Mémoires de la S. M. F.

Ivar Ekeland

Nonconvex duality

Mémoires de la S. M. F., tome 60 (1979), p. 45-55
<http://www.numdam.org/item?id=MSMF_1979__60__45_0>

© Mémoires de la S. M. F., 1979, tous droits réservés.

L'accès aux archives de la revue « Mémoires de la S. M. F. » (http://smf.emath.fr/Publications/Memoires/Presentation.html) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

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

Analyse non convexe [1977. Pau]
Bull. Soc. math. France,
Mémoire 60, 1979, p. 45-55.

NONCONVEX DUALITY

Ivar EKELAND

§.1 THE LEGENDRE TRANSFORM.

Let X be a reflexive Banach space, and X^* its dual. Typically, an element of X will be denoted by x and an element of X^* by x^*. The Legendre transform for X is the mapping $\mathcal{L}_X$ from $X \times X^* \times \mathbb{R}$ into $X^* \times X \times \mathbb{R}$ defined by :

$$\mathcal{L}_X(x,x^*,a) = (x^*,x, \langle x,x^*\rangle - a) .$$

Applying this definition to the Banach space X^*, with dual $X^{**} = X$, we get the map $\mathcal{L}_{X^*}$ from $X^* \times X \times \mathbb{R}$ into $X \times X^* \times \mathbb{R}$ defined by :

$$\mathcal{L}_{X^*}(x^*,x,b) = (x,x^*, \langle x,x^*\rangle - b) .$$

It is obvious that $\mathcal{L}_X$ and $\mathcal{L}_{X^*}$ are C^∞ maps (i.e. indefinitely differentiable). A simple calculation shows that they are inverse to each other :

$$\mathcal{L}_X \circ \mathcal{L}_{X^*} = \mathrm{Id}_{X^* \times X \times \mathbb{R}}$$

$$\mathcal{L}_{X^*} \circ \mathcal{L}_X = \mathrm{Id}_{X \times X^* \times \mathbb{R}}$$

This implies that they are non-linear diffeomorphisms, so that they map closed smooth subsets onto closed smooth subsets. I will now define what I mean by "extremizing" a subset of $X \times X^* \times \mathbb{R}$, and study the effect of the Legendre transform on such problems. This might seem unnatural (although very simple) for the time being, but will be fully justified in the next section.

DEFINITION 1. Let C be a subset of $X \times X^* \times \mathbb{R}$. A point $(x,x^*,a) \in C$ is extremal in C iff $x^* = 0$.
The problem of finding such points is denoted by :

$(\mathcal{P})$ ext C

The set of solutions to $(\mathcal{P})$ is :

$$\{\mathrm{ext}(\mathcal{P})\} = C \cap (X \times \{0\} \times \mathbb{R}) .$$

We now consider another reflexive Banach space Y, and a subset K of $(X \times Y) \times (X^* \times Y^*) \times \mathbb{R}$.

For any $y_0 \in Y$ there is a slice $K_{y=y_0}$ of K, which is a subset of $X \times X^* \times \mathbb{R}$:

$$K_{y=y_0} = \{(x,x^*,a) \mid \exists\, y^* \in Y^* : (x,y_0;x^*;y^*;a) \in K\}$$

and there is a corresponding extremization problem :

(q_y) $\quad$ ext $K_{y=y_0}$

If we have been careful to get $K_y = C$ for $y = 0$, then we have set up the original problem $(\mathcal{P})$ in a family of problems (q_y) , parametrized by $y \in Y$. In convex analysis, this is the standard situation for obtaining a duality result ; this is what we get here too. A new notation is needed ; for any $(x,x^*,a) \in K_{y=y_0}$, the (non-empty) set of $y^* \in Y^*$ such that $(x,y_0;x^*,y^*;a) \in K$ is denoted by $L_{y_0}(x,x^*,a)$.

PROPOSITION 2. The extremization problems :

$(\mathcal{P})$ $\quad$ ext $K_{y=0}$

and

$(\mathcal{P}^*)$ $\quad$ $\text{ext}(\mathcal{L}_{X \times Y} K)_{x^*=0}$

are dual to each other in the following way :

$$(\mathcal{P}^{**}) = (\mathcal{P})$$

$$(x,0,a) \in \{\text{ext}(\mathcal{P})\} \Longrightarrow L_0(x,0,a) \times \{0\} \times \{-a\} \subset \{\text{ext}(\mathcal{P}^*)\}.$$

Proof : Problem $(\mathcal{P}^{**})$ is stated as

$(\mathcal{P}^{**})$ $\quad$ $\text{ext}(\mathcal{L}_{X^* \times Y^*}\, \mathcal{L}_{X \times Y} K)_{y=0}$,

which boils down to $(\mathcal{P})$ since the two Legendre transforms are inverse to each other.

Consider a point $(x,0,a)$ which is extremal in $K_{y=0}$, and any $y^* \in L_{y_0}(x,0,a)$. By definition of these sets, we have :

$$(x,0;0,y^*;a) \in K .$$

Applying the Legendre transform, we get an equivalent statement :

$$(0,y^*;x,0;\ \langle x,0\rangle + \langle 0,y^*\rangle - a) \in \mathcal{L}_{X \times Y} K$$

$$(0,y^*;x,0;-a) \in \mathcal{L}_{X \times Y} K .$$

$$(y^*,0,-a) \in (\mathcal{L}_{X \times Y} K)_{x^*=0}$$

The latter means of course that $(y^*,0,-a)$ solves problem $(\mathcal{P}^*)$ □ .

Here we have taken the Legendre transform $\mathcal{L}_{X\times Y}$ with respect to both variables x and y . We can also take the Legendre transform $\mathcal{L}_Y$ with respect to the variable y only ; this leads us to the set $\mathcal{L}_Y K$ (which should really be $(Id_X \times \mathcal{L}_Y)K$) defined by :

$$\mathcal{L}_Y K = \{(x,y^*;x^*,y;\ \langle y,y^*\rangle - a) \mid (x,y;x^*,y^*;a) \in K\}.$$

In this way, we get a statement which encompasses both $(\mathcal{P})$ and $(\mathcal{P}^*)$:

PROPOSITION 3. Consider the extremization problem :

$(\mathcal{P}')$ $\qquad\qquad$ ext $\mathcal{L}_Y K$.

Then $(x,y^*;0,0;a)$ solves problem $(\mathcal{P}')$ if and only if $(x,0,-a)$ solves problem $(\mathcal{P})$ and $y^* \in L_0(x,0,-a)$ (which implies that $(y^*,0,a)$ solves problem $(\mathcal{P}^*)$) .

Proof : Suppose $(x,y^*;0,0;a) \in \mathcal{L}_Y K$. Going back to the definition of $\mathcal{L}_Y K$, this means that :

$$(x,0;0,y^*;-a) \in K$$

and the result follows immediately □ .

§.II – INTERPRETATION AND EXAMPLES.

We are quite used to associating subsets of $X \times \mathbb{R}$ with functions from X to $\mathbb{R}$. For instance, if f is a function from X to $\mathbb{R}$, its graph

$$\{(x,a) \mid x \in X\ ,\ a = f(x)\} = \text{graph } f$$

and its epigraph

$$\{(x,a) \mid x \in X\ ,\ a \geqslant f(x)\} = \text{epi } f$$

are subsets of $X \times \mathbb{R}$. Conversely, conditions can be stated for a subset of $X \times \mathbb{R}$ to be the graph (or the epigraph) of some function $f : X \to \mathbb{R}$.

If the function f is locally Lipschitzian, it has at every point $x \in X$ a generalized derivative $\partial f(x)$ (see Clarke [7] , [8]) , which is a non-empty, closed, convex, subset of X^* . This enables us to associate with f a subset of $X \times X^* \times \mathbb{R}$ its hypergraph ,

$$\{(x,x^*,a) \mid x \in X\ ,\ x^* \in \partial f(x)\ ,\ a = f(x)\} = \text{hyper } f\ .$$

In the particular case where f is C^p (p times continuously differentiable), its hypergraph is described by the equations :

$$x^* = f'(x) \ , \ a = f(x)$$

and it is a closed submanifold of $X \times X^* \times \mathbb{R}$, C^{p-1}-diffeomorphic to X .

It is to this hypergraph that we now apply the results of the preceding section. In this context, definition 1 becomes clear. The extremization problem

$$\text{ext hyper } f$$

simply consists in solving the equation $\partial f(x) \ni 0$ (or $f'(x) = 0$) , i.e. in looking for stationary points of f in X .

Let us try to figure out what happens to the Legendre transform. From the definitions, it follows that :

$$\mathcal{L}\,(\text{hyper } f) = \{(x^*, x, \langle x, x^*\rangle - f(x)) \mid x \in X \ , \ x^* \in \partial f(x)\}$$

and the question immediately arises : is that the hypergraph of some function $g : X^* \to \mathbb{R}$? If so, g will legitimately be denoted by $\mathcal{L}g$, and called the Legendre transform of f . Here are two important cases .

PROPOSITION 4.

Assume $f : X \to \mathbb{R}$ is convex and continuous. Define $f^* : X \times \mathbb{R} \cup \{+\infty\}$ by

$$f^*(x^*) = \sup \{ \langle x, x^* \rangle - f(x) \mid x \in X\}$$

and set :

$$\text{hyper } f^* = \{(x^*, \partial f^*(x^*), f^*(x^*)) \mid f^*(x^*) < +\infty\}$$

where $\partial f^*(x^*)$ denotes the subgradient of f^* at x (which coincides with the generalized derivative wherever f^* is locally Lipschitzian, i.e. on the interior of dom f^*). Then

$$\mathcal{L}\,(\text{hyper } f) = \text{hyper } f^* \ .$$

Proof : This follows from classical results in convex analysis. If $(x^*, x, a) \in \mathcal{L}$ (hyper f), then $x^* \ \partial f(x)$ and $a = \langle x, x^* \rangle - f(x)$. This implies that $x \in \partial f(x^*)$ and $a = f^*(x^*)$, so that $(x^*, x, a) \in$ hyper f^* . Conversely, if $(x^*; x, a) \in$ hyper f^*, then $x \in \partial f^*(x^*)$ and $a = f^*(x^*)$. This implies that $x^* \in \partial f(x)$ and $a = \langle x, x^* \rangle - f(x)$, so that $(x^*, x, a) \in \mathcal{L}$ (hyper f) $\square$.

PROPOSITION 5. Assume $f : X \to \mathbb{R}$ is C^2, and the linear operator $f''(x): X \to X^*$ is always invertible, with $\|f''(x)^{-1}\|$ uniformly bounded. If $f' : X \to X^*$ is proper (*), then it is a C^1-diffeomorphism, so that the equations :

$$g(x^*) = \langle x, f'(x)\rangle - f(x) \text{ , with } f'(x) = x^*$$

define a C^1 function $g : X^* \to \mathbb{R}$. Moreover :

$$\mathcal{L}(\text{hyper } f) = \text{hyper } g \text{ .}$$

Proof : Let us agree for the time being that f' is a C^1 diffeomorphism. Then g is defined by :

$$g(x^*) = \langle (f')^{-1}(x^*) , x^* \rangle - f\circ(f')^{-1}(x^*)$$

and is clearly well-defined and C^1. Writing $g(x^*) = \langle x, x^*\rangle - f(x)$ and differentiating this expression with respect to x^*, we get :

$$\langle g'(x^*), \xi^*\rangle = \langle x, \xi^*\rangle + \langle f''(x)^{-1}\xi^*, x^* - f'(x)\rangle \text{ .}$$

Taking into account the relation $x^* = f'(x)$, the last term vahishes, and we are left with $g'(x^*) = x$. The hypergraph of g is exactly $\mathcal{L}$ (hyper f).

We now prove the assertion about f'. By the implicit function theorem, since $f''(x)$ is invertible, f' is a C^1-diffeomorphism of some neighbourhood of x onto some neighbourhood of $f'(x)$. Moreover, I claim $f' : X \to X^*$ is surjective, so that it is a covering. Since X^* is simply connected, f' must be one-to-one, and a global diffeomorphism .

To see that f' is onto, take any $x^* \in X^*$, and consider the function $h : X \to \mathbb{R}$ defined by $h(x) = f(x) - \langle x, x^*\rangle$. Clearly, $h''(x) = f''(x)$, so that there is some $c > 0$ with :

$$\| h''(x)^{-1} \| \leq c \quad \text{, all } x \in X \text{ .}$$

I claim that there is some $\bar{x} \in X$ with $h'(\bar{x}) = 0$.

To see this, consider the function $\Psi(x) = \|g'(x)\|^*$, which is continuous and bounded from below, and apply [1], corollary 2.3. : there exists a sequence (x_n) such that, for every $n \in \mathbb{N}^*$ (see [2]) :

$$\forall\, x \in X \text{ , } \|h'(x)\|^* - \|h'(x_n)\|^* + \frac{1}{n}\|x - x_n\| \geq 0 \text{ .}$$

(*) i.e., the inverse image of a compact set must be compact.

By the implicit function theorem, the equation $h'(x_n^t) = (1-t)h'(x_n)$ defines a C^1 curve $t \mapsto x_n^t$ starting at $x_n^o = x_n$ and extending over some time interval $[0,\eta_n)$, with $\eta_n > 0$. The previous inequality then becomes :

$$- t \; \|h'(x_n)\|^* + \frac{1}{n} \|x_n^t - x_n\| \geq 0 \; .$$

Dividing by t, and letting $t \to 0$, we get :

$$- \|h'(x_n)\|^* + \frac{1}{n} \left\| \frac{d}{dt} x_n^t \Big|_{t=0} \right\| \geq 0$$

$$- \|h'(x_n)\|^* + \frac{1}{n} \| - h''(x_n)^{-1} h'(x_n)\| \geq 0 \; .$$

If $h'(x_n) \neq 0$, then $y_n^* = h'(x_n) \,/\, \|h'(x_n)\|^*$ is a unit vector, and the preceding inequality yields :

$$\|h''(x_n)^{-1} y_n^*\| \geq n$$

which contradicts the assumption that the $\|f''(x_n)^{-1}\|$ are uniformly bounded (remember that $f'' = h''$) .

Finally, we have the important special case when the function f splits as a sum $\sum f_i(x_i)$, the x_i being independent variables.

PROPOSITION 6. Assume the X_i, $1 \leq i \leq n$, are Banach spaces, and the $f_i : X_i \to \mathbb{R}$ are continuous functions.

Define $X = \prod_{i=1}^{n} X_i$ and $f : X \to \mathbb{R}$ by $f(x) = \sum_{i=1}^{n} f_i(x_i)$.

If, for each i, either prop. 4 or prop.5 applies to $f_i : X_i \to \mathbb{R}$, so that all the f_i have a well defined Legendre transform $\mathcal{L} f_i : X_i^* \to \mathbb{R}$, then so does f, and

$$\mathcal{L} f(x^*) = \sum_{i=1}^{n} \mathcal{L} f_i(x_i^*) \; .$$

Proof : Call K_i the hypergraph of f_i, and K the hypergraph of f. Clearly :

$$K = \{(x_1,\ldots,x_n; x_1^*,\ldots,x_n^* ; \sum_{i=1}^{n} a_i) \mid (x_i, x_i^*, a_i) \in K_i\}$$

so that :

$$\mathcal{L} K = \{(x_1^*,\ldots,x_n^* ; x_1,\ldots,x_n ; \sum_{i=1}^{n} a_i - \sum_{i=1}^{n} < x_i, x_i^* > \mid (x_i, x_i^*, a_i) \in K_i\}$$

$$= \{(x_1^*,\dots,x_n^* \; ; \; x_1,\dots,x_n \; ; \; \sum_{i=1}^n a_i^* \;\; | (x_i^* , x_i, a_i^*) \in \mathcal{L} K_i \} .$$

The latter means that $\mathcal{L} K$ is the hypergraph of the function $\sum_{i=1}^n \cdot \mathcal{L} f_i : \Pi X_i^* \to \mathbb{R}$. This is the desired result $\square$.

For a detailed study of the case when the Legendre transform $\mathcal{L}f$ cannnot be interpreted as a function, the reader is referred to [3] .

§.3. APPLICATIONS

We will show how various duality results follow from proposition 2. First, of course, we get the classical convex case. But we also get various non-convex ones, which are brought together for the first time.

A - THE CONVEX CASE.

PROPOSITION 7. Let $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ be two convex continuous functionals, and $A : X \to Y$ a linear continuous mapping between Banach spaces X and Y. Then the optimization problems

$$(\mathcal{P}) \qquad \inf\{f(x) + g(Ax) \mid x \in X\}$$

and

$$(Q) \qquad \sup \{ - g^*(-y^*) - f^*(A^*y^*) \mid y^* \in Y^* \}$$

are dual to each other in the following way : $\bar{x}$ solves $(\mathcal{P})$ if and only if any $\bar{y}^* \in \partial g(A\bar{x})$ solves (Q). If such is the case, we have

$$\min(\mathcal{P}) = \max(Q) .$$

Proof : Define the function $\Phi : X \times Y \to \mathbb{R}$ by :

$$\Phi(x,y) = f(x) + g(Ax-y).$$

Clearly Φ is convex and continuous. Problem $(\mathcal{P})$ can be written as :

$$(\mathcal{P}) \qquad \text{ext } \Phi(\cdot,0)$$

We then apply proposition 2 with K the hypergraph of Φ . By proposition 5, the set $\mathcal{L} K$ is simply the hypergraph of Φ^* , so that the dual problem becomes

$$(\mathcal{P}^*) \qquad \text{ext } \Phi^* (0,\cdot)$$

Computing $\Phi^*(x^*,y^*)$, we get :

$$\Phi^*(x^*,y^*)= \sup_{x,y} \{ <x,x^*> + <y,y^*> - f(x) - g(Ax-y)\}$$

Setting $Ax-y = z \in Y$, we get :

$$\Phi^*(x^*,y^*) = \sup_{x,z} \{ <x,x^* + A^*y^*> + <z,-y^*> - f(x) - g(z)\}$$

$$= f^*(x^* + A^*y^*) + g^*(-y^*) .$$

Proposition 2 now tells us that $\bar{x}$ solves problem $(\mathcal{P})$ (yielding $f(\bar{x})+g(A\bar{x})=a$) if and only if any $\bar{y}^*$ such that $(0,\bar{y}^*) \in \partial\Phi(\bar{x},0)$ solves problem $(\mathcal{P}^*)$ (yielding $g^*(-\bar{y}^*) + f^*(A^*\bar{y}^*) = - a$). This is the desired result $\square$.

B - TOLAND'S CASE

PROPOSITION 7. Let $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ be two convex continuous functionals, and $A : X \to Y$ a linear continuous mapping between Banach spaces X and Y. Then the extremization problems

$(\mathcal{P})$ $\quad\quad \text{ext}\{f(x) - g(Ax) \mid x \in X\}$

and

$(\mathcal{Q})$ $\quad\quad \text{ext}\{g^*(y^*) - f^*(A^*y^*) \mid y^* \in Y^*\}$

are dual to each other in the following sense : $\bar{x}$ solves $(\mathcal{P})$ if and only if any $\bar{y}^* \in \partial g(A\bar{x})$ solves $(\mathcal{Q})$. If such is the case, we have :

$$f(\bar{x}) - g(A\bar{x}) = g^*(\bar{y}^*) - f^*(A^*\bar{y}^*)$$

Proof : Define the function $\Phi : X \times Y \to \mathbb{R}$ by :

$$\Phi(x,y) = f(x) - g(Ax-y) .$$

Clearly, $(\mathcal{P})$ can be written as :

$(\mathcal{P})$ $\quad\quad \text{ext}\ \Phi(x,0)$.

We then apply proposition 2 with K the hypergraph of Φ . We now compute its Legendre transform

$$\mathcal{L}K = \{(x^*,y^* ; x,y; <x,x^*> + <y,y^*> - f(x) + g(Ax-y)) \mid x^* \in \partial f(x)-A^* y^* , y^* \in \partial g(Ax-y)\}$$

Setting $Ax-y = z$ this becomes :

$$\mathcal{L}K = \{(x^*,y^*;\ x, Ax-z\ ;\ \langle x, x^*\rangle + \langle Ax-z, y^*\rangle - f(x) + g(z)\ )$$

$$|\,x^* + A^*y^* \in \partial f(x),\ y^* \in \partial g(z)\}\quad .$$

The set $\mathcal{L}K$ is easily seen to be the hypergraph of the Legendre transform of the function

$$(x^*,y^*) \mapsto \mathcal{L}f(x^* + A^*y^*) + \mathcal{L}(-g)(-y^*)\ .$$

By proposition 5, we have $\mathcal{L}f = f^*$. We readily compute $\mathcal{L}(-g)$:

$$a^* \in \mathcal{L}(-g)(z^*) \Leftrightarrow \exists\, y \in X : (y, z^*, \langle y, z^*\rangle - a^*) \in \text{hyper}\ (-g)$$

$$\Leftrightarrow \exists\, y \in X : (y, -z^*, a^* - \langle y, z^*\rangle) \in \text{hyper}\ (g)$$

$$\Leftrightarrow \exists\, y \in X : (-z^*, y, -a^*) \in \text{hyper}\ (\mathcal{L}g)$$

$$\Leftrightarrow a^* = -\ \mathcal{L}g(-z^*) = -\ g^*(-z^*)\ .$$

Finally we get the dual problem :

$(\mathcal{P}^*)$ $\qquad\qquad \text{ext}\ \mathcal{L}\Phi(0,y^*)$

$(\mathcal{P}^*)$ $\qquad\qquad \underset{y^*}{\text{ext}}\ \{f^*(A^*y^*) - g^*(y^*)\}\ .$

Proposition 2 now tells us that $\bar{x}$ solves problem $(\mathcal{P})$ (yielding $f(\bar{x})-g(A\bar{x})=a$) if and only if any $y^* \in \partial g(A\bar{x})$ solves $(\mathcal{P}^*)$ (yielding $f^*(A^*y^*)-g^*(y^*) = -a$) changing signs from $(\mathcal{P}^*)$ to (q) , we get the desired result $\square$.

C. CALCULUS OF VARIATIONS.

Let us consider a classical (Bolza) problem in the calculus of variations :

$$(\mathcal{P})\quad \begin{cases} \text{ext} \displaystyle\int_0^T f(x(t), \dot{x}(t))\,dt \\ x(0) = \xi_0,\ x(T) = \xi_1\ . \end{cases}$$

The function $(\xi,\eta) \mapsto f(\xi,\eta)$ is assumed to have a well-defined Legendre transform $\mathcal{L}f$; i.e. some of the results of §.II apply to $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.

A solution x of problem $(\mathcal{P})$ is an absolutely continuous curve $t \mapsto x(t)$, satisfying the boundary conditions, and such that there exists another absolutely continuous curve $t \mapsto p(t)$ with :

$$(x(t),\dot{x}(t)\ ;\ \dot{p}(t),\ p(t)\ ;\ f(x(t),\ \dot{x}(t))) \in \text{hyper}\ (f) \quad \text{a.e.}$$

If for instance f is C^1, this condition boils down to $\frac{d}{dt} f'_\eta(x(t),\dot{x}(t)) = f'_\xi(\dot{x}(t),x(t))$, the usual Euler-Lagrange equation. We now have the result :

PROPOSITION 8. The problem

$$(Q) \qquad \text{ext} \left\{ \int_0^T - \mathcal{L} f(\dot{p}(t),p(t))dt + p(T)\xi_1 - p(0)\ \xi_0 \right\}$$

is dual to $(\mathcal{T})$ in the following sense : $\bar{x}$ is a solution to $(\mathcal{T})$ if and only if any $\bar{p}$ satisfying hyper $(f) \ni (\bar{x}(t), \frac{d\bar{x}}{dt}(t),\ \bar{p}(t)\ ,\ \frac{d}{dt}\bar{p}(t)$ $f(\bar{x}(t), \frac{d\bar{x}}{dt}(t)))$ for almost every $t \in [0,T]$ is a solution to (Q). Moreover, $\bar{x}$ and $\bar{p}$ assign the same value to $(\mathcal{T})$ and (Q).

Proof. The equation relating $\bar{x}$ and $\bar{p}$ can be written as :

$$\left(\frac{d\bar{p}}{dt}\ ,\ \bar{p}\ ;\ \bar{x}\ ,\ \frac{d\bar{x}}{dt}\ ;\ \bar{x}\frac{d\bar{p}}{dt} + \bar{p}\frac{d\bar{x}}{dt} - f(\bar{x}, \frac{d\bar{x}}{dt}) \right) \in \text{hyper}\ (f)$$

(to be understood for almost every t). But this simply means that $\bar{p}$ is an extremal of (Q). The value is computed to be :

$$\int_0^T - \mathcal{L}f\left(\frac{d\bar{p}}{dt}, \bar{p}\right)dt + p(T)\ \xi_1 - p(0)\ \xi_0 =$$

$$= \int_0^T f(\bar{x}, \frac{d\bar{x}}{dt})dt - \int_0^T \left\{ \bar{x}\frac{d\bar{p}}{dt} + \bar{p}\frac{d\bar{x}}{dt} \right\} dt + (p(T)\ \xi_1 - p(0)\xi_0)$$

The last two terms cancel, yielding the desired result.

D. CONCLUSION.

Of course, this is but a very rough outline of what can be done. The reader is refered to the original papers for the three preceding cases ([6] , [2] for A ; [4] , [5] for B ; [3] for C).

Let us also note that proposition 3 yields the Lagrangian formulation. In the convec case (A), problem $(\mathcal{T})$ and its dual (Q) are equivalent to the single problem :

$$\underset{x,y^*}{\text{ext}} \left\{ f(x) - \langle Ax,y^* \rangle - g^*(-y^*) \right\} .$$

which boils down to the classical saddle-point problem :

$$\inf_x \sup_{y^*} \{f(x) - \langle Ax,y^* \rangle - g^* (-y^*)\} .$$

In Toland's case (B) , problem $(\mathcal{T})$ and its dual $(\mathcal{Q})$ are equivalent to the simple problem :

$$\operatorname{ext}_{x,y^*} \{f(x) + \langle Ax,y^* \rangle + g^*(y^*)\} .$$

Finally, we can easily get a marginal interpretation of the dual variables. Indeed, consider the subset $V \subset Y \times Y^* \times \mathbb{R}$ defined by :

$$V = \{(y,y^*,a) \mid \exists x \in X : (x,y;0,y^*;a) \in \text{hyper } (\Phi) \} .$$

If V is a submanifold, it is easily checked that $da = \langle y^*, dy \rangle$. This can be interpreted as $y^* = \frac{\partial a}{\partial y}$, where $a(y)$ is a branch of extremal values for

$$(\mathcal{T}_y) \qquad \operatorname{ext}_x \Phi(x,y) .$$

BIBLIOGRAPHY

[1] I. EKELAND : "On the variational principle". J. Math. An. Appl. 47 (1974) p.324-353.

[2] I. EKELAND and R. TEMAM : "Convex analysis and variational problems". North-Holland - Elsevier, 1976.

[3] I. EKELAND : "Duality in non-convex optimization and the calculus of variations". SIAM J. Opt. Control, 15 (1977), p.905-934.

[4] J.F. TOLAND : "Duality in non-convex optimisation", University of Essex ; Fluid Mechanics Research Institute, Report n° 78, December 76.

[5] J.F. TOLAND : "On the stability of rotating heavy chains". University of Essex, Fluid Mechanics Research Institute, Report N° 82, May 77.

[6] R. T. ROCKAFELLAR : "Convex analysis", Princeton University Press, 1970.

[7] F. CLARKE : "Generalized gradients and applications", Transactions AMS 205, 1975, p.247-262.

[8] F. CLARKE : "Generalized gradients of Lipschitz functionals", Mathematics of Operations Research, to appear.

Centre de Recherche de mathématiques (CEREMADE) - Université Paris Dauphine
Place de Lattre de Tassigny
75775 PARIS - CEDEX 16.