INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392 I

SMR/92 - 21

AUTUMN COURSE

ON

VARIATIONAL METHODS IN ANALYSIS AND MATHEMATICAL PHYSICS

20 October - 11 December 1981

INTRODUCTION TO THE FENCHEL TRANSFORM

M. WILLEM
Institut Mathematique
2 Chemin du Cyclotron 2
B-1348 Louvain-la-Neuve
Belgium

INTRODUCTION TO THE FENCHEL TRANSFORM

M. Willem, Université Catholique de Louvain.

Introduction.

The basic idea of the LEGENDRE transform is to describe a smooth hypersurface

$$s = \varphi(u) \qquad s \in \mathbb{R},\ u \in \mathbb{R}^n$$

either as a set of points or as an envelope of tangent hyperplanes. The FENCHEL transform is an extension of this idea to (not necessarily smooth) convex functions.

Notation.

Every vector space will be real. For any normed space X, X^* will denote the dual of X, i.e. the space of all continuous linear functions from X into $\mathbb{R}$.
We will use the convention $0 \cdot \infty = \infty \cdot 0 = 0$.
If $(\varphi_i : X \to]-\infty, \infty])_{i \in I}$ is a family of functions

$$(\sup \varphi_i)(u) = \sup_i \varphi_i(u).$$

If φ is a function on a set S_1, $\varphi|_{S_2}$ is the restriction of φ to a subset S_2 of S_1.
If X is a normed space and if $f \in X^*$,

$$\langle f, u \rangle_{X^*, X} = \langle u, f \rangle_{X, X^*} = f(u).$$

Convex sets and functions.

Definition 1. A subset of a vector space X is convex when $\forall u, v \in C$, $\forall \lambda \in]0,1[$, $(1-\lambda)u + \lambda v \in C$.
A function $\varphi : X \to]-\infty, \infty]$ is convex if
$\forall u, v \in X$, $\forall \lambda \in]0,1[$ $\quad \varphi((1-\lambda)u + \lambda v) \le (1-\lambda)\varphi(u) + \lambda\varphi(v)$.

Remarks. 1. The effective domain of a function $\varphi : X \to]-\infty, \infty]$ is the set

$$D(\varphi) = \{ u \in X : \varphi(u) < \infty \}.$$

It is clear that, if φ is convex, $D(\varphi)$ is convex.

2. The epigraph of a function $\varphi : X \to]-\infty, \infty]$ is the set

$$\operatorname{epi} \varphi = \{ (u, \mu) \in X \times \mathbb{R} : \varphi(u) \le \mu \}.$$

A function φ is convex if and only if $\operatorname{epi} \varphi$ is a convex subset of $X \times \mathbb{R}$.

The following properties are easily verified:

Proposition 1. (i) The product of a convex function by a nonnegative real is convex.
(ii) The sum of two convex functions is convex.
(iii) The sup of a non empty family of convex functions is convex.

The following (equivalent) formulations of the HAHN-BANACH theorem are the basic tools of convexity theory.

Theorem A. (Analytic form) Let V be a subspace of a vector space X, let $l : V \to \mathbb{R}$ be linear and let $p : X \to \mathbb{R}$ be subadditive, positively homogeneous and such that $l \leq p|_M$. Then there exists a linear function $\tilde{l} : X \to \mathbb{R}$ such that $\tilde{l} \leq p$ on X and $\tilde{l}|_V = l$.

Theorem B. (Geometric form) Let C and D be convex subsets of a topological vector space such that $\operatorname{int} D \neq \emptyset$ and $C \cap \operatorname{int} D = \emptyset$. Then there is a closed hyperplane separating C and D.

4. Differentiable convex functions.

Definition 2. A real function φ on a normed space X is differentiable (or Gâteaux-differentiable) if for every $u \in X$ there exists $D\varphi(u) \in X^*$ such that

$$\forall v \in X \quad \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{\varphi(u+tv) - \varphi(u)}{t} = \langle D\varphi(u), v \rangle_{X^*,X}.$$

Proposition 2. Let φ be a differentiable function on a normed space X. The following statements are equivalent:

(i) φ is convex

(ii) $\forall u, v \in X$

$$\varphi(v) \geq \varphi(u) + \langle D\varphi(u), v-u \rangle_{X^*,X} \tag{1}$$

Proof. (i) $\Rightarrow$ (ii). By the definition of a convex function, for every $\lambda \in]0,1[$,

$$\frac{\varphi(u + \lambda(v-u)) - \varphi(u)}{\lambda} \leq \varphi(v) - \varphi(u).$$

If $\lambda \downarrow 0$ we obtain (1).

(ii) $\Rightarrow$ (i). It follows from (1) that, for every $\lambda \in]0,1[$, if $w = (1-\lambda)u + \lambda v$,

$$\cancel{\varphi((1-\lambda)u + \lambda v) \geq \varphi(w) + \lambda}$$

$$\varphi(u) \geq \varphi(w) - \lambda \langle D\varphi(w), v-u \rangle \tag{2}$$

$$\varphi(v) \geq \varphi(w) - (1-\lambda) \langle D\varphi(w), u-v \rangle. \tag{3}$$

It suffices to multiply (2) by $(1-\lambda)$ and (3) by λ to obtain by adding

$$(1-\lambda)\varphi(u) + \lambda\varphi(v) \geq \varphi(w). \quad \square$$

Remarks. 1. Inequality (1) was first used by WEIERSTRASS in the calculus of variations without reference to convexity.

2. The function $v \mapsto \varphi(u) + \langle D\varphi(u), v - u \rangle$ is a first order TAYLOR expansion at u.

Definition 3. A critical point of a differentiable function φ is a point u such that $D\varphi(u) = 0$. If u is a critical point, $\varphi(u)$ is a critical value.

Corollary 1. Every critical value of a convex differentiable function is a minimum.

Proof. If u is a critical point of the convex differentiable function φ it follows from (1) that $\forall v \in X$, $\varphi(v) \geq \varphi(u)$. $\square$

Remark. It is easy to prove that every (local) minimum (or maximum) of a differentiable function is a critical value.

The following definition generalizes to notion of increasing mapping from $\mathbb{R}$ into $\mathbb{R}$.

Definition 4. Let X be a normed space. A mapping $T: X \to X^*$ is monotone if $\forall u, v \in X$, $\langle Tu - Tv, u - v \rangle_{X^*, X} \geq 0$.

Corollary 2. If φ is a convex differentiable function then $D\varphi$ is a monotone mapping.

Proof. By (1)

$$\varphi(u) \geq \varphi(v) + \langle D\varphi(v), u - v \rangle$$
$$\varphi(v) \geq \varphi(u) + \langle D\varphi(u), v - u \rangle$$

The result follows by addition. $\square$

Exercise 1. (KAČUROVSKI) Prove the converse.

Corollary 3. (Quasi linearization). If φ is a convex differentiable function $\forall v \in X$, $\varphi(v) = \sup_{u \in X} (\varphi(u) + \langle D\varphi(u), v - u \rangle)$.

Proof. Obvious from (1). $\square$

Corollary 3 shows that a convex differentiable function is the sup of a family of affine continuous function. The FENCHEL duality follows from this observation. An appropriate setting for the theory is a pair of vector spaces in duality.

3. Vector spaces in duality.

Two vector spaces X and Y are in duality if there is a bilinear form $\langle \cdot, \cdot \rangle$ on $X \times Y$ such that

(4) $\forall u \in X \setminus \{0\} \quad \exists v \in Y : \langle u, v \rangle \neq 0$

and

(5) $\forall v \in Y \setminus \{0\} \quad \exists u \in X : \langle u, v \rangle \neq 0.$

Examples. 1. $X = Y = \mathbb{R}^m$ and

$$\langle (x_1, \dots, x_m), (y_1, \dots, y_m) \rangle = \sum_{j=1}^{m} x_j y_j$$

2. $X = Y = H$ when H is a prehilbertian space with inner product $(\cdot, \cdot)$ and

$$\langle u, v \rangle = (u, v).$$

3. X is a normed space, $Y = X^*$ and

$$\langle u, v \rangle = \langle u, v \rangle_{X, X^*}.$$

Exercise 2. Prove (4) for example 3 using theorem R.

In the following X and Y will play the same game.

4. The Fenchel transform.

Let X and Y be vector spaces in duality.

Definition 5. $\Gamma_0(X, Y)$ is the set of functions $\varphi : X \to]-\infty, \infty]$ such that

(a) $\varphi \not\equiv \infty$

(b) There is a (non empty) family $((v_i, \rho_i))_{i \in I} \subset Y \times \mathbb{R}$ such that

$$\varphi = \sup_i (\langle \cdot, v_i \rangle - \rho_i)$$

The definition of $\Gamma_0(Y, X)$ is similar. By proposition 1 the functions of $\Gamma_0(X, Y)$ and $\Gamma_0(Y, X)$ are convex.

Example. By corollary 3 every convex differentiable function on a normed space X belongs to $\Gamma_0(X, X^*)$.

The following lemma is almost self-evident but important.

Lemma 1. If $\varphi \in \Gamma_0(X, Y)$ then

$$\varphi(u) = \sup \{ \langle u, v \rangle - \rho : \varphi \geq \langle \cdot, v \rangle - \rho \}$$

Proof. If $\tilde{\varphi}(u) = \sup \{ \langle u, v \rangle - \rho : \varphi \geq \langle \cdot, v \rangle - \rho \}$ it is clear that $\varphi \geq \tilde{\varphi}$. If $\varphi = \sup_{i \in I} (\langle \cdot, v_i \rangle - \rho_i)$ and if $\varphi(u) > \eta$, then exists $i \in I$ such that

$$\langle u, v_i \rangle - \rho_i > \eta.$$

Thus $\tilde{\varphi}(u) > \eta$. But then $\tilde{\varphi}(u) \geq \varphi(u)$. $\square$

Definition 6. If $\varphi \in \Gamma_0(X,Y)$ then $\varphi^* : Y \to]-\infty, \infty]$ is defined by

$$\varphi^*(v) = \sup_{u \in D(\varphi)} (\langle u, v\rangle - \varphi(u))$$

If $\psi \in \Gamma_0(Y,X)$ the definition of ψ^* is similar.

Remarks. 1. Since $\varphi \in \Gamma_0(X,Y)$ there exist $(v, \rho) \in Y \times \mathbb{R}$ such that $\varphi \geq \langle \cdot, v\rangle - \rho$. Then $\varphi^*(v) \leq \rho$ and $\varphi^* \not\equiv \infty$. Thus $\varphi^* \in \Gamma_0(Y,X)$.

2. The FENCHEL inequality

$$\forall u, v \quad \varphi(u) + \varphi^*(v) \geq \langle u, v\rangle$$

follows directly from the definition.

3. It is clear that $\varphi_1 \leq \varphi_2 \Rightarrow \varphi_1^* \geq \varphi_2^*$.

Examples. 1. Let $X = Y = \mathbb{R}$ and let $\varphi(u) = e^u$. Then

$$\begin{aligned} \varphi^*(v) &= \infty && \text{for } v < 0 \\ &= 0 && \text{for } v = 0 \\ &= v(\ln v - 1) && \text{for } v > 0. \end{aligned}$$

2. If $\varphi(u) = \langle u, v\rangle - \rho$ then

$$\varphi^*(v) = \rho.$$
$$\varphi^*(w) = \infty \quad \text{for } w \neq v.$$

Theorem 1. If $\varphi \in \Gamma_0(X,Y)$ then $\varphi^{**} = \varphi$.

Proof. Since

$$\varphi \geq \langle \cdot, v\rangle - \rho \iff \rho \geq \sup_{u \in D(\varphi)} (\langle u, v\rangle - \varphi(u)) = \varphi^*(v)$$

it follows from lemma 1 that

$$\varphi(u) = \sup_{v \in D\varphi^*} (\langle u, v\rangle - \varphi^*(v))$$

i.e. $\varphi(u) = \varphi^{**}(u)$. □

Remarks 1. Theorem 1 is a sharpening of lemma 1.

2. The FENCHEL transform is an involution thus a bijection from $\Gamma_0(X,Y)$ onto $\Gamma_0(Y,X)$.

Definition 7. If $\varphi \in \Gamma_0(X,Y)$ and $u \in X$,
$\partial\varphi(u) = \{ v \in Y : \forall w \in X \quad \varphi(w) \geq \varphi(u) + \langle w - u, v\rangle \}$
is the subdifferential of φ at u. If $\partial\varphi(u) \neq \emptyset$ φ is subdifferentiable at u.

Remarks. 1. It follows from the definition that $\partial\varphi(u)$ is a convex set.

2. Clearly $\partial\varphi(u) \neq \emptyset \Rightarrow u \in D(\varphi)$.

3. The function $w \mapsto \varphi(u) + \langle w - u, v\rangle$ is similar to a first order TAYLOR expansion.

4. By definition $\varphi(u) = \inf_X \varphi \iff 0 \in \partial\varphi(u)$.

5. If $v_i \in \partial\varphi(u_i)$, $i = 1, 2$ we obtain as in corollary 2

$$\langle u_1 - u_2, v_1 - v_2 \rangle \geq 0.$$

Thus $\partial\varphi \subset X \times Y$ is a "monotone multimapping".

Example. If φ is a convex differentiable function on a normed space X then $\partial\varphi = D\varphi$.

Proof. By proposition 1, for every $u \in X$, $D\varphi(u) \in \partial\varphi(u)$.

If $v \in \partial\varphi(u)$ then

$$\varphi(u) - \langle u, v\rangle = \inf_{w \in X} \left(\varphi(w) - \langle w, v\rangle\right).$$

By the first order necessary condition for a minimum $D\varphi(u) - v = 0$. $\square$

Theorem 2. If $\varphi \in \Gamma_0(X,Y)$ the following statements are equivalent:

(α) $v \in \partial\varphi(u)$

(β) $\varphi(u) + \varphi^*(v) = \langle u, v\rangle$

(γ) $u \in \partial\varphi^*(v)$.

Proof. $(\alpha) \Leftrightarrow \forall w \in X \quad \langle w, v\rangle - \varphi(w) \le \langle u, v\rangle - \varphi(u)$

$$\Leftrightarrow \langle u, v\rangle - \varphi(u) = \sup_{w \in X}\left(\langle w, v\rangle - \varphi(w)\right)$$

$\Leftrightarrow (\beta)$.

By theorem 1, $(\beta) \Leftrightarrow \varphi^{**}(u) + \varphi^*(v) = \langle u, v\rangle$.
The equivalence between (β) and (γ) follows then from the equivalence between (β) and (α). $\square$

Remark. If φ is differentiable, we obtain

(α') $v = D\varphi(u)$

(β') $\varphi^*(v) = \langle u, v\rangle - \varphi(u)$

i.e. the implicit definition of the LEGENDRE transform.

Example. (Young duality) Let $\Phi : \mathbb{R}^+ \to \mathbb{R}^+$ be continuous, coercive (i.e. $\Phi(u) \to \infty$, $u \to \infty$), strictly increasing and such that $\Phi(0) = 0$. It is clear that $\Psi = \Phi^{-1}$ has the same properties. If

$$\varphi(u) = \int_0^{|u|} \Phi(s)\, ds$$

$$\psi(u) = \int_0^{|u|} \Psi(s)\, ds$$

then $\varphi, \psi \in \Gamma_0(\mathbb{R}, \mathbb{R})$ and it follows from (α') and (β') that, for $v \ge 0$,

$$v = \Phi(u)$$

$$\varphi^*(v) = uv - \int_0^u \Phi(s)\, ds$$

Thus

$$\varphi^*(v) = \Psi(v)\, v - \int_0^{\Psi(v)} \Phi(s)\, ds$$

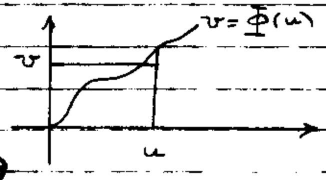
$$= \int_0^v \Psi(s)\, ds$$

$$= \psi(v).$$

Similarly if $v \le 0$, $\varphi^*(v) = \psi(v)$. For $u, v \ge 0$ the FENCHEL inequality is

$$\int_0^u \Phi(s)\, ds + \int_0^v \Psi(s)\, ds \ge uv$$

Geometrically



Exercise 1. Prove that if $p > 1$ and

$$\frac{1}{p} + \frac{1}{q} = 1$$

then, for every $u, v \geq 0$,

$$(*) \qquad \frac{u^p}{p} + \frac{v^q}{q} \geq uv \quad \text{(YOUNG inequality)}$$

2. Derive the HÖLDER inequality from (*).

Example. Let X be a normed space and let $\varphi : X \to \mathbb{R}$ be a continuous convex function. Then $\varphi \in \Gamma_0(X, X^*)$ and is everywhere subdifferentiable.

Proof. Let u be fixed in X. Theorem B applied in $X \times \mathbb{R}$ to

$$C = \{(u, \varphi(u))\} \quad \text{and} \quad D = \text{epi}\, \varphi$$

insures the existence of $v \in X^*$, $c_1, c_2 \in \mathbb{R}$ such that $\|v\| + |c_1| \neq 0$ and

$$(6) \qquad \langle v, u \rangle_{X^*, X} + c_1 \varphi(u) \leq c_2$$

$$(7) \qquad \forall (w, \mu) \in \text{epi}\, \varphi \quad \langle v, w \rangle_{X^*, X} + c_1 \mu \geq c_2$$

In particular, if $\mu \geq \varphi(u)$, $c_1(\mu - \varphi(u)) \geq 0$. Thus $c_1 \geq 0$. If $c_1 = 0$, we obtain

$$\forall w \quad \langle v, w - u \rangle \geq 0$$

i.e. $v = 0$, a contradiction. This shows that $c_1 > 0$. We choose $c_1 = 1$. It follows then from (6) and (7) that

$$\forall w \in H \quad \langle v, w \rangle + \varphi(w) \geq \langle v, u \rangle + \varphi(u)$$

This implies that $\varphi \in \Gamma_0(X, Y)$ and that $v \in \partial \varphi(u)$. □

References.

linear functional analysis

1) DUNFORD and SCHWARTZ, "linear operators", Interscience 1958
2) NIRENBERG "FUNCTIONAL ANALYSIS", Lecture Notes COURANT institute.

Convex analysis

1) EKELAND and TEMAM, "Convex analysis and variational problems" North Holland, 1976.
2) MOREAU "Fonctionnelles convexes", Lectures Notes Collège de France, 1966

Monotone operators

BREZIS, "Opérateurs maximaux monotones", North-Holland

Early duality

COURANT and HILBERT, "Methods of mathematical physics" Interscience 1953, I p. 231-242, 252-257, 268-272.