]\, dt$$

take the following simple form:

There exists an absolutely continuous $p(\cdot)$, a solution to

$$p'(t) = \nabla G(\tilde{x}(t))$$

such that, for a.e. t,

$$p(t) \in \partial F(\tilde{x}'(t))$$

In this formulation, the control set U is the effective domain of the convex function. The condition on $p(\cdot)$, solution to the differential equation, is that $p(t)$ is a *selection* from the set-valued map $t \to \partial F(\tilde{x}'(t))$. This formulation of the necessary conditions can be called the Euler Lagrange equation in semi-classical form.

We wish to point out that the Maximum Principle is more general than this last formulation: consider the problem of minimizing

$$\int_0^1 F(x'(t))\, dt;\ x(0) = 0, x(1) = 1$$

where F is the convex function

$$F(x') = \begin{cases} -\sqrt{1-(x')^2} & \text{se } |x'| \leq 1 \\ +\infty & \text{altrove} \end{cases}$$

The function $x(t) = t$ whose derivative $x'(t) \equiv 1$, is a solution. For this solution neither the classical Euler Lagrange equation is true (the derivative of F w.r.t. $x\prime$ does not exist along the solution) nor is true the semi-classical form (the subdifferential ∂F is empty along the solution). Still, the Maximum Principle holds: consider the vector $(p_0 \equiv 0, p \equiv 1)$, a non-trivial solution to $p_0' = F_{x_0} = 0$, $p' = F_x = 0$: one has

$$1 = \tilde{u}(t) = \max_{w \in U}\{w\}.$$

A sketch of the proof. We wish to present a sketch of the Proof of the Maximum Principle in its simplest form, for the problem with free right conditions, for two reasons: first, to point out the conceptual difference in taking variations; second, because it is overall a very beautiful proof. Consider the control system

$$y'(t) = F(t, y(t), u(t));\ y(0) = y^0,\ \text{ and } u(t) \in U$$

and, having fixed the final time T (but not the final state), assume that we wish to maximize $\psi(y(T))$, a given function of the final state. This special form of the optimization problem brings into evidence the geometric side of the proof. In the usual case, where we want to minimize

$$\int_a^b L(t, x(t), u(t))\, dt$$

subject to

$$x'(t) = f(t, x(t), u(t))$$

and $u(t) \in U$, it is enough to set: $y = (x^0, x)$; $F(t, (x^0, x), u) = (-L(t, x, u), f(t, x, u))$; $\psi((x^0, x)) = x^0$ to obtain a special case of the problem in this new formulation.

We will take a variation to the optimal control in the following way. Fix an arbitrary w in U. Let $\hat{y}(\cdot), \hat{u}(\cdot)$ a solution, and fix a time τ, a Lebesgue point for the maps $t \to F(t, \hat{y}(t), \hat{u}(t))$ and $t \to F(t, \hat{y}(t), w)$.

For every ε, define a new control u_ε as

$$u_\varepsilon(t) = \hat{u}(t) + \chi_{[\tau-\varepsilon,\tau]}(t)(w - \hat{u}(t))$$

so that we substitute w to $\hat{u}(t)$ on the interval $[\tau - \varepsilon, \tau]$. As opposite to the classical variations, there is no attempt to let w "tend to " $u(t)$. Our purpose is to estimate the effect of this variation, taken at time τ, at the final time T, so as to compare the result, $y_\varepsilon(T)$, with the solution $\hat{y}(T)$. Let us first compute the effect of the variation at time τ, i.e. let us estimate the difference $v_\varepsilon(\tau) = y_\varepsilon(\tau) - \hat{y}(\tau)$. We have

$$y_\varepsilon(\tau) - \hat{y}(\tau) = \int_{\tau-\varepsilon}^{\tau} [F(t, y_\varepsilon(t), w) - F(t, \hat{y}(t), \hat{u}(t))]\, dt =$$

$$= \int_{\tau-\varepsilon}^{\tau} [F(t, y_\varepsilon(t), w) - F(t, \hat{y}(t), w)]\, dt + \int_{\tau-\varepsilon}^{\tau} [F(t, \hat{y}(t), w) - F(t, \hat{y}(t), \hat{u}(t))]\, dt.$$

Assuming some boundedness (local, near the point $(\tau, \hat{y}(\tau))$ of the map $(t, y) \to F(t, y, w)$, the difference $y_\varepsilon(t) - \hat{y}(t)$ is bounded on $[\tau - \varepsilon, \tau]$ by $K\varepsilon$. Assuming that the map $y \to F(t, y, w)$ is Lipschitzean, we obtain the the first integral at the r.h.s. satisfies

$$\| \int_{\tau-\varepsilon}^{\tau} [F(t, y_\varepsilon(t), w) - F(t, \hat{y}(t), w)]\, dt\| \leq K_1\, \varepsilon^2.$$

Here, by the choice of τ, we have

$$\int_{\tau-\varepsilon}^{\tau} [F(t, \hat{y}(t), w) - F(t, \hat{y}(t), \hat{u}(t))]\, dt = \varepsilon\, ([F(\tau, \hat{y}(\tau), w) - F(\tau, \hat{y}(\tau), \hat{u}(\tau)))] + O(\varepsilon)).$$

Hence we have obtained that $v_\varepsilon(\tau) = \varepsilon\, ([F(t, \hat{y}(t), w) - F(t, \hat{y}(t), \hat{u}(t))] + O(\varepsilon)) = \varepsilon\, (v_\tau + O(\varepsilon))$.

On the interval $[\tau, T]$, the differential equations for y_ε and for $\hat{y}$ are the same, i.e.

$$y'(t) = F(t, y(t), \hat{u}(t))$$

and the two solutions differ by the initial (at time τ) condition.

By a basic result on the differentiability of a solution with respect to initial conditions, one obtains that the difference $y_\varepsilon(T) - \hat{y}(T)$ can be written as

$$y_\varepsilon(T) - \hat{y}(T) = \varepsilon[v(T) + O(\varepsilon)]$$

where $v(T)$ is the solution at time T to the Cauchy problem

$$v'(t) = D_y F(t, \hat{y}(t), \hat{u}(t))\, v(t);\;\; v(\tau) = v_\tau.$$

Here, by $D_y F(t, \hat{y}(t), \hat{u}(t))$ we mean the matrix of partial derivatives of F w.r.t. y, computed along the solution $\hat{y}$, $\hat{u}$.

Since $\hat{y}$ is a maximum, we must have that

$$\langle \nabla\psi(\hat{y}(T)), v(T)\rangle \leq 0.$$

It is this geometric condition at time T that we wish to transfer at time τ: let $p(\cdot)$ be a solution to the Cauchy problem:

$$p'(t) = -p(t)D_yF(t, \hat{y}(t), \hat{u}(t));\ y(T) = \nabla\psi(\hat{y}(T)).$$

By the product rule for derivatives, we obtain that

$$\frac{d}{dt}\langle p(t), v(t)\rangle = 0,$$

and, since the maximum condition gives $\langle p(T), v(T)\rangle \leq 0$, it follows that, for every t,

$$\langle p(t), v(t)\rangle \leq 0,$$

in particular at $t = \tau$. We have already obtained that, at almost every τ, $v(\tau) = F(t, \hat{y}(t), w) - F(t, \hat{y}(t), \hat{u}(t))$, so that we have obtained that, at almost every τ,

$$\langle p(\tau), F(t, \hat{y}(t), w)\rangle \leq \langle p(\tau), F(t, \hat{y}(t), \hat{u}(\tau))\rangle.$$

Since w was arbitrary in U, we have proved the Maximum Principle.

The variations we have used are called "needle variations"; as ε tends to zero, the graph of such a variation in effect looks like a needle.

Although some special results in the taste of the Maximum Principle have been proved for multi-dimensional problems, the reasoning used are *not* an extension of what has been presented above for one-dimensional domains.

5 Multi dimensional domains

Let us consider again Example 1 , but on a domain $\Omega \subset R^N$: we wish to minimize

$$\int_\Omega [F(v(x)) + u(x)]dx$$

where

$$F(\nabla u) = \begin{cases} 0 & \text{se } \|\nabla u\| \leq 1 \\ \|\nabla u\| - 2 & \text{se } 1 \leq \|\nabla u\| \leq 2, \\ +\infty & \text{altrove} \end{cases}$$

and zero boundary condition. In Example 1, in the one-dimensional case, when the domain of integration was "small", the solution was such that its derivative was $-dist(t, \partial I)$, i.e. the distance to the boundary with a negative sign in front. On a multi-dimensional domain will it be true that when the domain is "small"(in what sense?) the function "distance to the boundary" (with a negative sign in front) is a solution? And

how small the domain has to be?

If the proposed function $u(x) = -dist(x, \partial\Omega)$ is a solution, its gradient is a.e. such that $\|\nabla u\| = 1$ so that

$$\partial F(\nabla u(x)) = \{\lambda \nabla u(x) : 0 \leq \lambda \leq 1\}.$$

The semiclassical formulation of the Euler Lagrange equation, implies the existence of $p(x)$,a solution

to the differential equation

$$\text{div } p(x) = 1$$

such that $p(x) = \lambda(x)\nabla u(x)$ where $0 \leq \lambda \leq 1$.

Hence, we wish to solve

$$\text{div } p(x) = 1$$

$$p(x) = \lambda(x)\nabla u(x)$$

i.e., $\langle \nabla\lambda(x), \nabla u(x)\rangle + \lambda(x)\text{div } \nabla u(x) = 1$. Integrating along the trajectories of the ordinary differential equation

$$x'(t) = \nabla u(x(t))$$

we have that the Euler Lagrange partial differential equation becomes

$$\frac{d}{dt}\lambda(x(t)) + \lambda(x(t))\text{div } \nabla u(x(t)) = 1$$

The important point that we wish to make is that, by knowing the properties of $\partial\Omega$, we can compute the map $x \to div\nabla u(x)$, hence integrate the resulting equation for λ. As an example, let us consider (Example 2), the two-dimensional domain (depending on a parameter ℓ)

$$\Omega_\ell = \{(x, y) : |y| \leq \ell; |x| \leq \ell + \sqrt{\ell^2 - y^2}\}$$

We believe that the function $u(x) = -dist(x, \partial\Omega)$ is the solution when ℓ is sufficiently small.

To get insight on the problem, let us integrate the differential equation

$$\frac{d}{dt}\lambda(x(t)) + \lambda(x(t))\text{div } \nabla u(x(t)) = 1$$

from two different initial conditions, namely, $P_1 = (0, 0)$ e $P_2 = (1, 0)$; along the trajectories defined by the vector field $\nabla u(x)$ we have, in the first case, div $\nabla u(x) = 0$, while, in the second case, div $\nabla u(x) = \frac{1}{\|x\|}$. The solutions to the corresponding differential equations satisfying the initial condition $\lambda(0) = 0$, are: when issued from P_1, the solution is $\lambda_1 = t$, while, when issued from P_2, is $\lambda_2 = \frac{1}{2}t$. To provide an answer to the problem of "how large" Ω has to be, notice that the vector $p(x) = \lambda(x)\nabla u(x)$ ceases to belong to the subdifferential $\partial F(\nabla u(x))$, computed for $\|\nabla u\| = 1$, when $|\lambda| \geq 1$. This happens at $t = 1$ for the initial point P_1 and at $t = 2$ for P_2. The variable t in this case represents the distance from P_1 or from P_2. Summarizing what we have found, we can say that when $\ell \leq 1$, the map $u(x) = -dist(x, \partial\Omega_\ell)$ is a solution to the minimum problem.

These computations, for a generic convex Ω with piecewise smooth boundary, have been introduced in [6] in R^2 and in [15] in R^N; Celada-Cellina [5] show that, when Ω is a square, this condition cannot be improved.

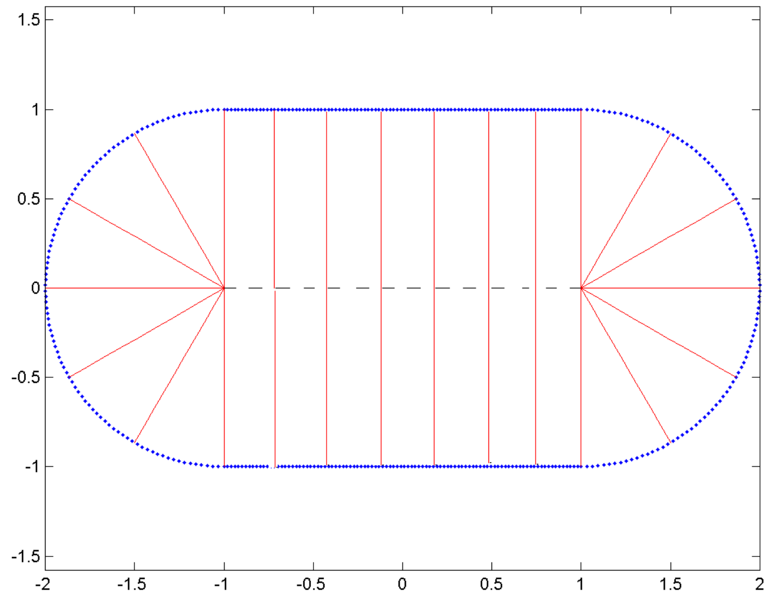


Figure 3: Ω_ℓ for $\ell = 1$

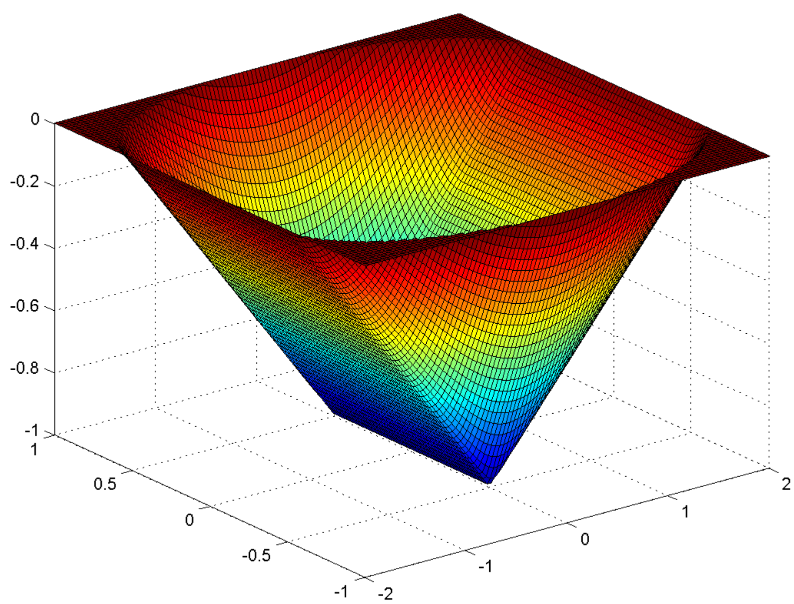


Figure 4: The function "‘- distance to the boundary of a square"‘

6 Minimum problems on multi-dimensional domains

Let F be a convex function, finite on R^N. Let $\tilde{u}$ be a solution to the minimum problem

$$\int_\Omega [F(\nabla u(x)) + G(u(x))]dx$$

under the constraint $\|\nabla u\| \leq 1$. There are three ways to look at this problem:

1. As an optimal control problem: to minimize the integral functional, depending on the control $v(x)$ and the state u

   $$\int_\Omega [F(v(x)) + G(u(x))]dx$$

   under the (Hamilton Jacobi) differential equation and the control set

   $$\nabla u = v \text{ con } v \in U = B_1.$$

   The natural conditions should be expressed in the form of a Maximum Principle.

2. As an unconstrained problem of the Calculus of Variations, that of minimizing

   $$\int_\Omega [\tilde{F}(\nabla u(x)) + G(u(x))]dx$$

   where $\tilde{F}$ is the extended-valued convex function

   $$\tilde{F}(\nabla u) = \begin{cases} F(\nabla u) & \text{if } \|\nabla u\| \leq 1 \\ +\infty & \text{otherwise} \end{cases}$$

   From the assumption that the original F was finite on R^N, it follows that the subdifferential of $\tilde{F}$ is non-empty on its effective domain, we expect that the necessary conditions be expressed by the validity of the Euler Lagrange equation in its semi-classical form.

3. As a minimization problem of a functional *constrained* to the closed, bounded and convex set K of those functions u satisfying

   $$u|_{\partial\Omega} = u^0|_{\partial\Omega} \text{ e } \|\nabla u\| \leq 1.$$

   The necessary conditions traditionally associated to this formulation of the problem consists in a *variational inequality*: we should prove that

   $$\int_\Omega [\langle \nabla F(\nabla \tilde{u}(x)), \nabla \eta(x) \rangle + G_u(\tilde{u}(x))\eta(x)]\, dx \geq 0$$

   for every *admissible* variation η ,i.e. such that
   $\eta = u - u^0$ with $u \in K$.

While the two first points of view are equivalent one to the other, and both lead to conditions like Pontriagin's conditions: there exists $p(x) \in \partial\tilde{F}(\nabla\tilde{u}(x))$:

$$\int_\Omega [\langle p(x), \nabla\eta(x)\rangle + G'(\tilde{u}(x))\eta(x)]\, dx = 0$$

for every sufficiently regular η that vanishes at the boundary of Ω, the third is not. In fact, it is enough to consider the case where the conditions

$$u|_{\partial\Omega} = u^0|_{\partial\Omega} \text{ and } \|\nabla u\| \le 1$$

are met by only one function, the boundary datum u^0 itself. Then, this datum is the solution to the minimum problem, since every other function gives the value $+\infty$ to the functional. The point of view of differential inequalities is useless in this case, since the condition that a certain inequality be satisfied by every admissible variation η gives no information since there are *no* non-trivial admissible variations. The condition involving a selection from the subdifferential computed along the solutions, instead, have to be verified for every variation η, independent on whether they are admissible or not, and provide useful information on the solution even in this case.

7 The few known results and a sketch of the proof of a basic result

Has the validity of these necessary conditions been established? Actually, only in a few special cases.

1. Brézis [3], [4], proves the validity of the semi-classical formulation of the Euler Lagrange equation for the problem of minimizing the integral functional where

$$F(\xi) = \tfrac{1}{2}\|\xi\|^2 \;;\; G(u) = u \;;\; u^0 = 0 \text{ and}$$
$$\Omega \text{ is a smooth convex set}$$

2. Cellina-Perrotta [10] prove the validity of these conditions for the problem where

$$F(\xi) = 0 \;;\; G(u) \text{ a strictly monotonic function} \;;\; u^0 \text{ Lipschitzian and}$$
$$\Omega \text{ a bounded open set}$$

3. Cellina [9] proves the condition when

$$F(\xi) = \tfrac{1}{2}\|\xi\|^2 \;;\; G(u) = 0 \;;\; \Omega \text{ a circular } R^2 \text{ having radius } \rho \text{ and}$$
$$u^0(x) = 1 - \tfrac{1}{\rho}\|x\|$$

Some details on the result of Cellina-Perrotta.

We wish to minimize

$$\int_\Omega [i_B(\nabla u(x)) + G(u(x))]dx$$

where i_B is indicator of the unit ball B, i.e.,

$$i_B(\nabla u) = \begin{cases} 0 & \text{when} \|\nabla u\| \leq 1 \\ +\infty & \text{elsewhere} \end{cases}$$

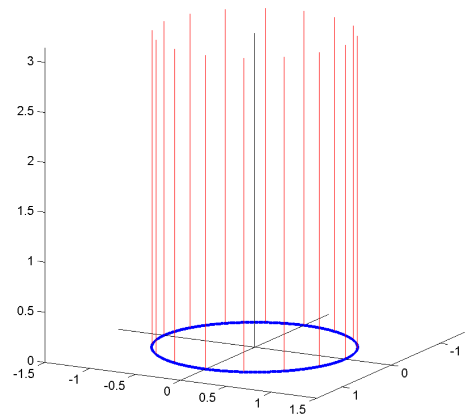


Figure 5: The function i_B

and $G(u)$ is strictly monotonic, the boundary condition u^0 is Lipschitzian and Ω is an open bounded set (no regularity is needed). As it will be evident in the proof, we might as well have that $G = G(x, u)$, as long as the strict monotonicity requirement w.r.t. u holds for a.e. $x \in \Omega$.

The main properties of the solution $\tilde{u}$, that will be used to prove the validity of the Euler-Lagrange equation are derived from two Lemmas. The first lemma says: let x_0 be any point in Ω and let $r > 0$ be such that $B(x_0, r) \subset \Omega$. Then:

$$\sup_{x \in \Omega : \|x - x_o\| = r} \tilde{u}(x) - \tilde{u}(x_0) = r.$$

In other words: given any point in Ω (*not*: almost every point!) and $r > 0$, there is a point x at a distance r from x_0, such that $u(x) - u(x_0)$ is the maximum possible; as a consequence (but this consequence is not strong enough for our purposes) the gradient of u equals 1 whenever it exists. The proof of this Lemma goes by showing that, if the claim of the Lemma is not true, one can build a variation $\eta_\lambda(\cdot)$, $\eta_\lambda(x) \leq 0$ and $\eta_\lambda(x) \not\equiv 0$, such that, for every sufficiently small $t > 0$, $\tilde{u}(\cdot) + t\eta_\lambda(\cdot)$ is admissible, i.e., for a.e. x, $\|\nabla\tilde{u}(x) + t\nabla\eta_\lambda(x)\| \leq 1$. Since this variation strictly decreases (pointwise) the value of $G(u(x))$, it strictly decreases the value of the functional, a contradiction. So the whole point relies on the construction of $\eta_\lambda(\cdot)$, and this is a delicate task: in general, one should expect that the gradient of the solution $\tilde{u}$ already satisfies $\|\nabla\tilde{u}(x)\| = 1$ a.e., so it is not easy to obtain a nontrivial η_λ giving $\|\nabla\tilde{u}(x) + t\nabla\eta_\lambda(x)\| \leq 1$.

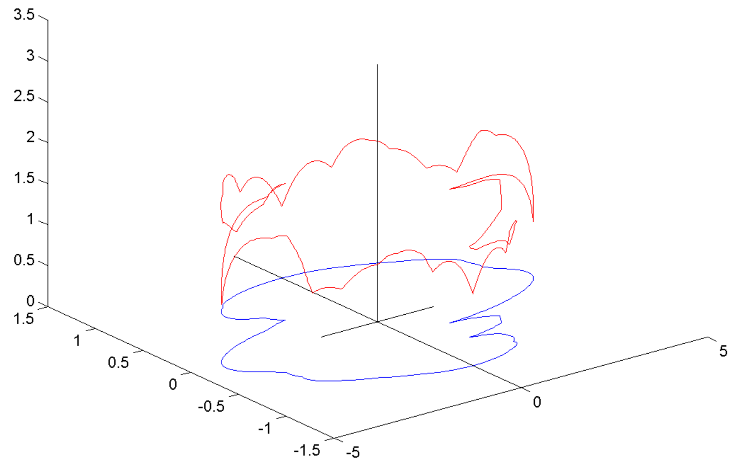


Figure 6: u^0

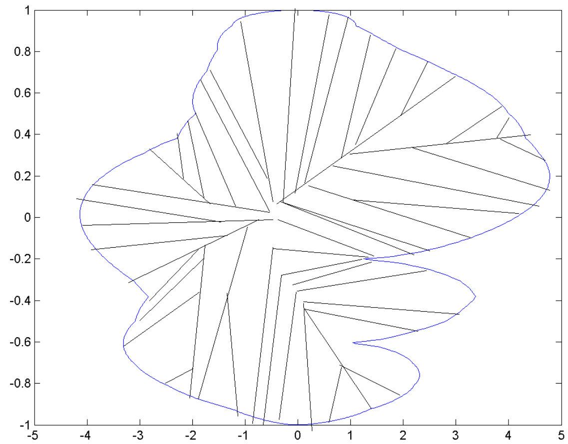
By this lemma, through each point x_0 there is (at least) a direction of maximal growth for the function $\tilde{u}$. Now, the lemma can be applied to the point x, at the boundary of $B(x_0, r)$, of maximal growth from x_0: from it, a new direction of maximal growth is defined: however, it turns out that this new direction cannot differ from the previous one, otherwise we would contradict the Lipschitzianity of $\tilde{u}$. Hence the process can be continued along the same direction until we reach the boundary of Ω, so that from the initial arbitrary x_0 there exists a half line of maximal growth reaching the boundary. These half lines cannot cross in Ω (but several half lines can be issued from the same point). Hence, the collection of these half lines (actually,their intersection with Ω), in addition to a set of measure zero, gives a partition of Ω.

The properties of the solution $\tilde{u}$ obtained by this Lemma are independent of the function $G(u(x))$, or $G(x, u(x))$, as long as it is pointwise monotonic in u. Notice that there is no problem in proving the existence of a solution to the minimum problem, since the functional is certainly coercive. Hence through this Lemma one proves the existence of a special solution to boundary value problem for a Hamilton Jacobi inclusion as

$$\|\nabla\tilde{u}(x)\| = 1;\ u|_{\partial\Omega} = u^0$$

having a special property of maximality among the solutions of the same inclusion; sometimes, this problem is faced through the definition of viscosity solutions.

Although the previous properties give an accurate description of the behaviour of the solution $\tilde{u}$, they are not sufficient to prove the validity of (E-L) for this problem: some additional regularity will have to be proved.



Since we have that $\partial F(\xi) = \alpha\frac{\xi}{\|\xi\|}$ and $\|\nabla\tilde{u}(x)\| = 1$ for a.e. x, to prove the validity of the Euler Lagrange equation, in its semiclassical form, we have to show that there exists a real valued $\alpha(x)$ such that:

1. $\alpha(x) \geq 0$

2. for every test function (or variation) $\eta \in C_c^\infty(\Omega)$, one has

$$\int_\Omega \alpha(x)\langle\nabla\tilde{u}(x), \nabla\eta(x)\rangle dx + \int_\Omega G'(\tilde{u}(x))\eta(x)dx = 0$$

Hence we have to define $\alpha(x)$ and, given a variation η, compute

$$\int_\Omega \alpha\langle\nabla\tilde{u}, \nabla\eta\rangle dx$$

and show that it is zero.

The idea is to define $\alpha(x)$ on each half line, and to perform the integration on Ω by a change of variables, integrating first along each line; this method will allow us to exploit all the properties of integrals defined on subsets of the real line.

Fix one component, say the k-th, of a $N-$ vector x to be q and consider E_q, the set of all vectors $x \in \Omega$ such that $x_k = q$. Call $\hat{x}$ the $N-1$ dimensional vector $x_1, ..., x_{k-1}, x_{k+1}, ..., x_N$. Set $X(\hat{x}, t)$ to be the solution at time t of the Cauchy problem

$$\frac{d}{dt}X(t) = \nabla u(X(t)); \;\; X(0) = x_1, ..., x_{k-1}, q, x_{k+1}, ..., x_N.$$

The map $x \to X(\hat{x}, t)$ can be seen as a change of variables from x to $(\hat{x}, t)$. In order to take advantage of this change of variables for our integration, we must investigate its regularity, hence the regularity of the vector field appearing at the r.h.s. of the differential equation, i.e., $\nabla u(x)$. We have already found that this is a field of unit vectors, or directions; however, in order to make use of the change of variables $x \to X(\hat{x}, t)$ under the integral sign, it must be differentiable a.e.

The second Lemma is about the regularity (differentiability a.e.) of a field of directions, described by the map $d(x)$, with d a unit vector, with the properties described above. This Lemma can be described as follows. Consider again the set E_q. From every point of this set is issued (at least) one half line (described by a direction), going to the boundary; we can consider $(x + a(x)d(x), x + b(x)d(x))$, the maximal open interval (of R^N) contained in Ω, so that the point $x + b(x)d(x)$ will be on $\partial\Omega$. On E_q, consider only (for the moment) those x's such that the corresponding interval $(x + a(x)d(x), x + b(x)d(x))$ has two properties: it is (uniformly) transversal to E_q, in the sense that d_k, the k-th component of $d(x)$, is at least $\frac{1}{\sqrt{N}}$, and it extends to both sides of E_q by at least ε: $(x - \varepsilon d(), x + \varepsilon d(x)) \subset (x + a(x)d(x), x + b(x)d(x))$. The Lemma then says that the map $x \to d(x)$ is Lipschitzian with Lipschitz constant $\frac{2\sqrt{N}}{\varepsilon}$.

By this Lemma one easily obtains that (for x restricted to this subset of points of E_q with the properties above), $JX(\hat{x}, t)$, the Jacobian of the map $X(\hat{x}, t)$, exists and it is non zero.

We have identified a subset of Ω, consisting of a family of intervals reaching the boundary of Ω and a map, that on this family is enough regular. We can obtain a countable family of sets, covering Ω, with the exception of a set of measure zero, by changing in a countable way the value q of the k-th coordinate and by considering all the coordinates k in $(1, ..., N)$. Hence, the integration over Ω will be carried on the union of this countable family of sets.

Assume we have defined α as

$$\alpha(\hat{x}, t) = \frac{1}{JX(\hat{x}, t)} \int_{a(\hat{x})}^{t} g'(u(X(\hat{x}, t)))J(X(\hat{x}, t)\, dt;$$

(actually, the true definition of α has to be that of a map on Ω, hence in the variable x, not in the variables $(\hat{x}, t)$: we should compose it with the inverse of X). Fix any η and perform the integral

$$\int_\Omega \alpha(x)\langle \nabla\tilde{u}(x), \nabla\eta(x)\rangle dx + \int_\Omega G'(\tilde{u}(x))\eta(x) dx.$$

Consider the first term: by the change of variables formula; the definition of α and Fubini's Theorem in the variables $(\hat{x}, t)$, one obtains

$$\int_\Omega \alpha(x)\langle \nabla\tilde{u}(x), \nabla\eta(x)\rangle dx =$$

$$= \int \left(\int_{a\hat{x})}^{b(\hat{x})} \left(\int_{a(\hat{x})}^{t} g'(u(X(\hat{x}, s)))J(X(\hat{x}, s)\, ds \right) \langle \nabla\eta(X(\hat{x}, t)), \nabla\tilde{u}(X(\hat{x}, t))\rangle\, dt \right) d\hat{x}$$

(the integral at the r.h.s. has to be performed over the union of the countable family of sets we have defined.)

Consider the map $t \to \eta(X(\hat{x}, t))$: we have

$$\frac{d}{dt}\eta(X(\hat{x}, t)) = \langle \nabla\eta(X(\hat{x}, t)), \frac{d}{dt}X(\hat{x}, t)\rangle = \langle \nabla\eta(X(\hat{x}, t)), \nabla\tilde{u}(X(\hat{x}, t))\rangle,$$

hence

$$\int_\Omega \alpha\langle\nabla\tilde{u},\nabla\eta\rangle dx =$$

$$= \int \left(\int_{a\hat{x})}^{b(\hat{x})} \left(\int_{a(\hat{x})}^{t} G'(u(X(\hat{x},s)))J(X(\hat{x},s)\,ds\right)\frac{d}{dt}\eta(X(\hat{x},t))\,dt\right) d\hat{x}.$$

Integrating by parts, and noticing that $\eta(X(\hat{x},b(\hat{x}))) = 0$, we have

$$\int_\Omega \alpha\langle\nabla\tilde{u},\nabla\eta\rangle dx =$$

$$= -\int \left(\int_{a(\hat{x})}^{b(\hat{x})} \eta(X(\hat{x},t))G'(\tilde{u}(X(\hat{x},t))J(\hat{x},t)dt\right) d\hat{x} =$$

$$= -\int_\Omega \eta(x))G'(\tilde{u}(x)\,dx.$$

We have proved that

$$\int_\Omega \alpha(x)\langle\nabla\tilde{u}(x),\nabla\eta(x)\rangle dx + \int_\Omega G'(\tilde{u}(x))\eta(x)dx = 0,$$

i.e., the validity of the Euler Lagrange equation.

An important remark: although, as we have noticed, the solution $\tilde{u}$ is *independent* of $G(u)$ or $G(x,u)$, as long as this map is strictly monotonic in u, the function α *does depend* on G.

A connection with problems of optimal transportation.

In Bouchitté lectures on Optimal Transportation [2] one finds the problem of maximizing

$$\langle f, u\rangle$$

on those functions u that are Lipschitzian, such that $\|\nabla u\| \le 1$, $u = 0$ on Σ, where f is a measure on Ω, and $\Sigma \subset \Omega$, under the condition the either $f(\Omega) = 0$ or that $\Sigma \ne 0$. The problem we have considered is very similar to the second case: take $\Sigma = \partial\Omega$, $f = -$ the Lebesgue measure, and the problem translates into the problem of minimizing

$$\int_\Omega [F(\nabla u(x)) + u(x)]dx$$

where

$$F(\nabla u) = \begin{cases} 0 & \text{when} \|\nabla u\| \le 1 \\ +\infty & \text{elsewhere} \end{cases}$$

and $u|_{\partial\Omega} = 0$. We have treated this problem for a general boundary condition $u = u^0$ at the boundary of Ω.

8 A connection with the fundamental theory of ordinary differential equations

Under what conditions a solution to an ordinary differential equation of the kind

$$x'(t) = \nabla\tilde{u}(x(t))$$

can be defined, in such a way that it is possible to integrate along the corresponding trajectories? The theory of differential equations during last century has considerably weakened the conditions to be imposed to the right hand side, passing from the condition of continuity to Carathéodory conditions, then to the requirement of measurability w.r.t. x (solutions in the sense of Filippov), but the problem one had in mind was the Cauchy problem for a given initial datum. Today, one would need a different theory.

Our purpose is to perform a multiple integration by means of successive integrations, by integrating first on the initial data and then on the trajectories to a given differential equation, arising from a variational problem. Since, in any case, we will integrate also on initial data, what happens on sets of zero measure of initial data it is of no importance, so that it is not the single Cauchy problem that matters. Moreover, the right hand side of the equation is a map in a Sobolev space, not a regular map.

The theory one wishes to develop is not a generalization of theorems like Peano's theorem for continuous r.h.s.: solutions provided by this theorem do not possess the properties required by a change of variables formula. It is enough to consider the equation

$$x'(t) = f(x(t))$$

where

$$f(x) = \begin{cases} \sqrt{x} & \text{when} x \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

The set $A \subset \Omega$, as presented in the figure, is of positive measure in the plane, but it is seen as a set of measure zero if we compute its area by successive integrations. The problem of the uniqueness of the solution to a Cauchy problem for almost all initial conditions is very difficult. Some partial results can be found in [7] and in [11]

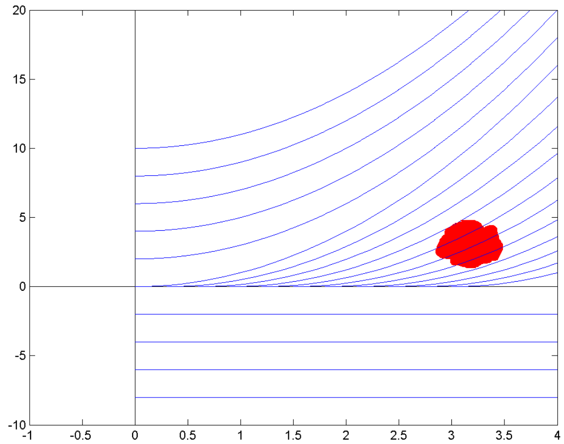


Figure 7:

References

[1] J.M. Ball, V.J. Mizel *One dimensional minimum problems whose minimizers do not satisfy the Euler Lagrange equation* Arch. Rat. Mech. Anal. 90 (1985), 325-388

[2] G. Bouchitté *Lectures on Optimal Transportation* Cortona, August 2004

[3] H. Brezis *Multiplicateur de Lagrange en torsion elasto-plastique* Arch. Rat. Mech. Anal. (1973), 32-40

[4] H. Brezis, M. Sibony *Equivalence de deux inequations variationelles et applications* Arch. Rat. Mech. Anal. 41 (1971), 254-265

[5] P. Celada, A. Cellina *Existence and non existence of solutions to a variational problem on a square* Houston J. of Math. 24 (1998), 345-375

[6] A. Cellina *Minimizing a functional depending on* ∇u *and* u , Ann. Inst. H. Poincaré, Anal. Non Lineaire, 14 (1997), 339-352

[7] A. Cellina *On uniqueness almost everywhere for monotonic differential inclusions* Nonlinear Analysis TMA 25 (1995), 899-903

[8] A. Cellina *On the validity of the Euler Lagrange equation* Jour. Diff. Equat. 171 (2001), 430-442

[9] A. Cellina *On a constrained Dirichlet problem* SIAM J. Control Optim. 41 (2002). 331-344

[10] A. Cellina, S. Perrotta *On the validity of the Maximum Principle and of the Euler Lagrange equation for a minimum problem depending on the gradient* SIAM J. Control Optim. 36 (1998), 1987-1998

[11] A.Cellina, M. Vornicescu *On gradient flows* Jour. Diff. Equat. 145 (1998), 489-501

[12] F. Clarke *Optimization and non smooth analysis* Wiley-Interscience, New York, 1983

[13] A. Ferriero, E.M. Marchini *On the validity of the Euler Lagrange equation*, a preprint

[14] L. Pontriagin, V. Boltyianskii, R. Gamkrelidze, E. Mishchenko *The Mathematical Theory of Control Processes* Interscience, New York, 1962

[15] M. Vornicescu *A variational Problem on subsets of* R^N , Proc. Royal. Soc. Edinburgh Sect. A, 127 (1997), 1089-1101