MATH. SCAND. 36 (1975), 131–136

CONVEXIFICATION OF CONJUGATE FUNCTIONS

ARNE BRØNDSTED

To Werner Fenchel on his 70th birthday.

1. Introduction.

The theory of conjugate convex functions, initiated by W. Fenchel [4], may be described as a duality theory — within the framework of paired linear spaces — for upper envelopes of continuous affine functions. (This description is justified by the fact that these upper envelopes are the lower semi-continuous convex functions with values in $]-\infty, +\infty]$ plus the constant function $-\infty$.) However, as has been pointed out by J. J. Moreau [5], [6] among others, the basic notions of the theory do not require the presence of a linear structure, — a general theory of conjugate functions may be developed for triples (X, Y, p), where X and Y are arbitrary sets and p is a function on $X \times Y$. It is the purpose of the present note to demonstrate that in a sense to be made precise, the duality theory for general triples (X, Y, p) may be embedded in the standard theory for triples (U, V, q), where U and V are real linear spaces and q is a non-degenerate bilinear form on $U \times V$, — provided that p takes only finite values.

2. Convexification.

In the following, let X and Y be arbitrary sets, and let $p \in \mathsf{R}^{X \times Y}$. Following J. J. Moreau [6] we shall call p a *pairing* of X and Y. Also, a function $f \in \overline{\mathsf{R}}^X$ (where $\overline{\mathsf{R}} := [-\infty, +\infty]$) is called *regular* if f is the supremum of functions $p(\cdot, y) - \beta$, where $y \in Y$ and $\beta \in \mathsf{R}$. The set of regular functions on X is denoted $\Gamma(X, Y)$. Regularity of functions $g \in \overline{\mathsf{R}}^Y$ is defined similarly. For any function $f \in \overline{\mathsf{R}}^X$, the *conjugate* $f^p \in \Gamma(Y, X)$ is defined by

$$f^p(y) := \sup_{x \in X} (p(x,y) - f(x)) .$$

The conjugate of a function $g \in \overline{\mathsf{R}}^Y$ is defined similarly. For any $f \in \overline{\mathsf{R}}^X$, the bi-conjugate $f^{pp} := (f^p)^p$ is the largest regular minorant of f, and hence f is regular if and only if $f = f^{pp}$.

Received October 15, 1974; in revised form December 9, 1974.

We shall denote the mapping $x \to p(x,\cdot)$ from X into R^Y by ξ, and the mapping $y \to p(\cdot,y)$ from Y into R^X by η. Note that ξ resp. η is injective if and only if $(p(\cdot,y))_{y\in Y}$ resp. $(p(x,\cdot))_{x\in X}$ separate the points of X resp. Y. Furthermore, we shall denote by U resp. V the linear span of $\xi(X)$ resp. $\eta(Y)$ in R^Y resp. R^X. We then have:

PROPOSITION 1. *For* $u=\sum_{i=1}^m \lambda_i \xi(x_i) \in U$ *and* $v=\sum_{j=1}^n \mu_j \eta(y_j) \in V$, *let*

$$q(u,v) := \sum_{j=1}^n \sum_{i=1}^m \lambda_i \mu_j p(x_i,y_j) .$$

Then q is a well-defined non-degenerate bilinear form on $U \times V$, i.e. U and V are in duality under q. Furthermore, $q(\xi(x),\eta(y))=p(x,y)$ for all $x \in X$, $y \in Y$.

PROOF. A straightforward verification shows that q is well-defined. Clearly, q is bilinear. If $u \in U \setminus \{o\}$, then $u(y_0) \neq 0$ for some $y_0 \in Y$. But then $q(u,v) \neq 0$ with $v := \eta(y_0)$. By the symmetry, this shows that q is non-degenerate. The last statement is obvious.

The preceding proposition describes an embedding of (X,Y,p) in (U,V,q). Denoting by $\sigma(X,Y)$ resp. $\sigma(Y,X)$ the coarsest topology on X resp. Y such that the functions $(p(\cdot,y))_{y\in Y}$ resp. $(p(x,\cdot))_{x\in X}$ are continuous, we have:

PROPOSITION 2. *The mapping ξ is a continuous and open mapping from $(X,\sigma(X,Y))$ onto its image $(\xi(X),\sigma(U,V))$ in $(U,\sigma(U,V))$. In particular, if $(p(\cdot,y))_{y\in Y}$ separate the points of X, then ξ is a homeomorphism. Similarly for η.*

PROOF. From the definition of $\sigma(U,V)$ it follows that ξ is continuous if (and only if) for each $v \in V$ the real valued function $q(\xi(\cdot),v)$ is continuous on $(X,\sigma(X,Y))$. But $q(\xi(x),v)=\sum_{j=1}^n \mu_j p(x,y_j)$ when $v=\sum_{j=1}^n \mu_j \eta(y_j)$, and therefore $q(\xi(\cdot),v)$ is continuous by the definition of $\sigma(X,Y)$.

To see that ξ is open, let $x_0 \in X$ and let $u_0 := \xi(x_0)$. Note that each neighbourhood of x_0 contains a neighbourhood of the form

$$V_{y_1,\ldots,y_n;\alpha}(x_0) := \{x \in X \mid |p(x,y_j)-p(x_0,y_j)| < \alpha,\ j=1,\ldots,n\} ,$$

where $y_1,\ldots,y_n \in Y$ and $\alpha \in \mathsf{R}$. Let $v_j := \eta(y_j)$, $j=1,\ldots,n$, and note that $q(\xi(x),v_j)=p(x,y_j)$ for each $x \in X$ and $j=1,\ldots,n$. This shows that the image of $V_{y_1,\ldots,y_n;\alpha}(x_0)$ under ξ is the set

$$\{u \in \xi(X) \mid |q(u,v_j)-q(u_0,v_j)| < \alpha,\ j=1,\ldots,n\}$$

which is a neighbourhood of u_0 in $(\xi(X),\sigma(U,V))$.

Clearly, when f is function on X with values in $\overline{\mathsf{R}}$, then $f\circ\xi^{-1}$ is a well-defined function on $\xi(X)$ if and only if $f(x_1)=f(x_2)$ for any $x_1,x_2\in X$ such that $p(x_1,\cdot)=p(x_2,\cdot)$. For functions f on X with this property we shall denote by $\xi(f)$ the function on U which equals $f\circ\xi^{-1}$ on $\xi(X)$ and takes the value $+\infty$ on $U\setminus\xi(X)$. For $g\in\overline{\mathsf{R}}^Y$, $\eta(g)$ is defined similarly.

PROPOSITION 3. *Let $f\in\overline{\mathsf{R}}^X$ be lower (or upper) semi-continuous on $(X,\sigma(X,Y))$. Then $\xi(f)$ is well-defined, and its restriction to $(\xi(X),\sigma(U,V))$ is lower (or upper) semi-continuous. Similarly for $g\in\overline{\mathsf{R}}^Y$.*

PROOF. It follows from the definition of $\sigma(X,Y)$ that if $p(x_1,\cdot)=p(x_2,\cdot)$, then x_1 and x_2 belong to the same open sets in $(X,\sigma(X,Y))$. Therefore, if f is semi-continuous, it takes the same value at any such two points, i.e. $\xi(f)$ is well-defined. The semi-continuity properties of $\xi(f)$ follows from Proposition 2.

By Proposition 3, or by a direct argument, $\xi(f)$ is well-defined for each regular function f on X. In order to obtain a representation of f by a regular function on U, we shall "regularize" $\xi(f)$. This regularization may be performed in the usual sense, i.e. by taking the supremum of all $\sigma(U,V)$-continuous affine minorants $q(\cdot,v)-\beta$ of $\xi(f)$. But it may also be performed using only those $q(\cdot,v)-\beta$ for which v belongs to some fixed subset A of V such that $A\supseteq\eta(Y)$. The case where $A=\eta(Y)$ is natural in the sense that regularizing $\xi(f)$ amounts to taking the supremum of those continuous affine functions on U which are extensions of the minorants $p(\cdot,y)-\beta$ of f, — where extension is to be understood in the sense that $q(\cdot,\eta(y))-\beta$ is the extension of $p(\cdot,y)-\beta$.

For $\eta(Y)\subseteq A\subseteq V$, the regularization mentioned above is the function $(\xi(f)^q+\psi_A)^q$, where ψ_A is the indicator function of A, i.e. ψ_A equals 0 on A and $+\infty$ on $V\setminus A$. To verify this, let $u\in U$. Then we have

$$\begin{aligned}&\sup\{q(u,v)-\beta \mid v\in A,\ \beta\in\mathsf{R},\ q(\cdot,v)-\beta\leqq\xi(f)\}\\&= \sup\{q(u,v)-\xi(f)^q(v) \mid v\in A\}\\&= \sup\{q(u,v)-(\xi(f)^q(v)+\psi_A(v)) \mid v\in V\}\\&= (\xi(f)^q+\psi_A)^q(u)\,.\end{aligned}$$

We shall denote this function by $\Xi_A(f)$, i.e.

$$\Xi_A(f)(u) = (\xi(f)^q+\psi_A)^q(u),\ u\in U\,.$$

Similarly, for $g\in\Gamma(Y,X)$ and $\xi(X)\subseteq A\subseteq U$, we let

$$H_A(g)(v) = (\eta(g)^q+\psi_A)^q(v),\ v\in V\,.$$

Note that in fact $\Xi_A(f) \in \Gamma(U, V)$. Also, note that $\Xi_V(f) = \xi(f)^{qq}$ as already mentioned, and that $A_1 \subseteq A_2$ implies $\Xi_{A_1}(f) \leqq \Xi_{A_2}(f)$.

PROPOSITION 4. *Let $\eta(Y) \subseteq A \subseteq V$. Then for any $f \in \Gamma(X, Y)$,*

$$\Xi_A(f)(\xi(x)) = \xi(f)(\xi(x)) = f(x)$$

for all $x \in X$. Similarly with X and Y interchanged.

The proof is a straightforward verification.

This proposition yields the desired representation of regular functions on X by regular functions on U. In fact, for each fixed subset A of V containing $\eta(Y)$, f is represented by $\Xi_A(f)$.

It turns out that q-conjugation of $\xi(f)$ reflects p-conjugation of f as nicely as one could hope for:

PROPOSITION 5. *For any $f \in \Gamma(X, Y)$ we have*

$$\eta(f^p) = \xi(f)^q + \psi_{\eta(Y)} .$$

Similarly for $g \in \Gamma(Y, X)$.

The proof is a straighforward verification.

Using this we obtain the following, which among other things shows that in some sense $\Xi_{\eta(Y)}$ and Ξ_V, as well as $H_{\xi(X)}$ and H_U, are equally natural:

PROPOSITION 6. *For any mutually conjugate functions $f \in \Gamma(X, Y)$ and $g \in \Gamma(Y, X)$ we have*

$$\Xi_{\eta(Y)}(f)^q = H_U(g) ,$$

$$\Xi_V(f)^q = H_{\xi(X)}(g) .$$

3. Remarks.

In the following, let (X, Y, p) and (U, V, q) be as in section 2.

A. Note that the construction in section 2 leads to a notion of "convexity on non-convex sets": A subset C of X might be called (Y,p)-*convex* if $\xi(C) = \xi(X) \cap C'$ for some convex subset C' of U, and a function f on X might be called (Y,p)-*convex* if $\xi(f)$ is well-defined and $\xi(f) = \varphi$

on $\xi(X)$ for some convex function φ on U. — The notion of a (Y,p)-convex set extends the notion of a Φ-convex set in the sense of K. Fan [3].

B. When adopting the convention that $(\pm\infty)+(\mp\infty)=-\infty$, $\Gamma(U,V)$ is closed under addition and multiplication by positive scalars, i.e. $\Gamma(U,V)$ is an "abstract" convex cone. One might ask when $\Gamma(X,Y)$ is also a convex cone, and, if so, to what extent it resembles a subcone of $\Gamma(U,V)$. If $\eta(Y)$ and A are convex cones, then the mapping Ξ_A from $\Gamma(X,Y)$ into $\Gamma(U,V)$ is positively homogeneous and superadditive. Only under strong conditions will Ξ_A be additive, and hence an isomorphism from $\Gamma(X,Y)$ onto a subcone of $\Gamma(U,V)$. — The main result of J. P. Aubin [2] is related to this question, although his point of view is different from ours.

C. By a *convex structure* on X we shall mean a mapping γ from the set

$$\bigcup_{n=1}^{\infty} \{(\lambda_1,\ldots,\lambda_n)\in \mathsf{R}^n \mid \lambda_i>0,\ \textstyle\sum_{i=1}^{n}\lambda_i=1\}\times X^n$$

into the set of non-empty subsets of X. It is easy to see that $\xi(X)$ is a convex subset of U if and only if there exists a convex structure γ on X such that each $p(\cdot,y)$, $y\in Y$, is γ-affine in the sense that

$$p(x,y) = \textstyle\sum_{i=1}^{n}\lambda_i p(x_i,y)$$

for each $x\in\gamma\big((\lambda_1,\ldots,\lambda_n),(x_1,\ldots,x_n)\big)$. — J. P. Aubin [1] studies γ-convex functions on a set equipped with a convex structure. His main result is an easy consequence of the results of section 2.

D. As an illustration, consider the following example. Let X be a compact topological space, let Y be a subspace of $\mathscr{C}_{\mathsf{R}}(X)$ separating the points and containing the constants, and let X and Y be paired by $p(x,\varphi)=\varphi(x)$ for $x\in X$ and $\varphi\in Y$. By section 2 there exists for every $f\in\Gamma(X,Y)$ a lower semi-continuous convex function g on U such that $g=f\circ\xi^{-1}$ on $\xi(X)$. And conversely, due to the particular circumstances, if $h\colon\xi(X)\to\overline{\mathsf{R}}$ admits a lower semi-continuous convex extension to U, then $f:=h\circ\xi$ is regular. Therefore, $\Gamma(X,Y)$ is the set of all lower semi-continuous functions on X if and only if every lower semi-continuous function on $\xi(X)$ may be extended to a lower semi-continuous convex function on U. The compact subsets of R^n having this extension property are the compact sets A such that $A=\operatorname{ext}\operatorname{conv}A$. Therefore, if Y is spanned by finitely many (linearly independent) functions $1,\varphi_1,\ldots,\varphi_n$, then every lower semi-continuous function on X is the supremum of linear combinations $\lambda_0+\lambda_1\varphi_1+\ldots+\lambda_n\varphi_n$, if and only if the mapping $x\to(\varphi_1(x),\ldots,\varphi_n(x))$ maps X homeomorphically onto a compact set $A\subseteq \mathsf{R}^n$ with $A=\operatorname{ext}\operatorname{conv}A$. In particular, X must be homeomorphic to a subset of some R^m. Conversely, if this is the case, then there exist such functions φ_i. In fact, if φ is a homeomorphism from X onto a subset of R^m, and X is not homeomorphic to a subset of R^k with $k<m$, then one may take $\varphi_i=\pi_i\circ\varphi$ for $i=1,\ldots,m$, where π_i is the ith projection in R^m, and take $\varphi_{m+1}=\psi\circ\varphi$, where ψ is any continuous strictly convex function on $\operatorname{conv}\varphi(X)$. The functions $1,\varphi_1,\ldots,\varphi_{m+1}$ will then have the properties described above.

REFERENCES

1. J. P. Aubin, *Characterization of cones of functions isomorphic to cones of convex functions*, 5th Conference on Optimization Techniques, R. Conti and A. Ruberti, editors, Lecture Notes in Computer Science 3, Springer-Verlag, Berlin, Heidelberg, New York 1973, 174–183.
2. J. P. Aubin, *Répresentation d'un cône convexe de fonctions par un cône de fonctions convexes*, Boll. Un. Mat. Ital. (4) 7 (1973), 221–228.
3. K. Fan, *On the Krein-Milman theorem*, Proc. Symp. Pure Math. 7, Convexity, Amer. Math. Soc. 1963, 211–219.
4. W. Fenchel, *On conjugate convex functions*, Canad. J. Math. 1 (1949), 73–77.
5. J. J. Moreau, *Fonctionelles Convexes*, Collège de France, Paris 1966–67.
6. J. J. Moreau, *Inf-convolution, sous-additivité, convexité des fonctions numeriques*, J. Math. Pures Appl. 49 (1970), 109–154.

UNIVERSITY OF COPENHAGEN
DENMARK