SIAM J. CONTROL AND OPTIMIZATION
Vol. 15, No. 6, November 1977

LEGENDRE DUALITY IN NONCONVEX OPTIMIZATION AND CALCULUS OF VARIATIONS*

IVAR EKELAND†

Abstract. A general duality theory is given for smooth nonconvex optimization problems, covering both the finite-dimensional case and the calculus of variations. The results are quite similar to the convex case; in particular, with every problem $(\mathcal{P})$ is associated a dual problem $(\mathcal{P}^*)$ having opposite value. This is done at the expense of broadening the framework from smooth functions $\mathbb{R}^n \to \mathbb{R}$ to Lagrangian submanifolds of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$.

Introduction. Duality methods are nowadays an important tool in the study of convex optimization problems. A systematic treatment within the framework of convex analysis can be found in the books of R. T. Rockafellar [14] and I. Ekeland and R. Temam [8]. However, it is easily forgotten that duality methods have been in use for quite a long time in classical mechanics, where people are used to stating a problem either in terms of x-phase variables, or of p-momentum variables, the mapping $x \to p$ being the Legendre transformation. A major difficulty lies in the fact that the Legendre transformation need not be one-to-one, except of course in the convex case.

This paper aims to provide people used to convex optimization problems with a systematic and updated treatment of duality theory for the smooth nonconvex case. The first two sections set up the general framework. It turns out that the framework of functions is not broad enough to cover our needs, because the Legendre transform of a smooth nonconvex function need not be a function. So we define Lagrangian submanifolds of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ as a better concept to work with, because the Legendre transform of a Lagrangian submanifold is still a Lagrangian submanifold, and because a Lagrangian submanifold comes very close to being a function from $\mathbb{R}^n$ to $\mathbb{R}$. Section 1 investigates the local properties of Lagrangian submanifolds, and § 2 studies the Legendre transform in this framework.

The duality theorems then follow quite easily, either in § 3 for the finite-dimensional case, or in § 4 for the calculus of variations. They are exactly what one would expect from the convex case. References to the bibliography are relegated to § 5.

1. Lagrangian submanifolds. Let f be a C^∞ real-valued function on $\mathbb{R}^n$. We can associate with f the following n-dimensional submanifold of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$:

$$V_f = \{(x, f'(x), f(x)) | x \in \mathbb{R}^n\}. \tag{1.1}$$

* Received by the editors September 10, 1976.

† Mathematics Research Center, University of Wisconsin—Madison, Madison, Wisconsin. Now at Centre de Recherche de Mathématiques de la Division, Université de Paris IX Dauphine, Paris, France. This work was supported by the United States Army under Contract No. DAAG29-75-C-0024.

This submanifold has the property of annihilating the differential form ω defined at any point (x, p, z) of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ by the formula

$$\omega = dz - \sum_{i=1}^{n} p_i \, dx_i. \tag{1.2}$$

Indeed, the restriction of ω to V_f reduces to $df - \sum_{i=1}^{n} (\partial f/\partial x_i) \, dx_i$ which is identically zero. This motivates the following definitions.

DEFINITION 1.1. A *Lagrangian submanifold* of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ is a closed n-dimensional C^∞-submanifold V such that

$$i_V^* \omega = 0 \tag{1.3}$$

where i_V: $V \to \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ is the canonical injection and i_V^*: $T^*(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}) \to T^* V$ the induced map of differential 1-forms. We shall say that $\bar{x} \in \mathbb{R}^n$ is a *critical point* of V and that $\bar{z} \in \mathbb{R}$ is a *critical value* whenever

$$(\bar{x}, 0, \bar{z}) \in V. \tag{1.4}$$

We shall associate with V a multivalued mapping F_V from $\mathbb{R}^n$ to $\mathbb{R}$:

$$F_V(x) = \{z \mid \exists \, p \in \mathbb{R}^n : (x, p, z) \in V\} \tag{1.5}$$

and call it the *characteristic map* of V.

In the following, we shall denote by π_x and π_{xz} respectively the restriction to V of the projections $(x, p, z) \to x$ and $(x, p, z) \to (x, z)$. The analogous notations π_p and π_{pz} will also be used. These maps send V into $\mathbb{R}^n$ and $\mathbb{R}^{n+1}$ respectively; note that:

$$\operatorname{graph} F_V = \pi_{xz}(V). \tag{1.6}$$

Particularly simple situations arise when these projections are proper. Recall that a continuous map π: $V \to \mathbb{R}^k$ is proper at $\xi \in \mathbb{R}^k$ iff every sequence ω_n in V such that $\pi(\omega_n) \to \xi$ is bounded. It is proper iff it is proper at every point $\xi \in \mathbb{R}^k$; this amounts to saying that $\pi^{-1}(K)$ is compact in V whenever K is compact in $\mathbb{R}^k$.

As a fundamental example of a Lagrangian submanifold, take the set V_f associated with a C^∞ function f: $\mathbb{R}^n \to \mathbb{R}$ by formula (1.1). Note that in this case π_x is a diffeomorphism from V on $\mathbb{R}^n$, and hence proper.

As a variant, consider a C^∞ function f defined on an open subset Ω of $\mathbb{R}^n$, and assume that $|f(x)| \to \infty$ whenever x converges to some point in the boundary of Ω. Then the set V_f defined by

$$V_f = \{(x, f'(x), f(x)) \mid x \in \Omega\} \tag{1.7}$$

is a Lagrangian submanifold. Note that in this case π_x is a diffeomorphism from V on Ω, but no longer on $\mathbb{R}^n$. Hence π_x is no longer proper, but π_{xz} is.

In both cases, the critical points/values of V_f are the critical points/values of f, and the characteristic map F_V of V_f coincides with f:

$$\forall x \in \mathbb{R}^n, \qquad F_V(x) = \{f(x)\}. \tag{1.8}$$

We now seek a partial converse: describe, at least locally, a given Lagrangian submanifold V, in terms of a smooth function f: $\mathbb{R}^n \to \mathbb{R}$. For that purpose, we introduce the set $\mathscr{R}$ of points $\bar{x} \in \mathbb{R}^n$ such that the 1-forms $i_V^* dx_1, \cdots, i_V^* dx_n$ are linearly independent at every point $(\bar{x}, \bar{p}, \bar{z})$ of V projecting on $\bar{x}$.

PROPOSITION 1.2. *The subset $\mathbb{R}^n \backslash \mathscr{R}$ has Lebesgue measure zero in $\mathbb{R}^n$. For every point $\bar{x} \in \mathscr{R}$ there exist a (possibly empty) countable set of indices A, a family $\mathscr{U}_\alpha$, $\alpha \in A$, of neighborhoods of $\bar{x}$ in $\mathbb{R}^n$, a family $f_\alpha: \mathscr{U}_\alpha \to \mathbb{R}$ of smooth functions, such that*

$$\pi_x^{-1}(\bar{x}) \subset \bigcup_{\alpha \in A} \mathscr{V}_\alpha \subset V \tag{1.9}$$

where

$$\mathscr{V}_\alpha = \{(x, f'_\alpha(x), f_\alpha(x)) | x \in \mathscr{U}_\alpha\}. \tag{1.10}$$

Note that (1.9) implies that $F_V(\bar{x}) = \{f_\alpha(\bar{x}) | \alpha \in A\}$. Intuitively, the part of F_V lying above $\bar{x}$ is decomposed into smooth branches f_α, $\alpha \in A$, with $z_\alpha = f_\alpha(\bar{x})$ and $p_\alpha = f'_\alpha(\bar{x})$. Two branches may intersect, but they must do so transversally: if $f_\alpha(\bar{x}) = f_\beta(\bar{x})$ with $\alpha \neq \beta$, then $f'_\alpha(\bar{x}) \neq f'_\beta(\bar{x})$.

Proof of Proposition 1.2. To say that the 1-forms $i_V^* dx_1, \cdots, i_V^* dx_n$ are linearly independent at $(\bar{x}, \bar{p}, \bar{z}) \in V$ means that $(\bar{x}, \bar{p}, \bar{z})$ is a regular point for the projection $\pi_x: V \to \mathbb{R}^n$. The set $\mathbb{R}^n \backslash \mathscr{R}$ is just the set of critical values for π_x, and it follows from Sard's theorem that it has measure zero.

Take $\bar{x} \in \mathscr{R}$, and let $\{(\bar{x}, \bar{p}_\alpha, \bar{z}_\alpha) | \alpha \in A\}$ be the (possibly empty) set of points of V projecting on $\bar{x}$. By the definition of $\mathscr{R}$, each $(\bar{x}, \bar{p}_\alpha, \bar{z}_\alpha)$, $\alpha \in A$, is a regular point for π_x. By the implicit function theorem, there are neighborhoods $\mathscr{U}_\alpha$ of $\bar{x}$ and $\mathscr{V}_\alpha$ of $(\bar{x}, \bar{p}_\alpha, \bar{z}_\alpha)$ such that $\pi_x: \mathscr{V}_\alpha \to \mathscr{U}_\alpha$ is a diffeomorphism. In other words, there are real-valued C^∞ functions f_α and $g_{\alpha i}$, $1 \leqq i \leqq n$, defined over $\mathscr{U}_\alpha$, such that

$$(x, p, z) \in \mathscr{V}_\alpha \Leftrightarrow \{x \in \mathscr{U}_\alpha, z = f_\alpha(x), p_i = g_{\alpha i}(x)\}. \tag{1.11}$$

The vanishing of $i_V^* \omega$ means that

$$df_\alpha - \sum_{i=1}^n g_{\alpha i}(x)\, dx_i = 0 \quad \text{over } \mathscr{U}_\alpha, \tag{1.12}$$

which yields

$$g_{\alpha i}(x) = \frac{\partial f_\alpha}{\partial x_i}(x) \quad \forall x \in \mathscr{U}_\alpha. \tag{1.13}$$

Writing (1.13) in (1.11), we get formula (1.10), with formula (1.9) being satisfied by construction. It only remains to prove that the set A is at most countable. For this, notice that

$$\pi^{-1}(\bar{x}) \cap \mathscr{V}_\alpha = \{(\bar{x}, \bar{p}_\alpha, \bar{z}_\alpha)\} \tag{1.14}$$

and hence that $\alpha \neq \beta \Rightarrow (\bar{x}, \bar{p}_\beta, \bar{z}_\beta) \notin \mathscr{V}_\alpha$. This shows that all points in $\pi^{-1}(\bar{x})$are isolated; hence any compact subset of V can contain only a finite number of them. As V is a closed subset of $\mathbb{R}^{2n+1}$, it can be written as a countable union of compact subsets, and the result follows. □

In the special case where the map π_x is proper at $\bar{x}$, it is easily seen that the set A has to be finite. Setting $\mathscr{U} = \bigcap_{\alpha \in A} \mathscr{U}_\alpha$, we get the following corollary.

COROLLARY 1.3. *Assume moreover the map π_x is proper. Then $\mathcal{R}$ is open in $\mathbb{R}^n$, and for every point $\bar{x} \in \mathcal{R}$ there is a neighborhood $\mathcal{U}$ of $\bar{x}$ and a (possibly empty) finite family of smooth functions $f_\alpha: \mathcal{U} \to \mathbb{R}$, $\alpha \in A$, such that*

$$\pi_x^{-1}(\mathcal{U}) = \bigcup_{\alpha \in A} \{(x, f'_\alpha(x), f_\alpha(x)) | x \in \mathcal{U}, \alpha \in A\}. \tag{1.15}$$

We now have a description of $\pi_x^{-1}(\bar{x})$ which is valid whenever $\bar{x} \in \mathcal{R}$, i.e. for almost every point $\bar{x} \in \mathbb{R}^n$. Points in $\mathbb{R}^n \backslash \mathcal{R}$ form a negligible subset, but they may nevertheless turn out to be important, so we will attempt a partial description in that case also.

PROPOSITION 1.4. *Let $t \mapsto (x(t), p(t), z(t))$ be a C^1 map from $]0, T]$ into V such that $x(t) \in \mathcal{R}$ $\forall t > 0$. Assume that, when $t \to 0$,*

$$x(t) \to \bar{x} \quad \text{and} \quad \frac{dx}{dt}(t) \to \xi, \tag{1.16}$$

$$z(t) \to \bar{z}, \tag{1.17}$$

$$\liminf \|p(t) - \bar{p}\| = 0, \tag{1.18}$$

with $(\bar{x}, \bar{p}, \bar{z})$ an isolated point of $\pi_{xz}^{-1}(\bar{x}, \bar{z})$. Then

$$p(t) \to \bar{p}, \tag{1.19}$$

$$\frac{dz}{dt}(t) \to \bar{p} \cdot \xi. \tag{1.20}$$

Proof. As $\bar{p}$ is an isolated point in $\pi_{xz}^{-1}(\bar{x}, \bar{z})$, there is a compact neighborhood $\mathcal{W}$ of $(\bar{x}, \bar{p}, \bar{z})$ in V such that

$$(\bar{x}, p, \bar{z}) \in \mathcal{W} \Rightarrow p = \bar{p}. \tag{1.21}$$

Assume $p(t)$ does not converge to $\bar{p}$. Then there is an open neighborhood $\mathcal{V}$ of $(\bar{x}, \bar{p}, \bar{z})$, contained in $\mathcal{W}$, and a sequence $t_n \to 0$ such that

$$(x(t_n), p(t_n), z(t_n)) \in \mathcal{W} \backslash \mathcal{V}. \tag{1.22}$$

Using (1.16) and (1.17), together with the fact that $\mathcal{W} \backslash \mathcal{V}$ is compact, we can extract a subsequence converging to some point

$$(\bar{x}, p', \bar{z}) \in \mathcal{W} \backslash \mathcal{V} \tag{1.23}$$

contradicting (1.21).

So $p(t)$ has to converge to $\bar{p}$, yielding (1.19). Setting $z(0) = \bar{z}$, we define a continuous real-valued function $t \mapsto z(t)$ on $[0, T]$. It follows from Proposition 1.2 and the fact that $x(t) \in \mathcal{R}$ for $t > 0$ that this function is derivable on $]0, T]$ with derivative:

$$\frac{dz}{dt}(t) = p(t) \frac{dx}{dt}(t). \tag{1.24}$$

When $t \to 0$, the right-hand side converges to $\bar{p} \cdot \xi$, and so does the left-hand side. □

Note that $(dp/dt)(t)$ need not converge. Note also that (1.16) and (1.20) imply that $(d^+x/dt)(0)=\xi$ and $(d^+z/dt)(0)=\bar{p}\cdot\xi$, with d^+/dt denoting the right-derivative. Equation (1.20) can be written

$$\frac{d^+z}{dt}(0)=\bar{p}\cdot\frac{d^+x}{dt}(0) \tag{1.25}$$

which expresses the vanishing of $dz-p\,dx$ above a point $\bar{x}$ not in $\mathcal{R}$.

Let us give a more accurate picture in a simple case:

PROPOSITION 1.5. *Assume π_x is proper and $\pi_x^{-1}(\bar{x})$ is finite. Let a simply connected subset Ω of $\mathcal{R}$ be given in the following way:*

$$\Omega=\{\bar{x}+t\xi\,|\,0<t<a,\ \xi\in S\} \tag{1.26}$$

with S an open subset of the unit sphere $\xi_1^2+\cdots+\xi_n^2=1$. There is a (possibly empty) finite family of C^1 functions $f_\alpha:\Omega\cup\{\bar{x}\}\to\mathbb{R}$, $\alpha\in A$, such that

$$\pi^{-1}(\Omega\cup\{\bar{x}\})=\{(x,f'_\alpha(x),f_\alpha(x))\,|\,x\in\Omega\cup\{\bar{x}\},\ \alpha\in A\}. \tag{1.27}$$

By a derivative of f_α at $\bar{x}$ we mean a linear functional $f'_\alpha(\bar{x})$ such that

$$\begin{aligned}&\forall\varepsilon>0,\ \exists\,\eta>0\colon \|x-\bar{x}\|\leqq\eta \quad\text{and}\quad x\in\Omega\\ &\Rightarrow|f_\alpha(x)-f_\alpha(\bar{x})-\langle f'_\alpha(\bar{x}),x-\bar{x}\rangle|\leqq\varepsilon\|x-\bar{x}\|.\end{aligned} \tag{1.28}$$

By a C^1 function on $\Omega\cup\{\bar{x}\}$ we mean a function f_α such that $f'_\alpha(x)$ is well-defined and continuous on $\{\bar{x}\}\cup\Omega$.

Proof of Proposition 1.5. The set $\pi_x^{-1}(x)$ has to be both compact (because π_x is proper) and discrete (because $x\in\mathcal{R}$), so it is finite. By Proposition 1.2, the map $\pi_x\colon\pi_x^{-1}(\Omega)\to\Omega$ is a covering. As Ω is simply connected, the restriction of π_x to each connected component of $\pi_x^{-1}(\Omega)$ is a diffeomorphism, hence the representation formula

$$\pi^{-1}(\Omega)=\{(x,f'_\alpha(x),f_\alpha(x))\,|\,x\in\Omega,\ \alpha\in A\}. \tag{1.29}$$

Now fix $\alpha\in A$ and let x converge to $\bar{x}$ in Ω. As π_x is proper, $(x,f'_\alpha(x),f_\alpha(x))$ has cluster points $(\bar{x},p,z)\in\pi_x^{-1}(\bar{x})$. As this set is finite, all its points are isolated. As in the preceding proof, we conclude that $f'_\alpha(x)\to p_\alpha$ and $f_\alpha(x)\to z_\alpha$. Setting $f_\alpha(\bar{x})=z_\alpha$ and $f'_\alpha(\bar{x})=p_\alpha$, we get a C^1 function as desired. □

Let us conclude this investigation of Lagrangian submanifolds by the following remark, which throws some light on the case where $\pi_x^{-1}(\bar{x})$ is not discrete. Let $t\to(x(t),p(t),z(t))$ be a C^1 path drawn on V along which $x(t)$ is constant: $x(t)=\bar{x}$, $0\leqq t\leqq T$. Then $z(t)$ has to be constant also: $z(t)=\bar{z}$, $0\leqq t\leqq T$, so in fact only $p(t)$ varies. This follows easily from the vanishing of $i_V^*\omega$, which yields in this case $(dz/dt)(t)=\sum_{i=1}^n p_i(dx_i/dt)(t)$. In particular, if $\mathcal{V}$ is an open path-connected subset of V projecting on $\bar{x}$, i.e. $\mathcal{V}\subset\pi_x^{-1}(\bar{x})$, then $\mathcal{V}$ is also contained in some hyperplane $H=\{(x,p,z)\,|\,x=\bar{x},\ z=\bar{z}\}$ as an open path-connected subset (openness follows from the fact that $\dim V=n=\dim H$).

2. The Legendre transformation. The mapping $\mathscr{L}$ of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ into itself defined by

$$\mathscr{L}(x, p, z) = (x', p', z'),\qquad x' = p, \quad p' = x, \quad z' = px - z \tag{2.1}$$

is called the *Legendre transformation.* Note the following.

PROPOSITION 2.1. *The Legendre transformation is a C^∞ involution:*

$$\mathscr{L}^2 = Id. \tag{2.2}$$

Proof. Using notations (2.1), we set $\mathscr{L}(x', p', z') = (x'', p'', z'')$, with

$$\begin{aligned} x'' &= p' = x, \\ p'' &= x' = p, \\ z'' &= p'x' - z' = px - (px - z) = z; \end{aligned}$$

hence we get the result. □

The fundamental fact about the Legendre transformation is that it preserves the 1-form ω, up to a change of sign.

THEOREM 2.2. $\mathscr{L}^*\omega = -\omega$.

Proof. Using notations (2.1), we get

$$\begin{aligned} \mathscr{L}^*\omega &= dz' - p'\,dx' \\ &= (x\,dp + p\,dx - dz) - x\,dp \\ &= p\,dx - dz \\ &= -\omega. \quad \square \end{aligned}$$

COROLLARY 2.3. *If V is a Lagrangian submanifold of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$, then so is $\mathscr{L}V$.*

Proof. It follows from Proposition 2.1 that $\mathscr{L}$ is a diffeomorphism of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ onto itself. Hence $\mathscr{L}V$ is a closed submanifold whenever V is. There only remains to check that $i^*_{\mathscr{L}V}\omega = 0$. To do that, we write the following diagram:

$$\begin{array}{ccc} V & \xrightarrow{\;i\;} & \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \\ {\scriptstyle l}\big\downarrow & & \big\downarrow{\scriptstyle \mathscr{L}} \\ \mathscr{L}V & \xrightarrow{\;j\;} & \mathbb{R}^n \times \mathbb{R} = \mathbb{R} \end{array} \tag{2.3}$$

where l is the restriction of $\mathscr{L}$ to V and j is the canonical injection. This diagram commutes, and gives rise to another commutative diagram relating 1-forms:

$$\begin{array}{ccc} T^*V & \xleftarrow{\;i^*\;} & T^*(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}\ ) \\ {\scriptstyle l^*}\big\uparrow & & \big\uparrow{\scriptstyle \mathscr{L}^*} \\ T^*(\mathscr{L}V) & \xleftarrow{\;j^*\;} & T^*(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}). \end{array} \tag{2.4}$$

Taking ω in the lower right-hand corner, and using formula (1.3) and Theorem 2.2, we get

$$i^* \circ \mathscr{L}^*(\omega) = i^*(-\omega) = -i^*(\omega) = 0; \tag{2.5}$$

going the other way around the diagram, we get

$$0 = l^* \circ j^*(\omega). \tag{2.6}$$

As l is a diffeomorphism, l^* is an isomorphism, and (2.6) implies that $j^*\omega = 0$, i.e. $\mathscr{L}V$ is Lagrangian. □

We now introduce a slight misuse of notations. Let V and W be Lagrangian submanifolds of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$, with $W = \mathscr{L}V$, and let F_V and F_W be the associated characteristic maps. We shall write freely $F_W = \mathscr{L}F_V$, and call F_W the Legendre transform of F_V. For instance, if $f\colon \mathbb{R}^n \to \mathbb{R}$ is a C^∞ function, then $\mathscr{L}f$ is the multivalued map from $\mathbb{R}^n$ to $\mathbb{R}$ defined by

$$\mathscr{L}f(x') = \{z' \mid \exists\, p' \in \mathbb{R}^n\colon (x', p', z') \in \mathscr{L}V_f\}. \tag{2.7}$$

Using (1.1) and (2.1), we get

$$\mathscr{L}f(p) = \{px - f(x) \mid f'(x) = p\}. \tag{2.8}$$

Several remarks are now in order. First of all, if f, in addition to being smooth, is *convex*, then the function $x \mapsto px - f(x)$ is concave, and the equation $p = f'(x)$ simply means that this function attains its maximum at x. Equation (2.8) then becomes

$$\mathscr{L}f(p) = \max\{px - f(x) \mid x \in \mathbb{R}^n\}. \tag{2.9}$$

Formula (2.9) shows that $\mathscr{L}f$ is single- or possibly empty-valued. In other words, $\mathscr{L}f$ is a real-valued function defined on some subset of $\mathbb{R}^n$. It is to be compared with the classical Fenchel transform of convex analysis:

$$f^*(p) = \sup\{px - f(x) \mid x \in \mathbb{R}^n\}. \tag{2.10}$$

Formulas (2.9) and (2.10) coincide whenever the function $x \to px - f(x)$ attains its maximum over $\mathbb{R}^n$. Define the effective domain of f^* as the set of points where it is finite:

$$\operatorname{dom} f^* = \{p \mid f^*(p) < \infty\}. \tag{2.11}$$

PROPOSITION 2.4. *$\mathscr{L}f(p) = f^*(p)$ if and only if f^* is subdifferentiable at p, i.e. $\partial f^*(p) \neq \phi$. This is the case at every interior point p of* $\operatorname{dom} f^*$:

$$p \in \operatorname{int} \operatorname{dom} f^* \Rightarrow \mathscr{L}f(p) = f^*(p). \tag{2.12}$$

Proof. Let us write down the definition of the subdifferential of f^*:

$$\partial f^*(p) = \{\bar{x} \in \mathbb{R}^n \mid p\bar{x} - f^{**}(\bar{x}) = \max_x\} \tag{2.13}$$

where the notation $= \max_x$ means that the left-hand side attains its maximum at $\bar{x}$.

But, as f is continuous and convex, it coincides with its biconjugate f^{**}; hence

$$\partial f^*(p)=\{\bar{x}\in\mathbb{R}^n \mid p\bar{x}-f(\bar{x})=\max_x\} \tag{2.14}$$

which proves the first part of the proposition.

It is a well-known fact from convex analysis that any convex function on $\mathbb{R}^n$ is continuous, and hence subdifferentiable, on the interior of its effective domain. Hence we have (2.12). □

In the general (smooth, nonconvex) case, formula (2.8) sets $\mathscr{L}f(p)$ in one-to-one correspondence with the sets of tangents to f having slope p.

PROPOSITION 2.5. *$z'\in\mathscr{L}f(p)$ if and only if $z=px-z'$ is a tangent hyperplane to graph f in $\mathbb{R}^n\times\mathbb{R}$.*

Proof. The hyperplane $z=px-z'$ in (x,z)-space is tangent to graph f if and only if there exists $\bar{x}\in\mathbb{R}^n$ such that $f'(\bar{x})=p$ and $f(\bar{x})=p\bar{x}-z'$. This reduces to $z'\in\mathscr{L}f(p)$ by (2.8). □

From Proposition 2.5 one sees instantly that $\mathscr{L}f$ can be multivalued. Indeed $\mathscr{L}f$ is a function, i.e. $\mathscr{L}f(p)$ is empty or a singleton for every p, if and only if f has only zero or one tangent of prescribed slope. In dimension $n=1$, this means exactly that f is convex. In higher dimensions, this also happens in the nonconvex case: take for instance $f(x_1,x_2)=x_1^2-x_2^2$; then $f'\colon(x_1,x_2)\mapsto(2x_1,-2x_2)$ is one-to-one. But the fact remains that, in contrast with the convex case, in the general case we have to deal with multivalued Legendre transforms. So let us attempt a description of $\mathscr{L}f$. We denote by V the Lagrangian submanifold (1.1) of $\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}$ associated with f, and by $A(x)$ the matrix of second derivatives of f at x:

$$A(x)=\left(\left(\frac{\partial^2 f}{\partial x_i\,\partial x_j}(x)\right)\right),\qquad 1\leqq i,j\leqq n. \tag{2.15}$$

PROPOSITION 2.6. *Assume $A(\bar{x})$ has full rank n. Then there exists a neighborhood $\mathscr{V}$ of $(f'(\bar{x}),\bar{x},\bar{x}f'(\bar{x})-f(\bar{x}))$ in $\mathscr{L}V$ projecting onto a neighborhood $\mathscr{U}$ of $f'(\bar{x})$ in $\mathbb{R}^n$, and a local inverse φ for f' such that*

$$\mathscr{V}=\{(p,[\mathscr{L}_{\mathscr{V}}f]'(p),[\mathscr{L}_{\mathscr{V}}f](p))\mid p\in\mathscr{U}\} \tag{2.16}$$

with $[\mathscr{L}_{\mathscr{V}}f](p)=p\varphi(p)-f\circ\varphi(p)$. In particular, we have

$$[\mathscr{L}_{\mathscr{V}}f]'(\bar{p})=\bar{x}. \tag{2.17}$$

Proof. It follows from the implicit function theorem that the map $x\mapsto f'(x)$ has a local inverse φ defined on some neighborhood $\mathscr{U}$ of $\bar{p}$. Setting

$$\mathscr{V}=\{(f'(x),x,xf'(x)-f(x))\mid x\in\varphi(\mathscr{U})\} \tag{2.18}$$

and using the definition of φ, we get

$$\mathscr{V}=\{(p,\varphi(p),p\varphi(p)-f\circ\varphi(p))\mid p\in\mathscr{U}\}. \tag{2.19}$$

Computing the derivative of $\mathscr{L}_\mathscr{V} f$, we get

$$[\mathscr{L}_\mathscr{V} f]'(p) = \varphi(p) + {}^t\varphi'(p)p - {}^t\varphi'(p)f' \circ \varphi(p)$$
$$(2.20) \qquad = (p) + {}^t\varphi'(p)p - {}^t\varphi'(p)p$$
$$= \varphi(p)$$

and formula (2.19) reduces to (2.16). □

$\mathscr{L}_\mathscr{V} f$ is a smooth branch of $\mathscr{L}f$ lying above $\bar{p}$. Note that $\bar{p}$ is a regular value for f': $\mathbb{R}^n \to \mathbb{R}^n$ if and only if it is a regular value for $\pi_x: \mathscr{L}V \to \mathbb{R}^n$. This is almost always the case, by Sard's theorem, and the part of $\mathscr{L}f$ lying above $\bar{p}$ then is a countable union of smooth branches such as $\mathscr{L}_\mathscr{V} f$ (this is a particular case of Proposition 1.2). If moreover f' is proper at $\bar{p}$, then so is π_x, and there are only a finite number of branches of $\mathscr{L}f$ lying above $\bar{p}$ (this is a particular case of Corollary 1.3).

We can of course apply Propositions 1.4 and 1.5 to get a description of $\mathscr{L}f$ above critical values of f'. But, in this particular case, we prefer another approach, which has the advantage of directly relating the shape of the Legendre transform above $f(\bar{x})$ to the degeneracy of the matrix of second derivatives at $\bar{x}$. We write the Taylor expansion of f at $\bar{x}$:

$$(2.21) \qquad f(\bar{x}+\xi) = f(\bar{x}) + \bar{p}\xi + \tfrac{1}{2}\langle A(\bar{x})\xi, \xi\rangle + \tfrac{1}{6}P_3(\bar{x}; \xi_1, \cdots, \xi_n) + O(|\xi|^4)$$

where $P_3(\bar{x};\ \cdot\ )$ is a homogeneous polynomial of degree 3 in n variables. Using the Euler formula, we may write

$$P_3(\bar{x}; \xi_1, \cdots, \xi_n) = \tfrac{1}{3}\sum_{i=1}^{n} \xi_i \frac{\partial P_3}{\partial \xi_i}(\bar{x}; \xi_1, \cdots, \xi_n) = \sum_{i=1}^{n} \xi_i \langle B_i(\bar{x})\xi, \xi\rangle$$

where $B_i(\bar{x})$ is the matrix with elements $\frac{1}{3}\partial^3 f/\partial x_i\, \partial x_j\, \partial x_k$, $1 \leqq j, k \leqq n$. Denote by $\langle B(\bar{x})\xi, \xi\rangle$ the n-vector with components $\langle B_i(\bar{x})\xi, \xi\rangle$.

PROPOSITION 2.7. *Assume that $A(\bar{x})$ has rank $(n-1)$ and that*

$$(2.22) \qquad \xi \neq 0, \qquad \xi \in \operatorname{Ker} A(\bar{x}) \Rightarrow \begin{cases} P_3(x; \xi_1, \cdots, \xi_n) \neq 0, \\ \langle B(\bar{x})\xi, \xi\rangle \notin \operatorname{Im} A(\bar{x}). \end{cases}$$

Then (possibly after reordering the linear coordinates $(p_1, \cdots, p_n)$ in $\mathbb{R}^n$ and changing p_n to $-p_n$) there is a neighborhood $\mathscr{V}$ of $(f'(\bar{x}), \bar{x}, f'(\bar{x})\bar{x} - f(\bar{x}))$ in $\mathscr{L}V$, a neighborhood $\mathscr{U} = \mathscr{U}' \times \mathscr{U}_n$ of $(\bar{p}_1, \cdots, \bar{p}_{n-1}, \bar{p}_n)$ in $\mathbb{R}^n$, C^∞ functions k_1, k_2: $\mathscr{U}' \to \mathbb{R}$ and h: $\mathscr{U} \to \mathbb{R}$, such that $\pi_{xz}\mathscr{V}$ is completely described by the set of conditions

$$(2.23) \qquad (p_1, \cdots, p_{n-1}, p_n) \in \mathscr{U}' \times \mathscr{U}_n \quad \text{and} \quad p_n \geqq k_1(p_1, \cdots, p_{n-1}),$$

$$(2.24) \qquad z \in \{z_+(p), z_-(p)\},$$

with

$$z_+(p) = k_2(p_1, \cdots, p_{n-1}) + (p_n - k_1)h(p_1, \cdots, p_{n-1}, \sqrt{p_n - k_1}),$$

$$z_-(p) = k_2(p_1, \cdots, p_{n-1}) + (p_n - k_1)h(p_1, \cdots, p_{n-1}, -\sqrt{p_n - k_1}).$$

Moreover $\partial z/\partial p_i = x_i$, $1 \leqq i \leqq n$, along the hypersurface

$$(2.25) \qquad p_n = k_1(p_1, \cdots, p_{n-1}).$$

Proof. The $(x_1, \cdots, x_n)$ are a system of coordinates in $\mathscr{L}V$ with formula (2.8) yielding $(p_1, \cdots, p_n, z)$ in terms of $(x_1, \cdots, x_n)$. In particular,

$$\frac{\partial f}{\partial x_i}(x) = p_i \quad \text{for } 1 \leqq i \leqq n. \tag{2.26}$$

The rank assumption on the matrix $A(\bar{x})$ implies that one of its $(n-1)\times(n-1)$ minors is invertible, for instance the one defined by the $(n-1)$ first rows and the $(n-1)$ first columns. Moreover, the nth row then is a linear combination of the $(n-1)$ first rows.

It follows from the implicit function theorem that the $(n-1)$ first equations of system (2.26) can be solved locally for $(x_1, \cdots, x_{n-1})$. In other words, $(p_1, \cdots, p_{n-1}, x_n)$ can be used as coordinates in some neighborhood $\mathscr{V}_1$ of $(\bar{p}, \bar{x}, \bar{z})$ in $\mathscr{L}V$.[1] Now consider the path $w(t) = (p(t), x(t), z(t))$ in $\mathscr{V}_1$ such that $p_1(t) = \bar{p}_1, \cdots, p_{n-1}(t) = \bar{p}_{n-1}, x_n(t) = \bar{x}_n + t$. There is some $T > 0$ such that $w(t)$ is well-defined for $-T \leqq t \leqq T$. Obviously $w(0) = (\bar{p}, \bar{x}, \bar{p}\bar{x} - f(\bar{x}))$; we shall write ξ' for $(dx/dt)(0)$ and ξ'' for $(d^2x/dt^2)(0)$. Equations (2.26) are satisfied along $w(t)$:

$$p_i(t) = \frac{\partial f}{\partial x_i}(x_1(t), \cdots, x_n(t)) \quad \text{for } 1 - T \leqq t \leqq T. \tag{2.27}$$

Writing Taylor expansions into (2.27), we get

$$p(t) - \bar{p} = tA(\bar{x})\xi' + \frac{t^2}{2}[\langle B(\bar{x})\xi', \xi'\rangle + A(\bar{x})\xi''] + O(t^3). \tag{2.28}$$

But $p_i(t) - \bar{p}_i = 0$ for $1 \leqq i \leqq n-1$, so that both sides of the $(n-1)$ first equations of system (2.28) are identically zero on $(-T, T)$. It follows that the $(n-1)$ first components of $A(\bar{x})\xi$ are zero, and, by the rank assumption, so is the last one

$$A(\bar{x})\xi' = 0. \tag{2.29}$$

Assumption (2.22) then yields

$$\langle B(\bar{x})\xi', \xi'\rangle + A(\bar{x})\xi'' \neq 0. \tag{2.30}$$

But again, both sides of the $(n-1)$ first equations (2.28) being identically zero on $(-T, T)$, the $(n-1)$ first components of vector (2.30) must be zero. It follows that the nth component must be nonzero. We summarize our results so far by stating that the nth equation of system (2.28) can be written as

$$p_n(t) - \bar{p}_n = \tfrac{1}{2}a_n t^2 + O(t^3), \qquad a_n \neq 0. \tag{2.31}$$

Similarly, we compute the Taylor expansion of $z(t)$ at $t = 0$. By definition, we have

$$z(t) = f'[x(t)]x(t) - f[x(t)]. \tag{2.32}$$

[1] From now on we set $\bar{p} = f'(\bar{x})$ and $\bar{z} = \bar{p}\bar{x} - f(\bar{x})$.

Successive derivations yield

$$\frac{dz}{dt}(0) = \langle A(\bar{x})\xi', \xi' \rangle \tag{2.33}$$

$$\frac{d^2z}{dt^2}(0) = 2\langle A(\bar{x})\xi', \xi'' \rangle + P_3(\bar{x}; \xi_1', \cdots, \xi_n'). \tag{2.34}$$

But we have seen that $A(\bar{x})\xi' = 0$, so that $(dz/dt)(0) = 0$ and $(d^2z/dt^2)(0) = b_n \neq 0$ by assumption (2.22). Finally we get

$$z(t) - \bar{z} = \tfrac{1}{2}b_n t^2 + O(t^3), \qquad b_n \neq 0. \tag{2.35}$$

Now $w'(0)$ is just the tangent vector $(\partial/\partial x_n)(\bar{p}_1, \cdots, \bar{p}_{n-1}, x_n)$ associated with the new coordinate systems. In other words, p_n and z, considered as functions of $(p_1, \cdots, p_{n-1}, x_n)$ in $\mathcal{V}_1$, satisfy

$$\frac{\partial p_n}{\partial x_n}(\bar{p}_1, \cdots, \bar{p}_{n-1}, \bar{x}_n) = 0, \tag{2.36}$$

$$\frac{\partial^2 p_n}{\partial x_n^2}(\bar{p}_1, \cdots, \bar{p}_{n-1}, \bar{x}_n) \neq 0, \tag{2.37}$$

$$\frac{\partial z}{\partial x_n}(\bar{p}_1, \cdots, \bar{p}_{n-1}, \bar{x}_n) = 0, \tag{2.38}$$

$$\frac{\partial^2 z}{\partial x_n^2}(\bar{p}_1, \cdots, \bar{p}_{n-1}, \bar{x}_n) \neq 0. \tag{2.39}$$

But other points (p, x, z) in $\mathcal{V}_1$ enjoy the property that $A(x)$ is of rank $(n-1)$ and satisfies (2.22). Indeed, consider the Jacobian determinant

$$\begin{aligned} \Delta(p_1, \cdots, p_{n-1}, x_n) &= \frac{D(p_1, \cdots, p_{n-1}, p_n)}{D(p_1, \cdots, p_{n-1}, x_n)}^{2} \\ &= \frac{\partial p_n}{\partial x_n}(p_1, \cdots, p_{n-1}, x_n) \end{aligned} \tag{2.40}$$

by a simple computation. Clearly rank $A(x_1, \cdots, x_n) < n$ if and only if $\Delta(p_1, \cdots, p_{n-1}, x_n) = 0$. But $\Delta = 0$ and $(\partial\Delta/\partial x_n) = (\partial^2 p_n/\partial x_n^2) \neq 0$ at point $(\bar{p}_1, \cdots, \bar{p}_{n-1}, \bar{x}_n)$. By the implicit function theorem, there are neighborhoods $\mathcal{U}_1$ of $(\bar{p}_1, \cdots, \bar{p}_{n-1})$ and $\mathcal{W}_1$ of $\bar{x}_n$ and C^∞ map g: $\mathcal{U}_1 \to \mathcal{W}_1$ such that

$$\Delta(p_1, \cdots, p_{n-1}, x_n) = 0 \Leftrightarrow x_n = g(p_1, \cdots, p_{n-1}) \forall (p_1, \cdots, x_n) \in \mathcal{U}_1 \times \mathcal{W}_1. \tag{2.41}$$

Conversely, $x_n = g(p_1, \cdots, p_{n-1})$ implies rank $A(x_1, \cdots, x_n) < n$. By a continuity argument, we can shrink $\mathcal{U}_1$ and $\mathcal{W}_1$ to $\mathcal{U}_2$ and $\mathcal{W}_2$ so that rank $A(x_1, \cdots, x_n)$ is exactly $n-1$ and assumption (2.22) is satisfied whenever $x_n = g(p_1, \cdots, p_{n-1})$ in $\mathcal{U}_2 \times \mathcal{W}_2$. We may even include in the bargain the fact that the first minor of $A(x)$ is invertible, so that $(p_1, \cdots, p_{n-1}, x_n)$ enjoys all the properties of $(\bar{p}_1, \cdots, \bar{p}_{n-1}, \bar{x}_n)$. By (2.36) and (2.39), it follows that $\partial p_n/\partial x_n = 0$, $\partial^2 p_n/\partial x_n^2 \neq$

[2] Recall that $D(f_1, \cdots, f_n)/D(x_1, \cdots, x_n)$ denotes the determinant $\|\partial f_i/\partial x_j\|$.

0, $\partial z/\partial x_n = 0$, $\partial^2 z/\partial x_n^2 \neq 0$ at every point $(p_1, \cdots, p_{n-1}, x_n) \in \mathscr{U}_2 \times \mathscr{W}_2$ such that $x_n = g(p_1, \cdots, p_{n-1})$.

It follows that

(2.42) $p_n = k_1(p_1, \cdots, p_{n-1}) + [x_n - g(p_1, \cdots, p_{n-1})]^2 h_1(p_1, \cdots, p_{n-1}, x_n)$,

(2.43) $z = k_2(p_1, \cdots, p_{n-1}) + [x_n - g(p_1, \cdots, p_{n-1})]^2 h_2(p_1, \cdots, p_{n-1}, x_n)$

with

(2.44) $x_n = g(p_1, \cdots, p_{n-1}) \Rightarrow h_1(p_1, \cdots, p_{n-1}, x_n) h_2(p_1, \cdots, p_{n-1}, x_n) \neq 0$.

The point of V defined by $(p_1, \cdots, p_{n-1}, x_n = g(p_1, \cdots, p_n))$ yields $p_n = k_1(p_1, \cdots, p_{n-1})$ and $z = k_2(p_1, \cdots, p_{n-1})$, so that k_1 and k_2 are C^∞ functions. It follows from the C^∞ division theorem of Malgrange that h_1 and h_2 can be chosen to be C^∞ functions also.

Assume that $h_1(\bar{p}_1, \cdots, \bar{p}_{n-1}, \bar{x}_n) > 0$. Then we can define $y_n = [x_n - g]\sqrt{h_1}$ and use $(p_1, \cdots, p_{n-1}, y_n)$ as a new system of local coordinates in some smaller neighborhood $\mathscr{V}_2$ of $(\bar{p}, \bar{x}, \bar{z})$ corresponding to $(p_1, \cdots, p_{n-1}, y_n) \in \mathscr{U}_3 \times \mathscr{W}_3$. Equations (2.42) and (2.43) become

$$p_n - k_1(p_1, \cdots, p_{n-1}) = y_n^2, \tag{2.45}$$

$$z - k_2(p_1, \cdots, p_{n-1}) = y_n^2 h_3(p_1, \cdots, p_{n-1}, y_n) \tag{2.46}$$

with $(p_1, \cdots, p_{n-1}) \in \mathscr{U}_3$ and $y_n \in \mathscr{W}_3$. This implies that $p_n - k_1$ is nonnegative. Conversely, whenever $p_n \geqq k_1$, we can solve (2.45) by $y_n = \pm\sqrt{p_n - k_1}$, getting two distinct values whenever the inequality is strict; possibly shrinking $\mathscr{U}_3$ to $\mathscr{U}_4$, we can arrange that both those values are in $\mathscr{W}_3$, so that (2.46) becomes

$$z - k_2 = (p_n - k_1) h(p_1, \cdots, p_{n-1}, \pm\sqrt{p_n - k_1}) \tag{2.47}$$

which, together with $(p_1, \cdots, p_{n-1}) \in \mathscr{U}_4$, completely describes $\pi_{xz}\mathscr{V}_2$.

If $h_1(\bar{p}_1, \cdots, \bar{p}_{n-1}, \bar{x}_n)$ should be negative, then we simply reverse p_n to $-p_n$, and we are back to the preceding case. So formulae (2.23) and (2.24) are proved.

For the sake of convenience, denote by Ω the set of points $(p_1, \cdots, p_n)$ such that $p_n > k_1$, and by Σ its boundary, the equation of which is $p_n = k_1$. Formula (2.24) yields along Σ

$$\begin{aligned} &\frac{\partial z_+}{\partial p_i} = \frac{\partial z_-}{\partial p_i} = \frac{\partial k_2}{\partial p_i}, \qquad 1 \leqq i \leqq n-1, \\ &\frac{\partial z_+}{\partial p_n} = \frac{\partial z_-}{\partial p_n} = h. \end{aligned} \tag{2.48}$$

It follows also from formula (2.24) that with any $p \in \Sigma$ and any vector $\pi' = (\pi'_1, \cdots, \pi'_n)$ pointing to the interior of Ω (i.e. $\pi'_n - \sum_{i=1}^{n-1} (\partial k_1/\partial p_i)\pi'_i > 0$) we can associate two continuous paths $t \to (p(t), x(t), z_+(t))$ and $t \to (p(t), x(t), z_-(t))$ in $\mathscr{L}V$ starting at (p, x, z) and satisfying $(dp/dt)(t) \to \pi'$ as $t \to 0$. From Proposition 1.4 (taking care that x- and p-coordinates are interchanged) it follows that, when $t \to 0$,

$$\frac{dz_+}{dt}(t) \to x \cdot \pi' \quad \text{and} \quad \frac{dz_-}{dt}(t) \to x \cdot \pi'. \tag{2.49}$$

But from formula (2.48) we get directly

$$\frac{dz_+}{dt}(t)\to\frac{\partial z}{\partial p}\cdot\pi' \quad\text{and}\quad \frac{\partial z_-}{\partial t}(t)\to\frac{\partial z}{\partial p}\cdot\pi' \tag{2.50}$$

where $\partial z/\partial p$ denotes the common value of the n-vectors (2.48). This yields $(\partial z/\partial p)\cdot\pi' = x\cdot\pi'$ for every vector π' in some half-space, and hence the desired formula $x=\partial z/\partial p$. □

In other words, $\mathscr{L}f$ is not defined locally for $p_n < k_1(p_1,\cdots,p_{n-1})$. In the region $p_n \geqq k_1(p_1,\cdots,p_{n-1})$, there are two well-defined branches for $\mathscr{L}f$. Along the boundary they coincide and have the same tangent hyperplane, and their shape away from the boundary is given by the following result.

COROLLARY 2.8. *We keep the assumptions and notations of Proposition* 2.7, *and we set* $q_n = p_n - k_1(p_1,\cdots,p_{n-1})$. *Then* $\mathscr{L}f$ *can be expanded near the boundary* $q_n=0$ *as*

$$z = k_2(p_1,\cdots,p_{n-1}) + q_n[a_0(p_1,\cdots,p_{n-1}) \pm a_1(p_1,\cdots,p_{n-1})\sqrt{q_n}] + O(q_n^{3/2}) \tag{2.51}$$

where the functions k_2, a_0, a_1 *are* C^∞. *Moreover*

$$\frac{\partial k_2}{\partial p_i}(p_1,\cdots,p_{n-1}) = x_i \quad \textit{for } 1\leqq i\leqq n-1, \tag{2.52}$$

$$a_0(p_1,\cdots,p_{n-1}) = x_n. \tag{2.53}$$

The proof consists simply of replacing h by its Taylor expansion in formula (2.24). We see that the two branches only intersect at the boundary $p_n = k_1$ of the admissible domain $p \geqq k_1$ (this is true even in the special case where $a_1 = 0$, because then the third order term $\pm a_3 q_n^{3/2}$ takes precedence). This is the classical "cusp" situation, so that Proposition 2.7 can be loosely stated as follows: a simple inflection point of f gives rise to a simple cusp of $\mathscr{L}f$.

Of course, more degenerate inflection points of f give rise to more complicated situations in $\mathscr{L}f$. A classification can be attempted along the lines of Proposition 2.7, but we are not going to conduct it any further. Let us only point out that, for all functions $f\in\mathscr{F}$, where $\mathscr{F}$ is a dense G_δ subset of $C^\infty(\mathbb{R}^n)$ in the Whitney topology, the space $\mathbb{R}^n$ can be partitioned as $\Sigma_0\cup\Sigma_1\cup\Sigma_2$ where:

Σ_0 consists of all points x where $A(x)$ is nondegenerate; it is an open subset of $\mathbb{R}^n$.

Σ_1 consists of all points x where $A(x)$ has rank $(n-1)$ and satisfies (2.22); it is a codimension one submanifold.

Σ_2 consists of all other points; it is a stratified subset of codimension $\geqq 2$.

Without going into details, this follows from Thom's transversality theorems. So, for most functions, the analysis performed thus far describes everything up to codimension two. In the one-dimensional case, $n=1$, that means precisely everything. Let us conclude by a simple example.

Define a function f on the real line by

$$f(x) = (x+x^2)^2. \tag{2.54}$$

We want to know what $\mathscr{L}f$ looks like. We need some data on f which are summarized in the following:

$$f'(x) = 4x(x+1)(x+\tfrac{1}{2}) = 4x^3+6x^2+2x,$$
$$f''(x) = 12x^2+12x+2,$$

x	$f(x)$	$p=f'(x)$	$f''(x)$	$z=f'(x)x-f(x)$
$-\infty$	$+\infty$	$-\infty$	*	$+\infty$
-1	0	0	*	0
-0.7887	$\frac{1}{36}$	0.19245	0	-0.1796
$-\frac{1}{2}$	$\frac{1}{16}$	0	*	$-\frac{1}{16}$
-0.2113	$\frac{1}{36}$	-0.19245	0	0.0129
0	0	0	*	0
$+\infty$	$+\infty$	$+\infty$	*	$+\infty$

We now can draw the graphs of f and $\mathscr{L}f$ (Figs. 1 and 2.) Note that the z-axis $p=0$ intersects $\mathscr{L}f$ at the simple point $z=-\frac{1}{16}$ and the double point $z=0$. This means that there are two distinct tangents to f with slope $p=0$: the first one is tangent to f at $x=-\frac{1}{2}$ only, the second one is tangent to f both at $x=-1$ and $x=0$. From formula (2.17), the tangent to $\mathscr{L}f$ at $(p=0, z=-\frac{1}{16})$ has slope $-\frac{1}{2}$, and the two branches of $\mathscr{L}f$ which intersect at $(p=0, z=0)$ have distinct tangents of slopes -1 and 0 respectively.

Moreover $\mathscr{L}f$ features two cusps at $(0.1945, -0.1796)$ and $(-0.1945, 0.0129)$. By Proposition 2.7, the tangents at those cusps are well-defined, and have slopes -0.7887 and -0.2113 respectively.

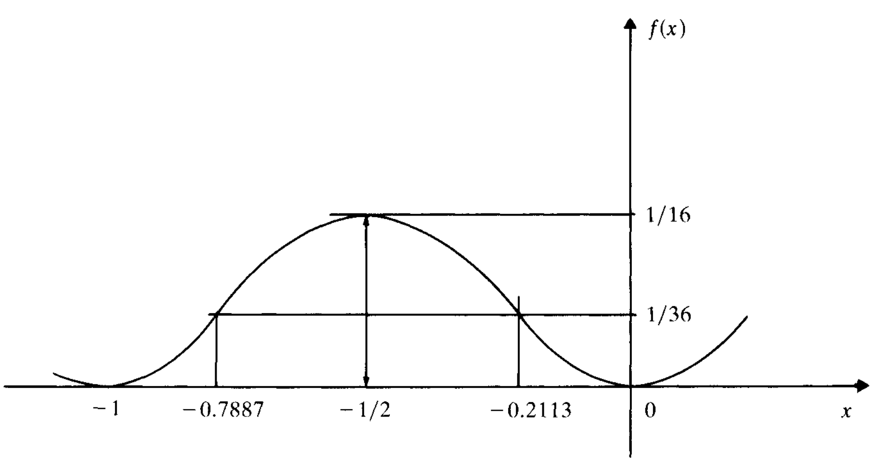

FIG. 1. $x \mapsto f(x)$. *Scale*: 0.01 ↑ → 0.1

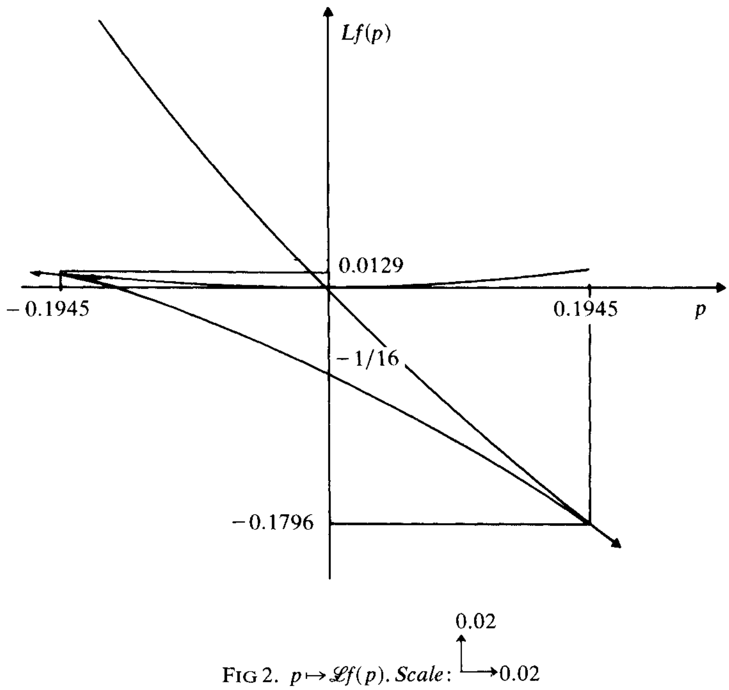

FIG 2. $p \mapsto \mathscr{L}f(p)$. *Scale*:

Note the parameteric equations for $\mathscr{L}f$:

$$p = 2x(x+1)(2x+1), \quad z = x(x+1)(3x^2+x). \tag{2.55}$$

Thus the graph of $\mathscr{L}f$ is the semi-algebraic set obtained by writing that the two algebraic equations (2.55) have a common solution in x, i.e. by eliminating x between the two equations.

3. Extremization problems and duality. Whenever V is a subset of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$, we shall denote by

$$\text{ext } V \tag{P}$$

the problem of determining all couples $(x, z) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$(x, 0, z) \in V. \tag{3.1}$$

$(\mathscr{P})$ will be termed an *extremization problem*, and any couple (x, z) satisfying (3.1) will be called a *solution* of $(\mathscr{P})$. The value of $(\mathscr{P})$, denoted by $\{\text{ext } \mathscr{P}\}$, will be the set of all $z \in \mathbb{R}$ such that there is an $x \in \mathbb{R}^n$ with (3.1) satisfied.

An important special case occurs when V is the Lagrangian submanifold associated with some C^∞ function $f: \mathbb{R}^n \to R$

$$V = \{(x, f'(x), f(x)) | x \in \mathbb{R}^n\}. \tag{3.2}$$

In that case formula (3.1) becomes

$$f'(x) = 0, \qquad z = f(x) \tag{3.3}$$

so that ($\mathscr{P}$) is simply the problem of determining the critical points and values of f. We shall write it

$$(\mathscr{P}) \qquad \operatorname*{ext}_x f(x)$$

and call it an *unconstrained smooth extremization problem.*

Another important special case occurs when

$$(3.4) \qquad V=\left\{\left(x, f'(x)-\sum_{j=1}^{k} \lambda_j g_j'(x), f(x)\right) \middle| g_j(x)=0, \lambda_j \in \mathbb{R}, 1 \leqq j \leqq k\right\}$$

where f and the g_j, $1 \leqq j \leqq k$, are C^∞ functions on $\mathbb{R}^n$. We set

$$(3.5) \qquad S=\pi_x V=\{x \mid g_j(x)=0, 1 \leqq j \leqq k\}.$$

LEMMA 3.1. *If the $g_j'(x)$, $1 \leqq j \leqq k$, are linearly independent at every $x \in S$, $x \in S$, then S is a closed submanifold of $\mathbb{R}^n$ and V is a Lagrangian submanifold of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$.*

Proof. The fact that S and V are closed $(n-k)$- and n-dimensional submanifolds follows easily from the implicit function theorem. We check condition (1.3) for V:

$$(3.6) \qquad \begin{aligned} i_V^* \omega &= df(x)-\left(f'(x)-\sum \lambda_j g_j'(x)\right) dx \\ &= (df(x)-f'(x)\, dx)+\sum \lambda_j g_j'(x)\, dx. \end{aligned}$$

The first term vanishes identically, and along V we have $g_j'(x)\, dx=0$ since $g_j(x)$ is a constant. □

The solutions of ($\mathscr{P}$) are all couples $(x, f(x))$ such that

$$(3.7) \qquad x \in S \quad \text{and} \quad \exists\, \lambda_1, \cdots, \lambda_k : f'(x)-\sum_{j=1}^{k} \lambda_j g_j'(x)=0.$$

If the $g_j'(x)$, $1 \leqq j \leqq k$, are linearly independent at every point $x \in S$, condition (3.7) means that x is a critical point of $f|_S$, the restriction of f to S. For that reason, we shall write ($\mathscr{P}$) as

$$(\mathscr{P}) \qquad \begin{aligned} &\operatorname{ext} f(x), \\ &g_j(x)=0, \qquad 1 \leqq j \leqq k, \end{aligned}$$

and call it a *constrained smooth extremization problem.*

Any critical point of a smooth convex (or concave) function is a minimum (or a maximum). For that reason, the various extremization problems we stated reduce to optimization problems when f is convex (or concave) and the g_j linear. So extremization is a natural generalization of optimization to the nonconvex case. Now it is a well-known fact that there is a duality theory of convex optimization problems, and we want to extend it to nonconvex extremization problems.

From now on we are given a linear map $A: \mathbb{R}^n \to \mathbb{R}^m$. We shall denote by x, p, y, q the vectors of $\mathbb{R}$, $(\mathbb{R}^n)^*$, $\mathbb{R}^m$, $(\mathbb{R}^m)^*$ respectively. With any subset V of

$\mathbb{R}^{n+m}\times\mathbb{R}^{n+m}\times\mathbb{R}$ we associate the subset V_A of $\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}$ defined by

$$V_A=\{(x,p+A^*q,z)|(x,Ax;p,q;z)\in V\}. \tag{3.8}$$

Applying this definition to the transpose $A^*:(\mathbb{R}^m)^*\to(\mathbb{R}^n)^*$,[3] and to any subset V^* of $\mathbb{R}^{n+m}\times\mathbb{R}^{n+m}\times\mathbb{R}$, we get

$$V^*_{A^*}=\{(q,y+Ax,z)|(A^*q,q;x,y;z)\in V^*\}\subset\mathbb{R}^m\times\mathbb{R}^m\times\mathbb{R}. \tag{3.9}$$

We now state the main result of this section.

THEOREM 3.2. *Let $A:\mathbb{R}^n\to\mathbb{R}^m$, be a linear map and V any subset of $\mathbb{R}^{n+m}\times\mathbb{R}^{n+m}\times\mathbb{R}$. Consider the extremization problems*

$$(\mathscr{P})\qquad \text{ext } V_A,$$

$$(\mathscr{P}^*)\qquad \text{ext } (\mathscr{L}V)_{-A^*}$$

The formulae

$$(x,Ax;-A^*q,q;z)\in V,\qquad z'=-z, \tag{3.10}$$

$$(-A^*q,q;x,Ax;z')\in\mathscr{L}V,\qquad z=-z' \tag{3.11}$$

are equivalent. Whenever (x,z) is a solution of $(\mathscr{P})$, the set of (q,z') satisfying (3.11) *or* (3.10) *is nonempty, and all of them are solutions of $(\mathscr{P}^*)$. Whenever (q,z') is a solution of $(\mathscr{P}^*)$, the set of (x,z) satisfying* (3.11) *or* (3.10) *is nonempty, and all of them are solutions of $(\mathscr{P})$.*

Proof. To say that (x,z) is a solution of $(\mathscr{P})$ means that there exists (p,q) such that

$$(x,Ax;p,q;z)\in V,\qquad p+A^*q=0 \tag{3.12}$$

which we may write in a more symmetric form:

$$(x,y;p,q;z)\in V,\quad y-Ax=0,\quad p+A^*q=0. \tag{3.13}$$

Applying the Legendre transformation, we obtain

$$(p,q;x,y;px+qy-z)\in\mathscr{L}V,\quad y-Ax=0,\quad p+A^*q=0. \tag{3.14}$$

The last two equations imply that

$$z'=px+qy-z=-A^*q\cdot x+q\cdot Ax-z=-z \tag{3.15}$$

and formula (3.14) becomes

$$(p,q;x,y;-z)\in\mathscr{L}V,\quad y-Ax=0,\quad p+A^*q=0. \tag{3.16}$$

Breaking the symmetry, we get

$$(-A^*q,q;x,y;-z)\in\mathscr{L}V,\qquad y-Ax=0 \tag{3.17}$$

which means precisely that $(q,-z)$ is a solution of $(\mathscr{P}^*)$. Since the Legendre transformation is an involution, formulae (3.12) and (3.17) are equivalent, and set up a one-to-one pairing between solutions (x,z) of $(\mathscr{P})$ and $(q,-z)$ of $(\mathscr{P}^*)$. But (3.12) is just (3.10), and (3.17) is (3.11). □

[3] From now on we shall omit the star.

The following is an easy consequence of the fact that the Legendre transformation $\mathscr{L}$ and the operation $A \to -A^*$ are involutions.

COROLLARY 3.3. $(\mathscr{P}^{**}) = (\mathscr{P})$.

Problems $(\mathscr{P})$ and $(\mathscr{P}^*)$ will be said to be *dual* to each other. Another easy consequence of Theorem 3.2 is the following.

COROLLARY 3.4. $\{\operatorname{ext} \mathscr{P}\} = -\{\operatorname{ext} \mathscr{P}^*\}$.

Theorem 3.2 is more readily understandable in the case of unconstrained smooth extremization problems. It reads as follows.

PROPOSITION 3.5. *Let $A: \mathbb{R}^n \to \mathbb{R}^m$ be a linear map and $f: \mathbb{R}^{n+m} \to \mathbb{R}$ a C^∞ function. Consider the extremization problems*:

$$(\mathscr{P}) \qquad \operatorname*{ext}_x f(x, Ax),$$

$$(\mathscr{P}^*) \qquad \operatorname*{ext}_q \mathscr{L}f(-A^*q, q).$$

The formulae

$$(3.18) \qquad -A^*q = f'_x(x, Ax), \quad q = f'_y(x, Ax), \quad z' = -f(x, Ax)$$

set up a one-to-one pairing between solutions $(x, f(x, Ax))$ of $(\mathscr{P})$ and (q, z') of $(\mathscr{P}^*)$. Whenever the matrix of second derivatives f'' has rank $(n+m)$ at (x, Ax), there is a neighborhood $\mathscr{U}$ of $(-A^*q, q)$ and a C^∞ selection $\mathscr{L}_{\mathscr{U}}f$ of $\mathscr{L}f$ over $\mathscr{U}$ such that

$$(3.19) \qquad f(x, Ax) = -(\mathscr{L}_{\mathscr{U}}f)(-A^*q, q),$$

$$(3.20) \qquad x = (\mathscr{L}_{\mathscr{U}}f)'_p(-A^*q, q), \qquad Ax = (\mathscr{L}_{\mathscr{U}}f)'_q(-A^*q, q).$$

This follows easily from taking $V = V_f$, the Lagrangian submanifold associated with f, in Theorem 3.2. The last part is a consequence of Proposition 2.6. Note that relations analogous to (3.20) hold whenever $(\mathscr{L}f)'$ can be defined in a consistent way at $(p, q; z')$; this would be the case for the cusp points described in Proposition 2.7.

Let us give an important special case.

COROLLARY 3.6. *Let $\varphi: \mathbb{R}^n \to \mathbb{R}$ and $\psi: \mathbb{R}^m \to \mathbb{R}$ be C^∞ functions, and consider the extremization problems*

$$(\mathscr{P}) \qquad \operatorname*{ext}_x \varphi(x) + \psi(Ax),$$

$$(\mathscr{P}^*) \qquad \operatorname*{ext}_q \mathscr{L}\varphi(-A^*q) + \mathscr{L}\psi(q).$$

Then $\{\operatorname{ext} \mathscr{P}\} = -\{\operatorname{ext} \mathscr{P}^*\}$, *and there is a one-to-one pairing between solutions $(x, \varphi(x) + \psi(Ax))$ of $(\mathscr{P})$ and (q^*, z') of $(\mathscr{P}^*)$, described by the relation*

$$(3.21) \qquad -A^*q = \varphi'(x), \quad q = \psi'(Ax), \quad -z' = \varphi(x) + \psi(Ax).$$

Whenever φ'' has rank n at x and ψ'' has rank m at Ax, there are neighborhoods $\mathscr{U}_1$ and $\mathscr{U}_2$ of $-A^*q$ and q, selections $\mathscr{L}_{\mathscr{U}}\varphi$ and $\mathscr{L}_{\mathscr{U}}\psi$ of $\mathscr{L}\varphi$ and $\mathscr{L}\psi$ over $\mathscr{U}_1$ and $\mathscr{U}_2$, *such that*

$$\mathcal{L}_{\mathcal{U}}\varphi(-A^*q)+\mathcal{L}_{\mathcal{U}}\psi(q)=\varphi(x)+\psi(Ax), \tag{3.22}$$

$$x=(\mathcal{L}_{\mathcal{U}}\varphi)'(-A^*q), \qquad Ax=(\mathcal{L}_{\mathcal{U}}\psi)'(q). \tag{3.23}$$

We now give two examples of applications of Theorem 3.2. They are both related to the problem of finding the eigenvectors and eigenvalues of a self-adjoint operator: we write it as an extremization problem in two different ways, and dualize both of them.

Let us start with the constrained smooth extremization problem

$$(\mathcal{P}) \qquad \begin{gathered} \operatorname{ext} \|Ax\|^2, \\ \|x\|^2=1. \end{gathered}$$

A solution to $(\mathcal{P})$ is a couple (x, z) such that

$$\|x\|^2=1, \qquad \exists\, \lambda\in\mathbb{R}\colon A^*Ax-\lambda x=0, \tag{3.24}$$

$$z=\|Ax\|^2=\lambda, \tag{3.25}$$

i.e. x is an eigenvector of A^*A and z is the corresponding eigenvalue.

Consider the subset $V\subset\mathbb{R}^{n+m}\times\mathbb{R}^{n+m}\times\mathbb{R}$ defined by

$$V=\{(x, y; -2\lambda x, 2y; \|y\|^2) \mid \|x\|^2=1, \lambda\in\mathbb{R}\}. \tag{3.26}$$

By Lemma 3.1 it is a Lagrangian submanifold. It is clear that problem $(\mathcal{P})$ is simply ext V_A. For the sake of convenience we will cut out part of V; indeed, it is apparent from formula (3.25) that $\lambda\geqq 0$ for any solution (x, z) of $\mathcal{P}$. So we introduce the "Lagrangian submanifold with boundary"

$$V'=\{(x, y; -2\lambda x, 2y; \|y\|^2) \mid \|x\|^2=1, \lambda\geqq 0\} \tag{3.27}$$

and we state problem $(\mathcal{P})$ as

$$(\mathcal{P}) \qquad \operatorname{ext} V'_A.$$

The Legendre transform of V' is again a Lagrangian submanifold with boundary. Going through the computations, we write it as a disjoint union $\mathcal{L}V=\Omega\cup\Gamma$, where Γ is the boundary

$$\Omega=\{(p, q; -p/\|p\|, q/2; -\|p\|+\|q\|^2/4) \mid p\neq 0\}, \tag{3.28}$$

$$\Gamma=\{(0, q; \xi, q/2; \|q\|^2/4) \mid \|\xi\|^2=1\}. \tag{3.29}$$

$\mathcal{L}V$ is clearly associated with the function $(p, q)\to -\|p\|+\|q\|^2/4$. The function $p\to -\|p\|$ is not differentiable at the origin, but let us agree that

$$\frac{d}{dp}(-\|p\|)|_{p=0}=\{\xi\in\mathbb{R}^n \mid \|\xi\|^2=1\}. \tag{3.30}$$

This being agreed upon, we can now state the dual problem $(\mathcal{P}^*)$ in the following way:

$$(\mathcal{P}^*) \qquad \operatorname*{ext}_q -\|A^*q\|+\|q\|^2/4.$$

Theorem 3.2 implies that whenever $(q, -\|A^*q\|+\|q\|^2/4)$ is a solution to $(\mathscr{P}^*)$, all couples $(x, \|Ax\|^2)$ given by

$$x = A^*q/\|A^*q\| \quad \text{if } A^*q \neq 0, \quad Ax = q/2, \quad \|x\|^2 = 1, \tag{3.31}$$

$$\|Ax\|^2 = \|A^*q\| - \|q\|^2/4 \tag{3.32}$$

are solutions to $(\mathscr{P})$; in other words x is an eigenvector of A^*A with norm one, and $\|A^*q\| - \|q\|^2/4$ is the corresponding eigenvalue. For instance, formula (3.30) shows us that $(0, 0)$ is a solution to $(\mathscr{P}^*)$ provided there exist $\xi \in \mathbb{R}^n$ with $\|\xi\|^2 = 1$ and $A\xi = 0$. Formulae (3.31) and (3.32) then yield the trivial fact that every such ξ is an eigenvector of A^*A with eigenvalue 0. Note as a conclusion that $-\{\text{ext } \mathscr{P}^*\}$ is just the spectrum of A^*A.

We now treat the same problem in another way. We define a subset W of $\mathbb{R}^{n+m} \times \mathbb{R}^{n+m} \times \mathbb{R}$ by

$$\begin{aligned} W = \{(x, y; -2x\|y\|^2/\|x\|^4, 2y/\|x\|^2; \|y\|^2/\|x\|^2) | x \neq 0\} \\ \cup \{(0, 0; 0, \eta; 0) | \eta \in \mathbb{R}^m\}. \end{aligned} \tag{3.33}$$

It can be checked that W is a Lagrangian submanifold. We asociate with it the extremization problem

$$(\mathscr{P}) \qquad \text{ext } W_A$$

which we state somewhat loosely as

$$(\mathscr{P}) \qquad \text{ext } \|Ax\|^2/\|x\|^2.$$

Of course, solving $(\mathscr{P})$ is just looking for the eigenspaces of A^*A. We now construct the dual problem $(\mathscr{P}^*)$. A simple computation yields

$$\begin{aligned} \mathscr{L}W = \{(p, q; -2p/\|q\|^2, 2q\|p\|^2/\|q\|^4; -\|p\|^2/\|q\|^2) | q \neq 0\} \\ \cup \{0, 0; \pi, 0; 0) | \pi \in \mathbb{R}^n\}. \end{aligned} \tag{3.34}$$

The dual problem $(\mathscr{P}^*)$, which is ext W_{-A^*}, will be stated somewhat loosely as

$$\text{ext } -\|A^*q\|^2/\|q\|^2. \tag{3.35}$$

We leave it to the reader to see what becomes of formulae (3.10)–(3.11). They tell us essentially that the eigenvalues of A^*A and AA^* coincide—a trivial fact.

We conclude this section by pointing out a technicality: even if V is a Lagrangian submanifold of $\mathbb{R}^{n+m} \times \mathbb{R}^{n+m} \times \mathbb{R}$, the set V_A need not be a Lagrangian submanifold of $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$. Indeed, it need neither be closed nor be a submanifold. As a simple example, take

$$V = \{(x, y; -y/x^2, 1/x; y/x) | \neq 0\} \tag{3.36}$$

a Lagrangian submanifold of $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}$. Setting $A: x \to mx$, we get

$$V_A = \{(x, 0, m) | x \neq 0\} \tag{3.37}$$

which is not closed in $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

However, we have the following.

LEMMA 3.7. *If V is a Lagrangian submanifold and if V_A is a closed submanifold, then V_A is Lagrangian.*

Proof. We check condition (1.3) for V_A:

$$(3.38)\qquad \begin{aligned} i_V^*\omega &= dz-(p+A^*q)\,dx \\ &= dz-p\,dx-qd(Ax) \end{aligned}$$

which is zero since $(x, Ax; p, q; z) \in V$, and the restriction of ω to V vanishes. □

Note also that if V is the Lagrangian submanifold associated with a C^∞ function $F:\mathbb{R}^n = \mathbb{R}^m \to \mathbb{R}$, then V_A is the Lagrangian submanifold associated with the C^∞ function $x \mapsto f(x, Ax)$ from $\mathbb{R}^n$ to $\mathbb{R}$—a fact we have used repeatedly in this section.

4. Applications to the calculus of variations. From now on, $\bar{\Omega} \subset \mathbb{R}^n$ will be an n-dimensional C^∞-submanifold with boundary Γ. We set $\Omega = \bar{\Omega} - \Gamma$, an open subset of $\mathbb{R}^n$; we endow Ω with the Lebesgue measure $d\omega$ and Γ with the induced $(n-1)$-dimensional measure $d\gamma$.

We consider a continuous linear map $A: V \to E$, where $E = L^2(\Omega; \mathbb{R}^m)$ and V is some Hilbertian subspace of $H = L^2(\Omega; \mathbb{R}^k)$ (i.e. V is a linear subspace of H endowed with some Hilbertian structure such that the inclusion mapping $V \to H$ is continuous). We assume that there is some Hilbert space T and some continuous liner map τ: $V \to T$ such that τ is surjective and $V_0 = \tau^{-1}(0)$ is dense in H. In practical examples, A will be some differential operator, V_0 will be $\overline{\mathscr{D}(\Omega)}$, the closure in V of the set of C^∞ functions with compact support in Ω, and T will associate with every function in V its "trace" on the boundary Γ. We shall state an abstract Green's formula for later use.

THEOREM 4.1. *There exist a Hilbertian subspace V^* of E, and continuous linear maps A^*: $V^* \to H$ and τ^*: $V^* \to T'$, the topological dual of T, such that, for every $x \in V$ and $q \in V^*$, we have*

$$(4.1)\qquad (q, Ax) - (A^*q, x) = \langle \tau^*q, \tau x\rangle$$

where $(\cdot\,,\,\cdot)$ denotes scalar product in L^2 and $\langle \cdot\,,\,\cdot \rangle$ denotes the duality pairing between T' and T.

We now turn to extremization problems in the calculus of variations. From now on, we are given a family W_ω, $\omega \in \Omega$, of Lagrangian submanifolds of $\mathbb{R}^{k+m} \times \mathbb{R}^{k+m} \times \mathbb{R}$, and we denote by $F_\omega(x, y)$ the associated characteristic maps. Moreover, we are given a convex lower semi-continuous function Φ: $T \to \mathbb{R} \cup \{+\infty\}$; as usual in convex analysis, its subdifferential will be denoted by $\partial\Phi$. We now state.

DEFINITION 4.2. The *calculus of variations problem*[4]

$$(\mathscr{P})\qquad \operatorname*{ext}_{x\in V} \int_\Omega F_\omega(x(\omega), Ax(\omega))\, d\omega + \Phi(\tau x)$$

consists in looking for all mappings $\omega \to (x(\omega), q(\omega), z(\omega))$ from Ω to

[4] Henceforth denoted by C.V. problem.

$\mathbb{R}^{k+m} \times \mathbb{R}^{k+m} = \mathbb{R}$ such that

(4.2) $\quad x \in V, \quad q \in V^*, \quad z \in L^1,$

(4.3) $\quad (x(\omega), Ax(\omega); -A^*q(\omega), q(\omega); z(\omega)) \in W_\omega$ for a.e. $\omega \in \Omega$,

(4.4) $\quad \tau^* q \in -\partial\Phi(\tau x)$.

Any pair $(x, z) \in V \times L^1$ such that there exists $q \in V^*$ satisfying (4.2)–(4.4) will be called an *extremal* of $(\mathcal{P})$. The number ζ defined by

$$\zeta = \int_\Omega z(\omega)\, d\omega + \Phi(\tau x) \tag{4.5}$$

will be the associated *value* of $(\mathcal{P})$. The set of values of problem $(\mathcal{P})$ will be denoted by $\{\text{ext}\, \mathcal{P}\}$.

The motivation for this definition is clear. In the case where $F_\omega(\xi, \eta) = f(\omega; \xi, \eta)$, a function which is C^∞ in (ξ, η) for almost every $\omega \in \Omega$, and measurable in ω for every $(\xi, \eta) \in \mathbb{R}^k \times \mathbb{R}^m$, then (4.2)–(4.6) become

(4.6) $\quad f'_\xi(\omega; x(\omega), Ax(\omega)) + A^* f'_\eta(\omega;, x(\omega), Ax(\omega)) = 0$ a.e.,

$$\tau^*[f'_\eta(x, Ax)] \in -\partial\Phi(x). \tag{4.7}$$

Equation (4.6) is the Euler–Lagrange equation on Ω associated with the integral

$$\int_\Omega f(\omega; x(\omega), Ax(\omega))\, d\omega \tag{4.8}$$

and formula (4.7) yields the so-called transversality conditions on the boundary Γ. In the case where f is convex in (ξ, η) for every ω, those are necessary and sufficient conditions for optimality. If f is not convex, but satisfies some growth condition at infinity, we get the first-order conditions for stationarity.

We now state the duality theorem.

THEOREM 4.3. *Consider the C.V. problems*

$$(\mathcal{P}) \qquad \underset{x \in V}{\text{ext}} \int_\Omega F_\omega(x(\omega), Ax(\omega))\, d\omega + \Phi(\tau x),$$

$$(\mathcal{P}^*) \qquad \underset{q \in V^*}{\text{ext}} \int_\Omega \mathcal{L}F_\omega(-A^*q(\omega), q(\omega))\, d\omega - \Phi^*(-\tau^* q).$$ [5]

Let (x, z) be an extremal of $(\mathcal{P})$ with value ζ; then, for any q satisfying (4.2)–(4.4), $(q, -xA^*q + qAx - z)$ is an extremal of $(\mathcal{P}^*)$ with value $-\zeta$. Conversely, let $(q, z') \in V^* \times L^1$ be an extremal of $(\mathcal{P}^*)$ with value ζ'; then, for any $x \in V$ satisfying

(4.9) $\quad (-A^*q(\omega), q(\omega); x(\omega), Ax(\omega); z'(\omega)) \in \mathcal{L}W_\omega$ *for a.e.* $\omega \in \Omega$,

$$\tau x \in \partial\Phi^*(-\tau^* q), \tag{4.10}$$

---

[5] Φ^* is the Fenchel conjugate of Φ in the sense of convex analysis:

$$\Phi^*(\delta') = \sup\{\langle\delta, \delta'\rangle - \Phi(\delta) | \delta \in T\} \quad \forall \delta' \in T'.$$

$(x, qAx - xA^*q - z')$ is an extremal of $(\mathcal{P})$ with value $-\zeta'$. Hence

$$\{\text{ext}\, \mathcal{P}\} = -\{\text{ext}\, \mathcal{P}^*\}. \tag{4.11}$$

Proof. The pointwise equation

$$(x(\omega), Ax(\omega); -A^*q(\omega), q(\omega); z(\omega)) \in W_\omega \tag{4.12}$$

can be written

(4.13)
$$(-A^*q(\omega), q(\omega); x(\omega), Ax(\omega); -x(\omega)A^*q(\omega) + Ax(\omega)q(\omega) - z(\omega)) \in \mathcal{L}W_\omega.$$

Moreover, formula (4.4) an also be written

$$\tau x \in \partial\Phi^*(-\tau^*q). \tag{4.14}$$

But equations (4.13) and (4.14), together with $x \in V$, $q \in V^*$, $z \in L^1$, simply mean that $(q, -xA^*q + Axq - z)$ is an extremal of $(\mathcal{P}^*)$. The associated value is

$$\zeta' = \int_\Omega (-x(\omega)A^*q(\omega) + Ax(\omega)q(\omega) - z(\omega))\, d\omega - \Phi^*(-\tau^*q). \tag{4.15}$$

Using Green's formula we have

$$\zeta' = -\int_\Omega z(\omega)\, d\omega + \langle \tau^*q, \tau x\rangle - \Phi^*(-\tau^*q). \tag{4.16}$$

Making use of (4.14), this becomes

$$\zeta' = -\int_\Omega z(\omega)\, d\omega - \Phi(\tau x) = -\zeta. \tag{4.17}$$

Hence the first part of the theorem is proved. The converse is proved along the same lines. □

Typical instances of such a mapping $A: V \to E$ are

$$\overrightarrow{\text{grad}}: H^1(\Omega) \to L^2(\Omega; \mathbb{R}^n), \tag{4.18}$$

$$\Delta: H^2(\Omega) \to L^2(\Omega; \mathbb{R}). \tag{4.19}$$

In the first case, T is $H^{1/2}(\Gamma)$, and Green's formula reads

$$\int_\Omega (\overrightarrow{\text{grad}}\, x \cdot \vec{q} + x \cdot \text{div}\, \vec{q})\, d\omega = \int_\Gamma \vec{n} \cdot \vec{q} x\, d\gamma. \tag{4.20}$$

In the second case, T is $H^{3/2}(\Gamma)$, and Green's formula reads

$$\int_\Omega (\Delta x \cdot q + x \cdot \Delta q)\, d\omega = \int_\Omega (q\vec{n} \cdot \overrightarrow{\text{grad}}\, x + x\vec{n} \cdot \overrightarrow{\text{grad}}\, q)\, d\gamma. \tag{4.21}$$

In both cases, we could define Φ as

$$\Phi(\delta) = \begin{cases} 0 & \text{if } \delta = \delta_0, \\ +\infty & \text{otherwise} \end{cases} \tag{4.22}$$

which gives a Dirichlet condition (fixed boundary values). We could also define

$$\Phi(\delta) = \begin{cases} 0 & \text{if } \int_\Gamma \delta = 0, \\ +\infty & \text{otherwise} \end{cases} \tag{4.23}$$

which is a kind of periodicity condition.

Let us give an example:

$$(\mathscr{P}) \qquad \begin{gathered} \operatorname{ext} \int_\Omega f(\omega; x(\omega), \operatorname{grad} x(\omega))\, d\omega, \\ x \in H^1(\Omega), \qquad \int_\Gamma x(\gamma)\, d\gamma = 0 \end{gathered}$$

has the following dual:

$$(\mathscr{P}^*) \qquad \begin{gathered} \operatorname{ext} \int_\Omega \mathscr{L}f(\omega; -\operatorname{div} q(\omega), q(\omega))\, d\omega, \\ q \in H(\Omega; \operatorname{div}), \qquad q = \text{const. on } \Gamma \end{gathered}$$

where $H(\Omega, \operatorname{div}) = \{u \in L^2(\Omega, \mathbb{R}^n) | \operatorname{div} u \in L^2(\Omega, \mathbb{R}^n)\}$. The task of rewriting (4.2)–(4.4) and (4.9)–(4.11) is left to the reader.

We are now going to show that we can get simultaneously the extremals (x, z) of $(\mathscr{P})$ and the extremals (q, z') of $(\mathscr{P}^*)$ from the extremals of a single C.V. problem.

PROPOSITION 4.4. *Consider the C.V. problems*

$$(\mathscr{Q}) \quad \underset{\substack{(x,y,q)\in \\ V\times E\times V^*}}{\operatorname{ext}} \int_\Omega [-A^*q(\omega)\cdot x(\omega) + q(\omega)y(\omega) - F_\omega(x(\omega), y(\omega))]d\omega - \Phi^*(-\tau^* q),$$

$$(\mathscr{Q}^*) \quad \underset{\substack{(x,p,q)\in \\ V\times E\times V^*}}{\operatorname{ext}} \int_\Omega [p(\omega)x(\omega) + q(\omega)\cdot Ax(\omega) - \mathscr{L}F_\omega(p(\omega), q(\omega))]\, d\omega + \Phi(\tau x).$$

The following are equivalent statements:

(a) (x, y, q, z') *is an extremal of* $(\mathscr{Q})$,

(b) (x, p, q, z) *is an extremal of* $(\mathscr{Q}^*)$,

(c) (x, q, z) *satisfy* (4.2)–(4.4),

(d) (q, x, z') *satisfy* (4.9)–(4.10) *and* $z' \in L^1$,

with $z + z' = -A^*q \cdot x + q \cdot Ax$. *In particular* (x, z) *is an extremal of* $(\mathscr{P})$ *and* (q, z') *an extremal of* $(\mathscr{P}^*)$.

Proof. We have already shown that (c) and (d) are equivalent. We shall be content with proving that (a) and (c) are equivalent; the proof that (b) and (d) are equivalent goes along the same lines.

Problem $(\mathscr{Q})$ can be written as

$$(\mathscr{Q}') \qquad \underset{\substack{(x,y,q)\in \\ V\times E\times V^*}}{\operatorname{ext}} \int_\Omega \mathscr{F}_\omega(x(\omega), y(\omega), -A^*q(\omega), q(\omega))\, d\omega$$

where $\mathscr{F}_\omega$ is the characteristic map associated wth the Lagrangian submanifold $\mathscr{W}_\omega$ of $\mathbb{R}^{2k+2m}\times\mathbb{R}^{2k+2m}\times\mathbb{R}$ defined by

(4.24) $$\mathscr{W}_\omega=\{((\xi,\eta,\pi,\rho;\pi-\sigma,\rho-\tau,\xi,\eta;\pi\xi+\rho\eta-\zeta)|\pi\in\mathbb{R}^k,\ \rho\in\mathbb{R}^m,(\xi,\eta;\sigma,\tau;\zeta)\in W_\omega\}.$$

We now apply Definition 4.2 to the Hilbert space $\mathscr{V}=V\times E\times V^*$ and the map $\mathscr{A}$: $\mathscr{V}\to E$ defined by $\mathscr{A}(x,y,q)=-A^*q$; its adjoint will be the map $\mathscr{A}^*$: $V\to H\times E\times H$ defined by $\mathscr{A}^*(x')=(0,0,-Ax')$. Conditions (4.2)–(4.4) then become

(4.25) $$x\in V,\quad y\in E,\quad q\in V^*,\quad x'\in V,\quad z'\in L^1,$$

(4.26) $$(x(\omega),y(\omega),-A^*q(\omega),q(\omega);0,0,x'(\omega),Ax'(\omega);z'(\omega))\in\mathscr{W}_\omega\quad\text{a.e.},$$

(4.27) $$\tau x'\in\partial\Phi^*(-\tau q).$$

So $(x;y,q,z')\in V\times E\times V^*\times L^1$ is an extremal of $(\mathscr{Q})$ if and only if there exists $x'\in V$ such that (4.26) and (4.27) are satisfied. Now, comparing (4.26) with (4.24), we get

(4.28) $$-A^*q(\omega)=\sigma,$$

(4.29) $$q(\omega)=\tau,$$

(4.30) $$x'(\omega)=x(\omega),$$

(4.31) $$Ax'(\omega)=y(\omega),$$

(4.32) $$z'(\omega)=-A^*q(\omega)\cdot x(\omega)+q(\omega)y(\omega)-\zeta,$$

(4.33) $$x(\omega),y(\omega);\sigma,\tau;\zeta)\in W_\omega.$$

All this boils down to

(4.34) $$(x(\omega),Ax(\omega);-A^*q(\omega),q(\omega);z(\omega))\in W_\omega\quad\text{a.e.}$$

with $z(\omega)+z'(\omega)=-A^*q(\omega)\cdot x(\omega)+q(\omega)\cdot Ax(\omega)$. With (4.30) taken into account, (4.27) becomes

(4.35) $$\tau x\in\partial\Phi^*(-\tau^*q)$$

which can be inverted to

(4.36) $$-\tau^*q\in\partial\Phi(\tau x).$$

But (4.34) and (4.36) are just (c), and we have proved our claim. □

Proposition 4.4 can be considered a smooth version of the saddle-point property for Lagrange multipliers in convex optimization. Note that in the case where $F_\omega(\xi,\eta)=f(\omega;\xi,\eta)$, measurable in ω, C^∞ in (ξ,η), problem $(\mathscr{P}^*)$ involves $\mathscr{L}f(\omega;\xi,\eta)$ which typically is multivalued and cusped; working with problem $(\mathscr{Q})$ is a way of circumventing this inconvenience at the cost of increasing the dimension.

We now apply this idea of "smoothing out" Legendre transforms to another example.

PROPOSITION 4.5. *We are given a C^∞ function $\varphi: [0, T]\times \mathbb{R}^n \to \mathbb{R}$, a measurable function $f: [0, T]\to \mathbb{R}^n$, and a point $\xi_0 \in \mathbb{R}^n$. We consider the differential equation*

$$(\mathscr{E}) \qquad \frac{dx}{dt}+\varphi'_\xi(t, x)=f \quad \text{a.e. on } [0, T],\ x(0)=\xi_0,$$

and the C.V. problems

$$(\mathscr{P}) \qquad \begin{aligned} &\operatorname{ext}\int_0^T\left[\varphi(t, x)+\mathscr{L}\varphi\left(t; f-\frac{dx}{dt}\right)+x\left(\frac{dx}{dt}-f\right)\right]dt,\\ &x\in H^1(0, T;\mathbb{R}^n), \qquad x(0)=x_0;\end{aligned}$$

$$(\mathscr{Q}) \qquad \begin{aligned} &\operatorname{ext}\int_0^T\left[\varphi(t, x)-\varphi(t, y)+\left(\frac{dx}{dt}-f\right)(x-y)\right]dt,\\ &x\in H^1(0, T;\mathbb{R}^n), \quad y\in H^1(0, T;\mathbb{R}^n), \quad x(0)=y(0)=\xi_0.\end{aligned}$$

If $(\mathscr{E})$ has no solution, then problems $(\mathscr{P})$ and $(\mathscr{Q})$ have no extremals. If $(\mathscr{E})$ has a solution $\bar{x}$, then problem $(\mathscr{P})$ has a unique extremal $(\bar{x}, 0)$, and problem $(\mathscr{Q})$ has a unique extremal $(\bar{x}, \bar{x}, 0)$.

Proof. Problem $(\mathscr{Q})$ arises from problem $(\mathscr{P})$ by replacing $\mathscr{L}\varphi(f-(dx/dt))$ by $y(f-(dx/dt))-\varphi(y)$, i.e. by smoothing out that part of the integrand which is a Legendre transform. Proposition 4.4 does not readily apply to this case, so we give a direct proof.

An extremal (x, y, z) of $(\mathscr{Q})$ is defined by the Euler equations

$$\varphi'_\xi(t, x)+\frac{dx}{dt}-f=\frac{d}{dt}[x-y], \tag{4.37}$$

$$-\varphi'_\xi(t, y)-\frac{dx}{dt}+f=0 \tag{4.38}$$

and the boundary conditions $x(0)=y(0)=x_0$. Together, they yield the system of differential equations on $[0, T]$

$$\frac{dy}{dt}+\varphi'_\xi(t, x)=f, \qquad y(0)=x_0, \tag{4.39}$$

$$\frac{dx}{dt}+\varphi'_\xi(t, y)=f, \qquad x(0)=x_0. \tag{4.40}$$

Now this is to be compared with equation

$$(\mathscr{E}) \qquad \frac{dx}{dt}+\varphi'_\xi(t, x)=f, \qquad x(0)=x_0.$$

The assumptions on φ imply that both system (4.39)–(4.40) and equation $(\mathscr{E})$ have at most one solution. If $\bar{x}$ is the solution of $(\mathscr{E})$, obviously $(\bar{x}, \bar{x})$ is the solution of (4.39)–(4.40). Conversely, if $(\bar{x}, \bar{y})$ is a solution of (4.39)–(4.40), then so is $(\bar{y}, \bar{x})$; from the uniqueness, it follows that $\bar{x}=\bar{y}$, obviously the solution of $(\mathscr{E})$. Writing $x=\bar{x}=\bar{y}=y$ in the integrand, we see that it is identically zero. We have proved the equivalence of equation $(\mathscr{E})$ and problem $(\mathscr{Q})$.

The equivalence of problems $(\mathscr{P})$ and $(\mathscr{Q})$ goes along the lines set up in Proposition 4.4. Indeed, (4.38) means simply that

$$-\varphi(t, y)-\left(\frac{dx}{dt}-f\right)y=\mathscr{L}\varphi\left(t; f-\frac{dx}{dt}\right) \tag{4.41}$$

and the integrands in $(\mathscr{P})$ and $(\mathscr{Q})$ become equal. With $x = y$, formula (4.41) yields, with a slight misuse of notations,

$$[\mathscr{L}\varphi]'_\xi\left(t, f-\frac{dx}{dt}\right)=x \tag{4.42}$$

and the Euler equation for $(\mathscr{P})$ turns out to be exactly equation $(\mathscr{E})$. □

Note that we have defined directly the extremals of a problem in the calculus of variations, without reference to any extremization problem. This is because the natural extremization problem involved is infinite-dimensional, and the results of the preceding sections do not extend readily to this case; indeed, smoothness assumptions which are natural in finite dimensions become preposterous in this new setting. In some particular cases, however, it can be made to work. Let us give an example, which will be recognized as an infinite-dimensional version of the example concluding § 3.

We consider the space $V=H_0^1(\Omega)$ and the function

$$f\colon V\backslash\{0\}\times L^2(\Omega)^n \to \mathbb{R}, \tag{4.43}$$

$$f(x, y)=|y|^2/|x|^2 \tag{4.44}$$

with $|\cdot|$ denoting the L^2-norm. Obviously f is a C^∞ function, with

$$p=f'_x(x, y)=-2x|y|^2/|x|^4\in L^2(\Omega), \tag{4.45}$$

$$q=f'_y(x, y)=2y/|x|^2\in L^2(\Omega)^n. \tag{4.46}$$

We now set

$$y=\operatorname{grad} x \tag{4.47}$$

to get the extremization problem

$$(\mathscr{P})\qquad \begin{aligned}&\operatorname{ext}|\operatorname{grad} x|^2/|x|^2,\\ &x\in H_0^1(\Omega),\qquad x\neq 0.\end{aligned}$$

Let us write out the equation for a critical point, taking into account the fact that the transpose of grad: $H_0^1(\Omega)\to L^2(\Omega)^n$ is $-$div: $L^2(\Omega)^n\to H^{-1}(\Omega)$:

$$0=p-\operatorname{div} q=-2(x|\operatorname{grad} x|^2/|x|^2+\operatorname{div}\operatorname{grad} x)/|x|^2. \tag{4.48}$$

Note that $|\operatorname{grad} x|$ cannot be zero unless x is, so (4.48) becomes

$$x=-\frac{|x|^2}{|\operatorname{grad} x|^2}\Delta x,\qquad x\neq 0. \tag{4.49}$$

In other words, the solutions of $(\mathscr{P})$ are the pairs $(x, 1/\lambda)$ where $-\lambda$ is a nonzero eigenvalue of the Laplacian under homogeneous boundary conditions, and x any nonzero eigenvector.

To get the dual problem, we note that (4.45) and (4.46) are invertible whenever $y \neq 0$; this yields

(4.50) $$x = -2p/|q|^2, \qquad y = 2q|p|^2/|q|^4,$$

so we are in the particularly simple case where the Legendre transformation is one-to-one. Equations (4.48) and (4.47) become

(4.51) $$p = \operatorname{div} q \in L^2,$$

(4.52) $$2(q|p|^2/|q|^2 + \operatorname{grad} \operatorname{div} q)/|q|^2 = 0.$$

But this means exactly that $q \neq 0$ is a critical point of the function $q \to -|\operatorname{div} q|^2/|q|^2$ over the space

(4.53) $$H(\Omega; \operatorname{div}) = \{q \in L^2(\Omega)^n | \operatorname{div} q \in L^2(\Omega)\}.$$

Finally, we get the dual problem

($\mathcal{P}^*$) $$\operatorname{ext} -|\operatorname{div} q|^2/|q|^2, \qquad q \in H(\Omega; \operatorname{div})$$

with the usual relationship (4.45)–(4.46) or (4.50). Note in particular that

(4.54) $$\{\operatorname{ext} \mathcal{P}\} = -\{\operatorname{ext} \mathcal{P}^*\}.$$

5. Comments. The notion of a Lagrangian submanifold is central to the theory of Fourier integral operators. It is attributed to V. Arnold [1] or V. Maslov [13], and has been painstakingly investigated [11, especially § 3], [16], [9]. However, these authors define a Lagrangian submanifold of a symplectic manifold (dimension $2n$, fundamental 2-form Ω) as an n-dimensional submanifold on which Ω pulls back to zero. In our framework, this would mean an n-dimensional submanifold of $\mathbb{R}^n \times \mathbb{R}^n$ on which $\hat{\Omega} = \sum_{i=1}^n dp_i \wedge dx_i$ pulls back to zero. Noting $\omega = dz - \sum_{i=1}^n p_i \, dx_i$ as in (1.3), we see that $\Omega = d\omega$. It follows that if $V \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ is a Lagrangian submanifold in the sense of Definition 1.1, if the projection $\pi_{xp}: V \to \mathbb{R}^n \times \mathbb{R}^n$ is proper and if its tangent map $T\pi_{xp}: TV \to T(\mathbb{R}^n \times \mathbb{R}^n)$ has rank n everywhere, then $\pi_{xp}V$ is a Lagrangian submanifold of $\mathbb{R}^n \times \mathbb{R}^n$ in the preceding sense. Our definition has the advantage of incorporating z, which is very useful for practical purposes.

For basic information about proper maps, we refer to any book on general topology, e.g. [4, Chaps. 1 and 2]. Sard's theorem in the C^∞ case, as well as basic information on submanifolds and the implicit function theorem, can be found in [12].

The definition (2.1) of the Legendre transformation is given in [6] as a particular case of a contact transformation. The contact transformation associated with a given C^∞ function $H(x, z; x', z')$ on $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ is the mapping which associates with any point $(x, p, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ the point (x', p', z') defined by the formulae

$$H(x', z'; x, z) = 0,$$
$$\partial H/\partial x' + p' \, \partial H/\partial z' = 0,$$
$$\partial H/\partial x + p \, \partial H/\partial z = 0.$$

From the two last equations it follows (formally) that $p=\partial z/\partial x$ and $p'=\partial z'/\partial x'$. It follows (still formally) from the first one that $dz'+p'\,dx'=0$ if and only if $dz+p\,dx=0$. In other words, if we have no trouble with cusps or closedness, a contact transformation will send a Lagrangian manifold onto a Lagrangian manifold. It need not be involutive. In the special case where $H(x',z';x,z)=z+z'-xx'$, we get the Legendre transformation.

Also related to the Legendre transform is the notion of dual varieties in algebraic geometry. Let a projective variety C be given by its equation $P(X_1,\cdots,X_n)=0$, where P is a homogeneous polynomial of degree d. The dual variety, $\hat{C}$ is the set of tangents to C; its equation $\hat{P}(u_1,\cdots,u_n)=0$ has as zeros all $(u_1,\cdots,u_n)$ such that the hyperplane $u_1X_1+\cdots+u_nX_n$ is tangent to C. In particular, $\hat{\hat{C}}=C$. For instance, if $f\colon\mathbb{R}^n\to\mathbb{R}$ is a polynomial, setting

$$z=\frac{X_{n+1}}{X_{n+2}},\qquad x_i=\frac{X_i}{X_{n+2}}$$

as is usual in projective geometry yields

$$\operatorname{graph} f=\left\{(X_1,\cdots,X_{n+2})\,\middle|\,\frac{X_{n+1}}{X_{n+2}}=f\left(\frac{X_1}{X_{n+2}},\cdots,\frac{X_n}{X_{n+2}}\right)\right\}.$$

The dual variety is simply the graph of the Legendre transform

$$\widehat{\operatorname{graph} f}=\operatorname{graph}\mathscr{L}f.$$

A particularly interesting case arises when $n=1$ and complex numbers are used. It can be shown that if C (resp. $\hat{C}$) is a complex algebraic curve of degree d (resp. $\hat{d}$), having r (resp. $\hat{r}$) double points and s (resp. $\hat{s}$) cusps, with no other singularities, then we have the following symmetric relationship (Plücker's formulae)

$$\hat{d}=d(d-1)-2r-3s,$$
$$d=\hat{d}(\hat{d}-1)-2\hat{r}-3\hat{s},$$
$$\hat{s}-s=3(\hat{d}-d).$$

(I am indebted to P. Deligne for this elementary algebraic geometry.)

Now let us proceed to providing §§ 2, 3, 4 with bibliographical references.

Fundamentals of convex analysis are given in [14] or [8]. Modern tools of different topology, included the Malgrange division theorem, Thom's transversality theorem and notions on stratifications, will be found in [15]; see [10] for a textbook on the subject. Note that the proof of Proposition 2.7 for $n=1$ does not require the C^∞ division theorem.

Condition (3.7) can be interpreted as a necessary condition for optimality in a much broader context than indicated, i.e. the space need not be finite-dimensional and the g_i' need not be linearly independent; see [7]. Duality theory for finite-dimensional convex optimization problems will be found in [14].

Theorem 4.1 is due to J.-P. Aubin. Its proof will be found in [2] or [3]. Duality theory for convex problems in the calculus of variations is treated in [8], but here we follow rather the approach of [3]. Proposition 4.5 is a nonconvex analogue of [5].

Acknowledgments. I am indebted to R. Temam for suggesting to me the eigenvalue examples concluding §§ 3 and 4. Also I wish to acknowledge long and Numerous conversations with J.-P. Aubin and F. Clarke, and the expert typing of Mrs. Sally Ross.

REFERENCES

[1] V. I. ARNOLD, *Characteristic class entering in quantization conditions*, J. Functional Anal. Appl., 1 (1967), pp. 1–13.

[2] J.-P. AUBIN, *Approximation of Elliptic Boundary-value Problems*, John Wiley, New York, 1972.

[3] ———, *Mathematical methods of game and economic theory*, North-Holland Elsevier, Amsterdam, to appear.

[4] N. BOURBAKI, *Topologie Générale*, 2éme edition, Hermann, Paris, 1960.

[5] H. BREZIS AND I. EKELAND, *Un principe variational associé à certaines équations paraboliques*, C.R. Acad. Sci. Paris, Sér. A-B, 282 (1976), pp. 971–974 and 1197–1198.

[6] C. CARATHEODORY, *Variations rechnung und partielle differentialgleichungen erster Ordnung*, Teubner, Leipzig, 1935.

[7] F. CLARKE, *A new approach to Lagrange multipliers*, Mathematics of Operations Res., 1 (1976).

[8] I. EKELAND AND R. TEMAM, *Convex Analysis and Variational Problems*, North-Holland Elsevier, Amsterdam, 1975.

[9] J. GUCKENHEIMER, *Catastrophes and partial differential equations*, Ann. Inst. Fourier (Grenoble), 23 (1973), 2, pp. 31–59.

[10] M. GOLUBISTSKY AND V. GUILLEMIN, *Stable Mappings and Their Singularities*, Springer-Verlag, Berlin, 1973.

[11] L. HÖRMANDER, *Fourier integral operators I*, Acta Math., 127 (1971), pp. 79–183.

[12] S. LANG, *Differentiable Manifolds*, Addison-Wesley, Reading, MA, 1972.

[13] V. MASLOV, *Theory of Perturbations and Asymptotic Methods*, Moskov. Gos. Univ., Moscow, 1965. (In Russian.)

[14] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1969.

[15] WALL, ed., *Proceeding of the Liverpool Singularities Symposium I*, Springer Lecture Notes in Mathematics 192, Springer-Verlag, Berlin, 1971.

[16] A. WEINSTEIN, *Symplectic manifolds and their Lagrangian submanifolds*, Advances in Math., 6 (1971), pp. 329–346.