SYMMETRIZED SEPARABLE CONVEX PROGRAMMING([1])

BY

L. MCLINDEN

ABSTRACT. The duality model for convex programming studied recently by E. L. Peterson is analyzed from the viewpoint of perturbational duality theory. Relationships with the traditional Lagrangian model for ordinary programming are explored in detail, with particular emphasis placed on the respective dual problems, Kuhn-Tucker vectors, and extremality conditions. The case of homogeneous constraints is discussed by way of illustration. The Slater existence criterion for optimal Lagrange multipliers in ordinary programming is sharpened for the case in which some of the functions are polyhedral. The analysis generally covers nonclosed functions on general spaces and includes refinements to exploit polyhedrality in the finite-dimensional case. Underlying the whole development are basic technical facts which are developed concerning the Fenchel conjugate and preconjugate of the indicator function of an epigraph set.

1. Introduction. One of the most useful model problems for convex optimization is the ordinary convex programming problem, in which a convex function is minimized subject to finitely many convex inequality constraints. The traditional dual approach to this problem involves the so-called Lagrangian duality model, in which solving the dual problem amounts to finding the optimal Lagrange multipliers. In 1972, E. L. Peterson [7] outlined an alternate duality theory for this problem, and various aspects of this model have been developed in a series of related papers [8]–[15]. Peterson's model provides a dual problem having more variables than the Lagrangian dual, but in a sense requiring no suboptimization for the dual objective function evaluations. The model is somewhat more general, too, in that its primal problem has additional structure built in which acts to force the dual to have the same form. That is, it is a "symmetric" duality model for ordinary convex programming. Additive separability plays a large role in the model, at least in a formal way.

In this paper Peterson's model is analyzed from the viewpoint of R. T. Rockafellar's perturbational duality theory (see, e.g. [19] or [21]). The setting is extended from R^n to general spaces, the lower semicontinuity requirements on the functions involved are relaxed, and the extremality or optimality conditions are sharpened. Particular emphasis is placed on comparing the "symmetrized" dual with the usual, Lagrangian dual, and it is shown in a precise way that the entire symmetrized duality model projects onto the traditional Lagrangian duality model. Enroute to this comparison, the standard facts about the Lagrangian model are developed here in a setting slightly more general than previously available in the literature. The case of homogeneous constraint functions is discussed as an illustration, with the present treatment extending recent work of C. R. Glassey [2] on explicit duality for such problems. As a technical byproduct of the present framework, a new existence criterion is established for optimal Lagrange multipliers in the case of R^n. Namely, it is shown essentially that it suffices to have a Slater point for which the polyhedral constraint inequalities are satisfied only weakly; this sharpens slightly Rockafellar's theorem [19, Theorem 28.2], which already handled the important case in which the polyhedral functions are actually affine.

The plan of the paper is as follows. In §2 we establish notation, state the primal, dual and saddlepoint problems associated with the duality model under study, and indicate some of the technical issues treated later concerning the "symmetrized" dual and the Lagrangian dual. In §3 we derive certain technical facts on which the entire development rests. These concern formulas for the Fenchel conjugate and "preconjugate" of the indicator function of an epigraph set, together with the associated subdifferential formulas. The preconjugate result is somewhat novel technically, in that it requires the reverse of the common proof technique of dualizing a result; it is a "predual" type of result. In §4 Peterson's model is derived in broad outline form from a notationally streamlined cone-constrained model. Appeal is freely made here to the various facts concerning Rockafellar's perturbational duality approach. This viewpoint provides the basis for distinguishing quite clearly in §5 between the symmetrized duality model and the Lagrangian duality model. There, a detailed comparison of the two models is made. The focus is on a comparison of the respective dual problems, Kuhn-Tucker vectors, and extremality conditions. This section can be viewed as extending somewhat, to the case of general convex functions, certain ideas introduced by Rockafellar in 1964 [18] and later developed by him for the case in which the functions involved are faithfully convex [20]. In §6 certain projection phenomena noted in §5 are examined further, and it is shown the precise sense in which the entire symmetrized model (consisting of primal, dual and associated saddlepoint problems) projects onto the ordinary duality model.

Throughout, we use rather freely the general definitions and background material concerning convexity which is found in [19], [21].

2. Notation and statement of model. The model we shall study is heavily dependent upon additive separability, which may either be already available naturally or else be induced artificially. In either case it is therefore necessary for us to consider, at least in a formal way, a number of distinct spaces. Thus, let there be given convex functions f_0, f_i, f_j defined on spaces X_0, X_i, X_j, where the indices i and j range over finite (possibly empty) index sets I and J. The spaces may in general be any real linear spaces equipped with locally convex Hausdorff topologies. In particular, they might all be R^n, or Hilbert spaces, or reflexive Banach spaces under the norm topologies. Let the functions take values in $(-\infty, +\infty]$ without being identically $+\infty$, and assume each function has lower semicontinuous hull somewhere finite. (The latter is fulfilled automatically in the case of R^n.)

Throughout, when referring to product spaces we use a convenient and transparent notational abbreviation. Namely, the product space $\times \{X_i | i \in I\}$ will be denoted simply by X^I; its elements, the ordered $|I|$-tuples $(x_i)_{i \in I}$ where $x_i \in X_i$ for each $i \in I$, will be denoted simply by x^I. Similarly for X^J and x^J, R^I and ξ^I, R^J and ξ^J. It is convenient also to let X denote the product space

$$X = X_0 \times X^I \times X^J \times R^I \times R^J.$$

When $J = \varnothing$, for example, as is the case in (2.5) below, we agree to interpret X^J and R^J as the degenerate vector space consisting of just the zero vector. Such trivial "factor" spaces can clearly be carried along at no cost, and are effectively suppressed when the general model is specialized, such as in §§5, 6 for instance.

Now to each f_i associate a function ψ_i on $X_i \times R$ defined by

$$\psi_i(x_i, \xi_i) = \begin{cases} 0 & \text{if } f_i(x_i) + \xi_i \leq 0, \\ +\infty & \text{otherwise,} \end{cases} \tag{2.1}$$

and to each f_j associate a function σ_j on $X_j \times R$ defined by

$$\sigma_j(x_j, \xi_j) = \begin{cases} (f_j\xi_j)(x_j) & \text{if } \xi_j > 0, \\ (f_j0^+)(x_j) & \text{if } \xi_j = 0, \\ +\infty & \text{if } \xi_j < 0. \end{cases} \tag{2.2}$$

To within a minus sign, ψ_i is just the indicator of the epigraph of f_i, and as such carries a complete description of f_i. In the definition of σ_j, the terms $f_j\xi_j$ are the right scalar multiples of f_j, defined by

$$(f_j\xi_j)(x_j) = \xi_j f_j(\xi_j^{-1}x_j).$$

The term $f_j 0^+$ is the recession function of f_j, defined by

$$(f_j 0^+)(x_j) = \sup\{ f_j(x_j' + x_j) - f_j(x_j') | x_j' \in \operatorname{dom} f_j \}.$$

It serves to describe the growth properties or asymptotic nature of f_j. Clearly σ_j carries a complete description of f_j. Less obvious is the fact that σ_j is positively homogeneous and convex in (x_j, ξ_j) jointly. The σ_j's will be seen to play a role dual to that of the ψ_i's. Finally, define on X an additively separable function f by the formula

$$f(x) = f(x_0, x^I, x^J, \xi^I, \xi^J)$$

$$= f_0(x_0) + \sum_I \psi_i(x_i, \xi_i) + \sum_J \sigma_j(x_j, \xi_j). \tag{2.3}$$

The primal problem of the duality model studied in this paper can now be stated:

$$\min\{ f(x) | x \in K \}, \tag{P}$$

where f is given by (2.3) and K is a convex cone of the form

$$K = P \times Q, \quad P \subset X_0 \times X^I \times X^J, \quad Q = \{ (\xi^I, \xi^J) | \xi^I = 0 \}, \tag{2.4}$$

for some given nonempty convex cone P. In other words, (P) is the problem

$$\text{minimize } f_0(x_0) + \sum_I \psi_i(x_i, \xi_i) + \sum_J \sigma_j(x_j, \xi_j)$$

$$\text{subject to } (x_0, x^I, x^J) \in P, \quad \xi^I = 0, \ \xi^J \in R^J,$$

where the ψ_i and σ_j are as in (2.1) and (2.2). Essentially this problem was introduced by Peterson [7] in 1972, and various technical aspects of it are treated in [8]–[15].

Notice that the only way in which the variables in (P) can be coupled, or made dependent on one another, is by means of the cone P. This high degree of formal additive separability makes possible some technical simplifications in deriving certain required formulas and also in carrying out some of the proofs. To illustrate how coupling of variables can be achieved, consider what is probably the most important case, that in which

$$J = \varnothing \quad \text{and} \quad P = \{ (x_0, x^I) | x_0 = x_i, \forall i \in I \}. \tag{2.5}$$

It is immediate from (2.3), (2.1) and (2.4) that (P) is then just the ordinary convex program

$$\min\{ f_0(x_0) | f_i(x_0) \leqslant 0, \forall i \in I \}, \tag{P_0}$$

provided we ignore the presence of the trivial variables ξ_i, which must be zero, and also the fact that there are really $1 + |I|$ copies of the variable x_0. Other special cases of problem (P) are described in [8].

The problem dual to (P) in the symmetrized model involves the Fenchel

conjugate of f,

$$f^*(y) = \sup_x \{\langle x, y \rangle - f(x)\}.$$

Due to additive separability, cf. (2.3), this can be expressed in terms of the conjugates of the individual terms f_0, f_i, f_j making up f:

$$f^*(y) = f^*(y_0, y^I, y^J, \eta^I, \eta^J)$$

$$= f_0^*(y_0) + \sum_I \psi_i^*(y_i, \eta_i) + \sum_J \sigma_j^*(y_j, \eta_j). \tag{2.6}$$

Here the variables y, y_0, y_i, y^I, η^I, etc., range over the spaces Y, Y_0, Y_i, Y^I, R^I, etc., which are dual to the original spaces (or more generally, which are paired in duality with them [1], [3]). The value of a continuous linear functional $y \in Y$ at a vector $x \in X$ is denoted by $\langle x, y \rangle$, etc. In particular, in the case of R^n this bracket notation simply means the usual dot product of two vectors.

The problem dual to (P) can now be stated:

$$- \min\{f^*(y) | y \in K^*\}, \tag{D}$$

where K^* is the cone dual to K (i.e. the negative polar of K). According to (2.4), we have

$$K^* = P^* \times Q^*, \quad Q^* = \{(\eta^I, \eta^J) | \eta^J = 0\}, \tag{2.7}$$

and so by (2.6) this problem is therefore of the form

$$- \text{minimize} \ f_0^*(y_0) + \sum_I \psi_i^*(y_i, \eta_i) + \sum_J \sigma_j^-(y_j, \eta_j)$$

$$\text{subject to} \ (y_0, y^I, y^J) \in P^*, \quad \eta^I \in R^I, \eta^J = 0.$$

It is not fully apparent from this that (D) really has the same form as (P). It follows, however, from the facts that the functions conjugate to ψ_i and σ_j are given by

$$\psi_i^*(y_i, \eta_i) = \begin{cases} (f_i^* \eta_i)(y_i) & \text{if } \eta_i > 0, \\ (f_i^* 0^+)(y_i) & \text{if } \eta_i = 0, \\ +\infty & \text{if } \eta_i < 0, \end{cases} \tag{2.8}$$

and

$$\sigma_j^*(y_j, \eta_j) = \begin{cases} 0 & \text{if } f_j^*(y_j) + \eta_j \leq 0, \\ +\infty & \text{otherwise,} \end{cases} \tag{2.9}$$

respectively. This shows that (D) is formed from convex functions f_0^*, f_i^*, f_j^* together with finite index sets I, J and convex cones P^*, Q^* in essentially the same manner as was (P), by just interchanging the roles of I and J.

For the important case (2.5), in which (P) becomes essentially the ordinary

convex program (P_0), we have

$$P^* = \left\{ (y_0, y^I) | y_0 + \sum_I y_i = 0 \right\}. \tag{2.10}$$

This means that (D) assumes the form

$$\max \left\{ -f_0^*(y_0) - \sum_I (f_i^* \eta_i)(y_i) | y_0 + \sum_I y_i = 0 \text{ and } \eta^I \geq 0 \right\}, \tag{2.11}$$

provided the terms $f_i^* \eta_i$ are interpreted as $f_i^* 0^+$ whenever $\eta_i = 0$. We shall present in §5 a mild condition under which this problem can be simplified significantly, namely by interpreting the terms $f_i^* \eta_i$ as $\psi_{\{0\}}$ whenever $\eta_i = 0$. On the other hand, the dual problem associated to (P_0) by the ordinary, i.e. Lagrangian, model is

$$\max_{\eta^I \geq 0} \left\{ \inf_{x_0 \in C} \left\{ f_0(x_0) + \sum_I \eta_i f_i(x_0) \right\} \right\}, \tag{D_0}$$

where $C = \text{dom} f_0 \cap \cap_I \text{dom} f_i$. It will be shown, among other things, that the optimal values in problems (D) and (D_0) always satisfy $\text{val}(D) \leq \text{val}(D_0)$, where (D) here is understood to be the problem (2.11). Under relatively weak assumptions which vary slightly with the situation being treated, this inequality is actually an equality, and furthermore a vector η^I solves (D_0) if and only if there exist vectors y_0, y^I such that (y_0, y^I, η^I) solves (D), i.e. (2.11). In other words, under mild conditions the ordinary dual (D_0) is the image of the "symmetrized" dual (2.11) under the projection transformation $(y_0, y^I, \eta^I) \to \eta^I$. Peterson [14] has also observed essentially this connection between the two duals. The connection was also indicated earlier by Rockafellar rather implicitly in [19, p. 322] and explicitly in [20], where it forms a key element in the computational approach suggested there.

In 1964 Rockafellar [18, p. 88, Theorem 5] established a strong duality theorem relating (P_0) to problem (2.11). He invoked a Slater-type constraint qualitication and worked with constraint functions which were continuous and everywhere finite; a refinement to cover affine constraints was included. Also, in [19, pp. 322–323] and [20] Rockafellar generated essentially problem (2.11) as a dual to (P_0) via the general perturbational duality theory. This involved equipping (P_0) with a certain rather full class of perturbations. In the development below we exploit the perturbational duality approach fully in deriving the symmetrized model. By this means it will be clear that the present, symmetrized model stems from the same, full class of perturbations, but with the added elements of separability and symmetry built in as well.

Underpinning the entire symmetrized model are the conjugacy and subdifferential formulas for the functions ψ_i and σ_j defined in (2.1) and (2.2). The conjugacy formulas, (2.8) and (2.9), have already been invoked in stating

the dual problem, and the subdifferential formulas will be essential in developing the model's extremality conditions. We turn now to the task of deriving the specific facts which will be needed.

3. The indicator of an epigraph: its preconjugate and conjugate. To a large extent, the material in this section builds on facts and techniques concerning recession vectors contained in [17], [19]. Proposition 2 is somewhat novel in that it amounts to a "predual" version of Proposition 1, as opposed to the usual dual type of result. The reader can skip over the proofs in this section on first reading.

LEMMA 1. *If f is a proper convex function with* lsc *f somewhere finite, then* $\psi^*_{\mathrm{dom}\,f} = f^*0^+$.

PROOF. For f actually closed, this was established in [17, Corollary 3C(d)], while for f not necessarily closed but the underlying space Euclidean, the result was established in [19, Theorem 13.3]. The latter proof really carries over to the general situation, as we now demonstrate. Since f^* is by definition the pointwise supremum of the family of continuous affine functions of the form $h(y) = \langle x, y \rangle - \xi$, where $(x, \xi) \in \mathrm{epi}\,f$, it follows that epi f^* is the (nonempty) intersection of the corresponding closed halfspaces epi h. Hence, $\mathrm{epi}(f^*0^+) = 0^+(\mathrm{epi}\,f^*)$ is the recession cone of this intersection. By [17, Theorem 2A(b)], it follows that

$$0^+ \cap \{\mathrm{epi}\,h | (x, \xi) \in \mathrm{epi}\,f\} = \cap \{0^+\mathrm{epi}\,h | (x, \xi) \in \mathrm{epi}\,f\}.$$

Now for $h(y) = \langle x, y \rangle - \xi$ it is clear that

$$0^+\mathrm{epi}\,h = \mathrm{epi}\,h0^+ = \mathrm{epi}\langle x, \cdot \rangle.$$

Combining these facts, we obtain

$$\mathrm{epi}(f^*0^+) = \cap \{\mathrm{epi}\langle x, \cdot \rangle | (x, \xi) \in \mathrm{epi}\,f\}$$
$$= \mathrm{epi}(\sup\{\langle x, \cdot \rangle | x \in \mathrm{dom}\,f\}) = \mathrm{epi}(\psi^*_{\mathrm{dom}\,f}),$$

which completes the proof.

We now deal with the functions of type ψ_i, given in (2.1). Our first result establishes formula (2.8) and more.

PROPOSITION 1. *Let f be a proper convex function with* lsc *f somewhere finite, and write* $E = \{(x, \xi) | f(x) + \xi \leqslant 0\}$. *Then* cl ψ_E *is the indicator function of* $\{(x, \xi) | (\mathrm{cl}\,f)(x) + \xi \leqslant 0\}$, *and*

$$\psi^*_E(y, \eta) = \begin{cases} (f^*\eta)(y) & \text{if } \eta > 0, \\ (f^*0^+)(y) & \text{if } \eta = 0, \\ +\infty & \text{if } \eta < 0. \end{cases}$$

Hence, one has the inequality

$$(f^*\eta)(y) \geqslant \langle x, y \rangle + \xi\eta \quad whenever\ f(x) + \xi \leqslant 0 \leqslant \eta,$$

where $f^*\eta$ is interpreted as f^*0^+ when $\eta = 0$. Moreover, equality occurs in this inequality if and only if $(y, \eta) \in \partial\psi_E(x, \xi)$, and this is equivalent to the conditions

$$y \in \partial(\eta f)(x), \quad f(x) + \xi \leqslant 0 \leqslant \eta, \quad (f(x) + \xi) \cdot \eta = 0,$$

where ηf is interpreted as $\psi_{\mathrm{dom} f}$ when $\eta = 0$. (Note: $\partial(\eta f)(x) = \eta\partial f(x)$ when $\eta > 0$.)

PROOF. First, cl $\psi_E = \psi_{\mathrm{cl}\,E}$. Since cl(epi f) = epi(cl f), it follows that cl E has the form asserted. Next, we compute $\psi_E^*(y, \eta)$. If $\eta < 0$, then using any $(\bar{x}, \bar{\xi}) \in E$ we obtain

$$\psi_E^*(y, \eta) \geqslant \sup_{\xi < \bar{\xi}} \{\langle \bar{x}, y \rangle + \xi\eta\} = +\infty.$$

If $\eta = 0$, we have

$$\psi_E^*(y, 0) = \sup\{\langle x, y \rangle | (x, \xi) \in E\}$$
$$= \sup\{\langle x, y \rangle | x \in \mathrm{dom}\, f\} = (f^*0^+)(y),$$

where the last equality is by Lemma 1. If $\eta > 0$, we have

$$\psi_E^*(y, \eta) = \sup\{\langle x, y \rangle + \xi\eta | f(x) + \xi \leqslant 0\} = \sup_x\{\langle x, y \rangle - \eta f(x)\}$$
$$= \eta \cdot \sup_x\{\langle x, \eta^{-1}y \rangle - f(x)\} = (f^*\eta)(y).$$

This establishes the conjugacy formula. Using it, Fenchel's inequality

$$\psi_E(x, \xi) + \psi_E^*(y, \eta) \geqslant \langle x, y \rangle + \xi\eta$$

for the function ψ_E reduces to the inequality asserted. Moreover, we know that the case of equality is characterized by the condition $(y, \eta) \in \partial\psi_E(x, \xi)$ on the variables involved. By the particular form of ψ_E and ψ_E^*, this is equivalent to having either

$$f(x) + \xi \leqslant 0 < \eta \quad \text{and} \quad (f^*\eta)(y) = \langle x, y \rangle + \xi\eta \qquad \text{(a)}$$

or else

$$f(x) + \xi \leqslant 0 = \eta \quad \text{and} \quad (f^*0^+)(y) = \langle x, y \rangle. \qquad \text{(b)}$$

Now we claim that (a) is equivalent to

$$f(x) + \xi = 0 < \eta \quad \text{and} \quad y \in \partial(\eta f)(x), \qquad \text{(a$'$)}$$

and (b) is equivalent to

$$f(x) + \xi \leqslant 0 = \eta \quad \text{and} \quad y \in \partial\psi_{\mathrm{dom} f}(x). \qquad \text{(b$'$)}$$

Indeed, suppose (a) holds. Since $\eta > 0$, the easy identity $(\eta f)^* = f^*\eta$ holds. From this and the definition of $(\eta f)^*(y)$ follows $\langle x, y \rangle - \eta f(x) \leqslant (f^*\eta)(y)$.

Combined with the rest of the information in (a), this yields

$$f(x) + \xi = 0.$$

Now using this, we can rewrite the equation in (a) as

$$\eta f(x) + (f^*\eta)(y) = \langle x, y \rangle,$$

which, in view of the identity cited earlier, is equivalent to $y \in \partial (\eta f)(x)$. The converse implication, starting from (a′), follows by reversing the last part of the argument just given. Now suppose (b) holds. Then $x \in \text{dom } f$, so that

$$\psi_{\text{dom } f}(x) + (f^*0^+)(y) = \langle x, y \rangle.$$

Since $\psi^*_{\text{dom } f} = f^*0^+$ by Lemma 1, this means $y \in \partial\psi_{\text{dom } f}(x)$. This argument, too, reverses to establish the converse. To complete the proof, just observe that the disjunction (a′) or (b′) is equivalent to the characterization we needed to prove.

We turn now to functions of type σ_j, given by (2.2). When f_j is closed, the basic information needed later for the symmetrized model can be obtained by "dualizing" Proposition 1 in the usual manner (i.e. by applying it to f^* in place of f and using $f^{**} = \text{cl } f = f$). In order to cover f_j's which are not necessarily closed, however, we now derive another, more general result, one somewhat parallel to Proposition 1. It covers the application to the σ_j's arising from nonclosed f_j's as in (2.2) and also a bit more. It may be useful in applying the symmetrized model to situations in which the recession functions f_j0^+ are not readily available.

PROPOSITION 2. *Let f be a proper convex function with* lsc f *somewhere finite, and let h be any function (not necessarily convex) satisfying* (cl $f)0^+ \leqslant h \leqslant \psi_{\{0\}}$. *(The choices $h = \psi_{\{0\}}$ and $h = f0^+$ are the main ones.) Then the (not necessarily convex) function σ defined by*

$$\sigma(x, \xi) = \begin{cases} (f\xi)(x) & \text{if } \xi > 0, \\ h(x) & \text{if } \xi = 0, \\ +\infty & \text{if } \xi < 0, \end{cases}$$

satisfies

$$\sigma^* = \psi_E, \quad \text{where } E = \{(y, \eta) \,|\, f^*(y) + \eta \leqslant 0\},$$

and in fact

$$(\text{lsc } \sigma)(x, \xi) = \psi^*_E(x, \xi) = \begin{cases} ((\text{cl } f)\xi)(x) & \text{if } \xi > 0, \\ ((\text{cl } f)0^+)(x) & \text{if } \xi = 0, \\ +\infty & \text{if } \xi < 0, \end{cases}$$

where lsc σ *is closed proper convex. Hence, one has the inequality*

$$(f\xi)(x) \geqslant \langle x, y \rangle + \xi\eta \quad \text{whenever } f^*(y) + \eta \leqslant 0 < \xi,$$

where $f\xi$ is interpreted as h when $\xi = 0$. Equality occurs in this inequality if and only if $(y, \eta) \in \partial\sigma(x, \xi)$, and this is equivalent to the conditions

$$y \in \partial(f\xi)(x), \quad f^*(y) + \eta \leqslant 0 \leqslant \xi, \quad (f^*(y) + \eta) \cdot \xi = 0,$$

where again $f\xi$ is interpreted as h when $\xi = 0$. Moreover, the function σ is convex if and only if h is convex and satisfies $\psi_{\{0\}} \geqslant h \geqslant f0^+$. (Thus, the choice $h = f0^+$ is the least which will make σ convex.) Finally, in the event that $X = R^n$, the function σ is polyhedral convex if and only if f is polyhedral convex and $h = f0^+$.

PROOF. First we show why h can be taken to be $f0^+$, i.e. why $(\text{cl } f)0^+ \leqslant f0^+$. This is equivalent to epi $f0^+ \subset$ epi$(\text{cl } f)0^+$. Since

$$\text{epi } f0^+ = 0^+ \text{ epi } f, \quad \text{epi}(\text{cl } f)0^+ = 0^+ \text{ epi cl } f = 0^+ \text{ cl epi } f,$$

the question reduces to whether $0^+ S \subset 0^+(\text{cl } S)$ holds for a nonempty convex set S. But this follows from [17, Theorem 2A(b)] (or [19, Theorem 8.3] in the Euclidean case). Next, we work towards the formulas for lsc σ and σ^*. To this end, let σ_1 denote the "sigma" function corresponding to the choice $h = \psi_{\{0\}}$, let σ correspond to an arbitrary h of the sort specified in the hypotheses, and finally let σ_2 denote the asserted form of lsc σ. We have $\sigma_2 \leqslant \sigma \leqslant \sigma_1$ by these definitions (since cl $f \leqslant f$), so that

$$\text{lsc } \sigma_2 \leqslant \text{lsc } \sigma \leqslant \text{lsc } \sigma_1, \quad \sigma_1^* \leqslant \sigma^* \leqslant \sigma_2^*.$$

From Proposition 1 we have

$$-\infty < \text{lsc } \sigma_2 = \sigma_2 = \psi_E^*, \quad \sigma_2^* = \psi_E$$

for $E = \{(y, \eta) \mid f^*(y) + \eta \leqslant 0\}$. We also have that

$$\sigma_1^*(y, \eta) = \sup_{x, \xi} \{\langle x, y \rangle + \xi\eta - \sigma_1(x, \xi)\} = \max\{0, \alpha\},$$

where

$$\alpha = \sup_{\xi > 0, x} \{\langle x, y \rangle + \xi\eta - (f\xi)(x)\}$$

$$= \sup_{\xi > 0} \sup_{x} \{\langle x, y \rangle + \xi\eta - \xi f(\xi^{-1}x)\}$$

$$= \sup_{\xi > 0} \left\{\xi \cdot \left(\sup_{x}\{\langle x, y \rangle - f(x)\} + \eta\right)\right\}$$

$$= \sup_{\xi > 0} \{\xi \cdot (f^*(y) + \eta)\} = \psi_E(y, \eta),$$

so that $\sigma_1^* = \psi_E = \sigma_2^*$. Since lsc $\sigma_1 > -\infty$, it follows that lsc $\sigma_1 = \text{cl } \sigma_1 = \sigma_1^{**} = \psi_E^* = \sigma_2$. In view of the above inequalities bounding lsc σ and σ^*, it follows that

$$\text{lsc } \sigma = \psi_E^*, \quad \sigma^* = \psi_E.$$

Next, using the formula $\sigma^* = \psi_E$, we can write Fenchel's inequality for the

function σ as

$$\sigma(x, \xi) + \psi_E(y, \eta) \geqslant \langle x, y \rangle + \xi\eta.$$

This is the general inequality asserted. Equality occurs in it if and only if $(y, \eta) \in \partial\sigma(x, \xi)$. Taking into account the special form of σ and ψ_E, this is equivalent to having either

$$f^*(y) + \eta \leqslant 0 < \xi \quad \text{and} \quad (f\xi)(x) = \langle x, y \rangle + \xi\eta \qquad \text{(a)}$$

or else

$$f^*(y) + \eta \leqslant 0 = \xi \quad \text{and} \quad h(x) = \langle x, y \rangle. \qquad \text{(b)}$$

By an argument similar to that used in Proposition 1, one can verify that (a) is equivalent to

$$f^*(y) + \eta = 0 < \xi \quad \text{and} \quad y \in \partial (f\xi)(x). \qquad \text{(a')}$$

(One uses the identity $(f\xi)^* = \xi f^*$.) Now consider (b). From $(\text{cl } f)0^+ \leqslant h \leqslant \psi_{\{0\}}$ it follows that

$$\psi_Y = \psi_{\{0\}}^* \leqslant h^* \leqslant ((\text{cl } f)0^+)^* = \psi_{\text{cl}(\text{dom } f^*)},$$

where the last equality is by Lemma 1 applied to f^*. Hence, h^* is identically zero on $\text{cl}(\text{dom } f^*)$. From this fact and the observation that $f^*(y) + \eta \leqslant 0$ forces $y \in \text{dom } f^*$, it follows that (b) is the same as

$$f^*(y) + \eta \leqslant 0 = \xi \quad \text{and} \quad y \in \partial h(x). \qquad \text{(b')}$$

Hence, $(y, \eta) \in \partial\sigma(x, \xi)$ is equivalent to having either (a') or (b') hold, which is essentially the assertion.

Now we tackle the convexity characterization. Since σ is convex on $X \times R$ if and only if its restriction to each line in $X \times R$ is convex, let us examine the behavior of the one-dimensional function

$$\sigma_L(\lambda) = \sigma(x_\lambda, \xi_\lambda), \qquad -\infty < \lambda < +\infty,$$

on lines L consisting of points of the form

$$(x_\lambda, \xi_\lambda) = (1 - \lambda)(x_0, \xi_0) + \lambda(x_1, \xi_1)$$

for distinct pairs (x_0, ξ_0) and (x_1, ξ_1) in $X \times R$. Each such L can be regarded as being of one of three types: (i) $\xi_0 = \xi_1 = 0$; (ii) $\xi_0 = \xi_1 \neq 0$; and (iii) $\xi_0 = 0$ and $\xi_1 = 1$. The convexity of $\sigma_L(\lambda)$ for all lines L of type (i) is clearly equivalent to convexity of h on X. Next, consider the function

$$\sigma_1(x, \xi) = \begin{cases} (f\xi)(x) & \text{if } \xi > 0, \\ \psi_{\{0\}}(x) & \text{if } \xi = 0, \\ +\infty & \text{if } \xi < 0, \end{cases}$$

which is convex because f is (see [19, p. 35]). Since $\sigma_L(\lambda)$ for lines L of type (ii) coincides with the restriction of $\sigma_1(x, \xi)$ to L, it follows that convexity of $\sigma_L(\lambda)$ is automatic on all lines of type (ii). Now consider any fixed line L of

type (iii); that is, let pairs $(x_0, 0)$ and $(x_1, 1)$ be given (notice that $\xi_\lambda = \lambda$ in this case). Now $\sigma_L(\lambda)$ coincides with the restriction of $\sigma_1(x, \xi)$ to L everywhere, *except* perhaps at the point $(x_0, 0)$ corresponding to $\lambda = 0$. Since for such an L the function $\sigma_L(\lambda)$ has the value $+\infty$ for negative λ and is known to be convex for positive λ, it will be convex on all of R if and only if

$$\text{either } \sigma_L(0) \geqslant \lim_{\lambda \downarrow 0} \sigma_L(\lambda) \quad \text{or} \quad \sigma_L(\lambda) = +\infty, \, \forall \lambda > 0.$$

Now the first of these alternatives is the same as

$$h(x_0) \geqslant \lim_{\lambda \downarrow 0} (f\lambda)(x_\lambda),$$

while the second can easily be seen equivalent to

$$x_1 \notin \text{dom } f + (-\infty, 1)x_0.$$

It follows that $\sigma_L(\lambda)$ is convex on *all* lines L of type (iii) *having the same x_0 parameter* if and only if

$$h(x_0) \geqslant \sup\left\{ \lim_{\lambda \downarrow 0} (f\lambda)(x_\lambda) | x_1 \in \text{dom } f + (-\infty, 1)x_0 \right\}.$$

Let $s(x_0)$ denote this supremum. We can rewrite it as

$$s(x_0) = \sup_{x \in \text{dom } f} \sup_{\beta > 0} \left\{ \lim_{\lambda \downarrow 0} (f\lambda)(x_\lambda) | x_1 = x + x_0 - \beta x_0 \right\}$$

$$= \sup\{ t(x_0, x) | x \in \text{dom } f \},$$

where for each $x \in \text{dom } f$

$$t(x_0, x) = \sup_{\beta > 0} \left\{ \lim_{\lambda \downarrow 0} \lambda f\big(x + (\lambda^{-1} - \beta)x_0\big) \right\}$$

$$= \sup_{\beta > 0} \left\{ \lim_{\lambda \uparrow \infty} \lambda^{-1} f\big(x + (\lambda - \beta)x_0\big) \right\}$$

$$= \sup_{\beta > 0} \left\{ \lim_{\tau \uparrow \infty} (\beta + \tau)^{-1} f(x + \tau x_0) \right\}.$$

Now it is not hard to show that

$$\lim_{\tau \uparrow \infty} (\beta + \tau)^{-1} f(x + \tau x_0) = \lim_{\tau \uparrow \infty} \tau^{-1} f(x + \tau x_0)$$

for each $\beta > 0$, $x \in \text{dom } f$, $x_0 \in X$. (To see this, let $\theta(\tau) = f(x + \tau x_0)$ and $\gamma = \lim_{\tau \uparrow \infty} \tau^{-1} \theta(\tau)$. If γ is finite,

$$\lim_{\tau \uparrow \infty} (\beta + \tau)^{-1} \theta(\tau) = \lim_{\tau \uparrow \infty} \left[(\beta + \tau)^{-1} \tau \right] \cdot \left[\tau^{-1} \theta(\tau) \right]$$

$$= \lim_{\tau \uparrow \infty} \left[(\beta + \tau)^{-1} \tau \right] \cdot \lim_{\tau \uparrow \infty} \left[\tau^{-1} \theta(\tau) \right] = 1 \cdot \gamma.$$

If γ is $-\infty$, then $(\beta + \tau)^{-1}\theta(\tau) < \tau^{-1}\theta(\tau)$ implies

$$\lim_{\tau \uparrow \infty} (\beta + \tau)^{-1} \theta(\tau) \leqslant \gamma = -\infty.$$

If γ is $+\infty$, let $M > 0$ be given. We can choose $\bar{\tau} \geqslant \beta$ so that $\tau^{-1}\theta(\tau) \geqslant 2M$ whenever $\tau \geqslant \bar{\tau}$. Since $(\beta + \tau)^{-1}\tau \geqslant 1/2$ for $\tau \geqslant \bar{\tau}$, it follows that

$$(\beta + \tau)^{-1}\theta(\tau) = \left[(\beta + \tau)^{-1}\tau\right] \cdot \left[\tau^{-1}\theta(\tau)\right] \geqslant (1/2) \cdot (2M) = M$$

whenever $\tau \geqslant \bar{\tau}$. This shows $\lim_{\tau\uparrow\infty}(\beta + \tau)^{-1}\theta(\tau) = +\infty$ also, completing the verification.) Using this fact, we have that

$$t(x_0, x) = \lim_{\tau\uparrow\infty} \tau^{-1}f(x + \tau x_0),$$

and hence

$$
\begin{aligned}
s(x_0) &= \sup\left\{\lim_{\tau\uparrow\infty} \tau^{-1}f(x + \tau x_0) \,|\, x \in \mathrm{dom}\, f\right\} \\
&= \sup\left\{\lim_{\tau\uparrow\infty} \tau^{-1}\left[f(x + \tau x_0) - f(x)\right] \,|\, x \in \mathrm{dom}\, f\right\} \\
&= \sup\left\{\sup_{0 < \tau < \infty}\left\{\tau^{-1}\left[f(x + \tau x_0) - f(x)\right]\right\} \,|\, x \in \mathrm{dom}\, f\right\} \\
&= \sup\left\{f(x + x_0) - f(x) \,|\, x \in \mathrm{dom}\, f\right\} = (f0^+)(x_0).
\end{aligned}
$$

It follows that $\sigma_L(\lambda)$ is convex on *all* lines L of type (iii) if and only if $h(x_0) \geqslant f0^+(x_0)$, $\forall x_0 \in X$. Combining this fact with the earlier ones concerning lines of type (i) and (ii) completes the proof of the convexity characterization.

Finally, assume $X = R^n$ and suppose that σ is polyhedral convex. Since σ is clearly proper, σ is therefore closed [**19**, Corollary 19.1.2] and agrees with lsc σ. From the formula already established for lsc σ, it follows that $f\xi = (\mathrm{cl}\, f)\xi$ for each $\xi > 0$ and $h = (\mathrm{cl}\, f)0^+$. The choice $\xi = 1$ in the first fact yields f closed, and so the second fact yields $h = f0^+$. The fact that f is polyhedral is immediate from the fact that epi f is essentially just the intersection of epi σ with the hyperplane

$$\left\{(x, \xi, \mu) \in R^n \times R \times R \,|\, \xi = 1\right\},$$

and hence is a (nonempty) polyhedral convex set. Suppose conversely, now, that f is polyhedral proper convex and $h = f0^+$. Then f is closed [**19**, Corollary 19.1.2], so the formula for lsc σ shows that $\sigma = \mathrm{lsc}\, \sigma$, hence σ is closed proper convex. Therefore $\sigma = \sigma^{**}$, where we know that

$$\sigma^*(y, \eta) = \begin{cases} 0 & \text{if } f^*(y) + \eta \leqslant 0, \\ +\infty & \text{otherwise.} \end{cases}$$

This shows σ^* is, to within a minus sign, the indicator of epi f^*. Since f^* is polyhedral [**19**, Theorem 19.2], this means σ^* is, and hence $\sigma = (\sigma^*)^*$ is also [**19**, Theorem 19.2].

We remark that Proposition 1 for the closed case can be deduced from Proposition 2 by choosing $h = f0^+$ and then dualizing the resulting facts.

4. The symmetrized separable duality model. Here we derive the overall structure and prove the basic facts concerning the general symmetrized model. That is, starting from the primal problem (P) given in §2, we generate (D) and an associated saddlepoint problem. Then we develop the associated Kuhn-Tucker theorem, and provide several criteria for strong duality, including an especially sharp criterion for the finite-dimensional case.

Our approach, as implied in §2, is to think of (P) initially as simply a problem of minimizing a convex function f over a convex cone K, suppressing for the moment the particular structure of f and K. Symmetric duality for this very general form of problem dates back to 1963 (Rockafellar [16], see also [19, Theorem 31.4]). Although one can, of course, regard this as a special case of the original Fenchel duality theorem, for what follows we prefer to present it instead within the broader perspective of perturbational duality. In doing this, we follow the general formulation given in [21]. This has the advantage that many of the general results from [21] can immediately, or at least fairly readily, be translated into the present situation. For this reason, in the development below we limit ourselves to writing down only the most central results for the present model, leaving to the interested reader the task of stating the many other, more special results which can be so obtained.

Let Y denote the space paired with X, and let V be another copy of Y and U be another copy of X. From now on, we regard X as paired with V and U as paired with Y. Now consider the convex function F on $X \times U$ given by

$$F(x, u) = f(x) + \psi_K(x + u), \tag{4.1}$$

and the associated minimization problem

$$\min_x \{ F(x, 0) \}, \quad \text{i.e.} \quad \min_{x \in K} \{ f(x) \}. \tag{4.2}$$

From F form a concave function G on $Y \times V$ by means of the formula

$$G(y, v) = \inf_{x, u} \{ F(x, u) + \langle u, y \rangle - \langle x, v \rangle \}$$

[21, Equation (4.17)]. In view of (4.1), a straightforward computation yields that

$$G(y, v) = -\psi_{K^*}(y) - f^*(y + v), \tag{4.3}$$

where $K^* = \{ v | \langle x, v \rangle \geqslant 0, \ \forall x \in K \}$. The problem dual to (4.2) is now formed using G:

$$\max_y \{ G(y, 0) \}, \quad \text{i.e.} \quad - \min_{y \in K^*} \{ f^*(y) \}. \tag{4.4}$$

The "generalized Lagrangian" saddlepoint problem associated with (4.2) and (4.4) is that which corresponds to the saddle function H on $X \times Y$ given by

the formula

$$H(x, y) = \inf_u \left\{ F(x, u) + \langle u, y \rangle \right\}$$

[21, Equation (4.2)]. Another easy calculation using (4.1) yields that

$$H(x, y) = \begin{cases} f(x) - \psi_{K^*}(y) - \langle x, y \rangle & \text{if } x \in \text{dom } f, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.5}$$

The associated Lagrangian problem (L) is thus

$$\underset{\substack{x \in \text{dom} f \ y \in K^*}}{\text{minimax}} \left\{ f(x) - \langle x, y \rangle \right\}. \tag{4.6}$$

Notice that this problem is essentially linear in y (disregarding the relatively simple cone constraint K^*). The abstract Kuhn-Tucker condition (cf. [19], [21]), i.e. extremality or optimality condition, corresponding to problems (4.2), (4.4) and (4.6) is $(0, 0) \in \partial H(x, y)$. This is equivalent to the property that the pair (x, y) is a saddlepoint of H, i.e. a solution of (L), and this is easily seen to be equivalent, in view of (4.5), to the pair of subdifferential relations

$$y \in \partial f(x), \qquad -x \in \partial \psi_{K^*}(y). \tag{4.7}$$

The symmetrized model now essentially follows from the above, upon choosing f and K as in (2.3) and (2.4) and invoking the information contained in Propositions 1 and 2. In view of the very general formulation of Proposition 2, though, we can actually treat with no additional effort a form of (P) slightly more general. Specifically, we assume henceforth that f is defined as in (2.3) but with the slight change that, instead of taking the σ_j's to be of the form (2.2), we permit them to be of the form

$$\sigma_j(x_j, \xi_j) = \begin{cases} (f_j \xi_j)(x_j) & \text{if } \xi_j > 0, \\ h_j(x_j) & \text{if } \xi_j = 0, \\ +\infty & \text{if } \xi_j < 0, \end{cases} \tag{4.8}$$

where h_j is any given extended-real-valued convex function satisfying

$$\psi_{\{0\}} \geqslant h_j \geqslant f_j 0^+. \tag{4.9}$$

With this choice of f and K, problem (4.2) becomes the primal problem (P) introduced following (2.3), where it is understood that the σ_j's may be of the more general form (4.8), (4.9).

Turning now to the dual, in problem (4.4) we have K^* given by (2.7). By Propositions 1 and 2, formulas (2.8) and (2.9) are valid for ψ_i^* and σ_j^*, respectively. When substituted into formula (2.6) for f^*, these serve to express f^* directly in terms of the individual conjugates f_0^*, f_i^*, f_j^*. The resulting expression for f^* reveals f^* as having structure essentially identical (i.e. "symmetric") with that of f, except for the roles of i and j having been interchanged. Note that in Q^* also the roles of i and j are interchanged. This

shows that problem (4.4) does reduce to problem (D) of §2, and that this problem is indeed a symmetric dual of (P).

The reader can easily imagine the saddlepoint problem (4.6) written also in terms of the additional structure now present.

Consider now the perturbational aspect of the model. Formula (4.1) shows that problem (4.2) is embedded in a whole class of parametrized problems

$$u \to \min_{x} \{ F(x, u) \},$$

in which the perturbation parameter u measures the amount by which the cone K and the graph of f have been shifted horizontally in relation to each other. An examination of (4.3) and (4.4) reveals the (symmetric) fact that the same type of perturbations are involved with the dual problem. It is an important fact about perturbational duality that, when there is no duality gap, the dual optimal solutions serve to describe, in terms of directional derivatives and subgradient theory, the sensitivity or instantaneous rate of change of the primal optimal value with respect to small perturbations in a given direction u, say. (See [19, Theorem 29.1] and [21, Theorems 16 and 17].) Viewed in relation to the specific structure involved in (P), the above class of perturbations for (4.2) are as follows. Regarding the cone P as stationary (in relative terms), the epigraph of f_0 is translated horizontally, the epigraphs of the f_i's are each translated both horizontally and vertically, and the epigraphs of the f_j's are each both translated horizontally and also subjected to a positive scalar magnification with respect to the origin. A dual optimal solution will, according to the general theory, usually describe primal problem sensitivity with respect to any vector combination of the above types of problem perturbations.

A dual optimal solution y which satisfies the condition

$$\text{val}(P) = \text{val}(D) = -f^*(y) \in R, \quad y \in K^*, \qquad (4.10)$$

is called a *Kuhn-Tucker vector for* (P). The importance of such vectors stems from the following type of result.

PROPOSITION 3. *Let y be a Kuhn-Tucker vector for* (P). *Then the solutions of* (P), *if any, occur among the global minimizers of the function*

$$x \to H(x, y) = f(x) - \langle x, y \rangle.$$

In particular, x solves (P) *if and only if*

$$f(x) - \langle x, y \rangle = \inf_{X} \{ f - \langle \cdot, y \rangle \} \quad \text{and} \quad x \in K, \langle x, y \rangle = 0.$$

PROOF. From the way G and H were introduced above, we have that

$$G(y, v) = \inf_{x} \{ \inf_{u} \{ F(x, u) + \langle u, y \rangle \} - \langle x, v \rangle \}$$

$$= \inf_{x} \{ H(x, y) - \langle x, v \rangle \}.$$

For any $y \in K^*$, it therefore follows from (4.3) and (4.5) that

$$- f^*(y) = \inf_x \{ H(x, y) \} = \inf_X \{ f - \langle \cdot, y \rangle \}. \qquad (4.11)$$

Now assume y is a Kuhn-Tucker vector for (P), and write μ for the common optimal value in (4.10). Suppose x solves (P). Then $f(x) = \mu$, $x \in K$. But by (4.10) and (4.11) we also have $\mu = - f^*(y) \leqslant f(x) - \langle x, y \rangle$, $y \in K^*$. Using $\mu \in R$, we can combine these facts to deduce both $0 \leqslant -\langle x, y \rangle$ and $0 \leqslant \langle x, y \rangle$. Hence $0 = \langle x, y \rangle$, and so it follows (using (4.11) again) that

$$f(x) - \langle x, y \rangle = \mu = \inf_X \{ f - \langle \cdot, y \rangle \}.$$

On the other hand, suppose that x satisfies

$$x \in K, \quad \langle x, y \rangle = 0, \quad f(x) - \langle x, y \rangle = \inf_X \{ f - \langle \cdot, y \rangle \}.$$

Using (4.10) and (4.11), we can then write $\mu = - f^*(y) = f(x) - \langle x, y \rangle$, and hence $\mu = f(x)$. Since also $x \in K$, this shows that x solves (P).

The next result includes an explicit characterization of the Kuhn-Tucker conditions.

PROPOSITION 4. *In order that x solve* (P) *and y be a Kuhn-Tucker vector for* (P), *it is necessary that (x, y) solve* (L), *and this occurs if and only if (x, y) satisfies the Kuhn-Tucker conditions*

$$\left(x_0, x^I, x^J \right) \in \mathrm{cl}\, P, \qquad \left(y_0, y^I, y^J \right) \in P^*,$$

$$\left\langle \left(x_0, x^I, x^J \right), \left(y_0, y^I, y^J \right) \right\rangle = 0, \qquad y_0 \in \partial f_0(x_0),$$

$$\left. \begin{array}{l} f_i(x_i) \leqslant 0 \leqslant \eta_i \\ f_i(x_i) \cdot \eta_i = 0 \\ y_i \in \partial (\eta_i f_i)(x_i) \end{array} \right\} \forall i \in I, \qquad \left. \begin{array}{l} f_j^*(y_j) \leqslant 0 \leqslant \xi_j \\ f_j^*(y_j) \cdot \xi_j = 0 \\ y_j \in \partial (f_j \xi_j)(x_j) \end{array} \right\} \forall j \in J,$$

where $\eta_i f_i$ is to be interpreted as $\psi_{\mathrm{dom}\, f_i}$ whenever $\eta_i = 0$, and $f_j \xi_j$ is to be interpreted as h_j whenever $\xi_j = 0$. Conversely, if (x, y) satisfies the Kuhn-Tucker conditions just listed and if the cone P is closed, then x solves (P) *and y is a Kuhn-Tucker vector for* (P).

PROOF. The condition $- x \in \partial \psi_{K^*}(y)$ in (4.7) is equivalent (using the bipolar theorem) to $x \in \mathrm{cl}\, K$, $y \in K^*$, $\langle x, y \rangle = 0$. The special form of K, cf. (2.4), permits this to be written as the two conditions

$$\left(x_0, x^I, x^J \right) \in \mathrm{cl}\, P, \left(y_0, y^I, y^J \right) \in P^*, \left\langle \left(x_0, x^I, x^J \right), \left(y_0, y^I, y^J \right) \right\rangle = 0$$

and

$$\xi^I = 0, \quad \xi^J \in R^J, \quad \eta^I \in R^I, \quad \eta^J = 0.$$

Now consider the other condition of (4.7), $y \in \partial f(x)$. It is an elementary consequence of the additive separability of f that this splits into the

conditions

$$y_0 \in \partial f_0(x_0),$$

$$(y_i, \eta_i) \in \partial \psi_i(x_i, \xi_i), \ \forall i \in I, \quad (y_j, \eta_j) \in \partial \sigma_j(x_j, \xi_j), \ \forall j \in J.$$

By the subdifferential formulas established in Propositions 1 and 2, the last two sets of conditions can be broken down further. Combining all the resulting facts with the cone information gives the characterization of the Kuhn-Tucker conditions asserted. Next, the necessity assertion is immediate from the implication (e) $\Rightarrow$ (f) of [21, Theorem 15]. When P is closed, so is K, and hence the underlying function F defined in (4.1) is closed convex in u for each fixed x. Nearly everything in the converse assertion now follows immediately from the implication (f) $\Rightarrow$ (e) of [21, Theorem 15]. The only remaining item to prove is finiteness of the saddlevalue. But it is easy to check, using (4.5), that if H has a saddlepoint then its saddlevalue must be finite.

The following corollary corresponds to the classical Kuhn-Tucker theorem. The traditional role of some sort of constraint qualification is played here by the relation

$$\inf(P) = \max(D), \tag{4.12}$$

which is sometimes called "strong duality." We understand (4.12) to mean that

$$\operatorname{val}(P) = \operatorname{val}(D) = -f^*(y) \quad \text{for some } y \in K^*. \tag{4.12'}$$

Note that, while the y's fulfilling (4.12') are necessarily solutions to (D), they need not in general be Kuhn-Tucker vectors for (P), cf. (4.10). This is because the common optimal value in (4.12') could perhaps be $-\infty$.

COROLLARY 4A (KUHN-TUCKER THEOREM). *Assume that the strong duality relation* (4.12) *holds. In order that x solve* (P), *it is necessary that there exist a y such that (x, y) satisfies the Kuhn-Tucker conditions listed in Proposition* 4. *When the cone P is closed, this necessary condition is also sufficient.*

PROOF. Let x solve (P) and y satisfy the condition in (4.12'). Since f is never $-\infty$, $\operatorname{val}(P) > -\infty$; since f^* is never $-\infty$, $\operatorname{val}(D) < +\infty$. Hence the common optimal value in (4.12') is finite, so that y is in fact a Kuhn-Tucker vector for (P). Now apply the proposition.

For the necessary condition in Corollary 4A to have substance, we must provide conditions which guarantee that relation (4.12) holds. The section concludes with two such conditions. The first is an all-purpose condition, applicable to the most general spaces.

PROPOSITION 5. *The strong duality relation* (4.12) *holds if there exist* $(x_0, x^I, x^J) \in P$ *and* $\xi^J > 0$ *such that*

(i) f_0 *is bounded above on some neighborhood of* x_0;

(ii) *for each* $i \in I$ *there exists* $\alpha_i > 0$ *such that* $f_i \leq -\alpha_i$ *on some neighborhood of* x_i; *and*

(iii) *for each* $j \in J$, σ_j *is bounded above on some neighborhood of* (x_j, ξ_j).

We remark that when an f_i is continuous, the corresponding condition in (ii) can of course be replaced by the condition $f_i(x_i) < 0$. Also, when the topology on X_j is determined by a norm, it can be shown that the corresponding condition in (iii) can be relaxed to simply requiring that f_j be bounded above on some neighborhood of $\xi_j^{-1}x_j$.

PROOF. By [21, Theorem 17], it suffices to ensure that the primal optimal value function $\varphi(u) = \inf_x \{ F(x, u) \}$ is bounded above on a neighborhood of the origin. Since by (4.1) φ satisfies

$$\varphi(u) = \inf_x \{ f(x) + \psi_K(x + u) \}$$

$$= \inf_x \{ f(x - u) + \psi_K(x) \} \leq \psi_K(x) + f(x - u)$$

for any x, it suffices to have an $x \in K$ such that the function $u \to f(x - u)$, is bounded above on some neighborhood of the origin. The rest of the proof consists of translating this condition into the form asserted, using the particular form of f and K given in (2.3), (2.1), (4.8) and (2.4).

The second result along these lines deals rather fully with the basic case in which all the spaces are finite-dimensional. In §5, as an application of this result, we shall obtain a new, refined criterion for the existence of optimal Lagrange multipliers for ordinary convex programming in R^n.

The relative interior of a set will be denoted by "ri".

PROPOSITION 6. *Assume all the spaces are finite-dimensional and that* val(P) $> -\infty$. *Then a Kuhn-Tucker vector for* (P) *exists (and hence the strong duality relation* (4.12) *holds a fortiori) if there exist an* $\xi^J \geq 0$ *and* $(x_0, x^I, x^J) \in P$ *such that the conditions*

$$\left(x_0, x^I, x^J \right) \in \text{ri } P, \tag{4.13}$$

$$x_0 \in \text{ri dom } f_0, \tag{4.14}$$

$$x_i \in \text{ri dom } f_i \quad \text{and} \quad f_i(x_i) < 0, \tag{4.15i}$$

$$\xi_j > 0 \quad \text{and} \quad x_j \in \xi_j \text{ ri dom } f_j, \tag{4.16j}$$

hold for all $i \in I$ *and* $j \in J$. *Moreover, when any or all of the problem elements* P, f_0, f_i, f_j *are polyhedral, corresponding weakenings in conditions* (4.13) *through* (4.16) *can be made as follows:*

$$\left(x_0, x^I, x^J \right) \in P, \tag{4.13'}$$

$$x_0 \in \text{dom } f_0, \tag{4.14'}$$

$$f_i(x_i) \leqslant 0, \tag{4.15$'_i$}$$

$$\left. \begin{array}{l} h_j = f_j 0^+, \xi_j \geqslant 0, x_j \in \text{dom } f_j \xi_j, \\ \text{where } f_j \xi_j \text{ is } f_j 0^+ \text{ if } \xi_j = 0. \end{array} \right\} \tag{4.16$'_j$}$$

PROOF. Since the hypotheses entail finiteness of $\mu = \text{val}(P)$ and since the optimal value in (D) cannot exceed μ, it suffices to prove that

$$\mu = -f^*(y) \quad \text{for some } y \in K^*. \tag{4.17}$$

We shall first provide an argument leading to (4.17), introducing hypotheses as needed, and afterwards indicate why the condition in the proposition guarantees the hypotheses needed by our argument. We begin by scanning all the functions f_0, ψ_i, σ_j involved in f (given by (2.3), (2.1), (4.8), (4.9)) to see which are polyhedral and which are not. Note that ψ_i is polyhedral if and only if f_i is, and (by Proposition 2) σ_j is polyhedral if and only if f_j is and also $h_j = f_j 0^+$. By a suitable relabeling (permutation) of the coordinates of X, we can partition it as $X = X_1 \times X_2$ in such a way that f can be written in the form

$$f(x) = f(x_1, x_2) = p(x_1) + q(x_2), \tag{4.18}$$

where p consists of all the polyhedral terms of f and q consists of all the nonpolyhedral terms. Next, effectively extend the domains of definition of p and q to all of X by introducing functions $\tilde{p}$ and $\tilde{q}$ via

$$\tilde{p}(x) = \tilde{p}(x_1, x_2) = p(x_1) \quad (\forall x_2) \tag{4.19}$$

and

$$\tilde{q}(x) = \tilde{q}(x_1, x_2) = q(x_2) \quad (\forall x_1). \tag{4.20}$$

For definiteness in what follows, we suppose that P is polyhedral. Then K is, too, and so $\psi_K(x) = \psi_K(x_1, x_2)$ is also (even under the relabeling of variables, since that operation is accomplished via a linear transformation). We can now write

$$\mu = \inf_x \{ f(x) + \psi_K(x) \} = \inf_{(x_1, x_2)} \{ p(x_1) + q(x_2) + \psi_K(x_1, x_2) \}$$

$$= \inf_{(x_1, x_2)} \{ \tilde{p}(x_1, x_2) + \tilde{q}(x_1, x_2) + \psi_K(x_1, x_2) \}$$

$$= -[(\tilde{p} + \psi_K) + \tilde{q}]^*(0, 0). \tag{4.21}$$

Since $\tilde{p} + \psi_K$ is polyhedral, by [19, Theorem 20.1] we have

$$[(\tilde{p} + \psi_K) + \tilde{q}]^*(0, 0) = (\tilde{p} + \psi_K)^* \square \tilde{q}^*(0, 0),$$

where the infimal convolution on the right is actually attained, provided the hypothesis

$$\varnothing \neq \text{dom}(\tilde{p} + \psi_K) \cap \text{ri dom } \tilde{q} \tag{4.22}$$

is satisfied. This means, according to (4.21), that

$$\mu = -\left[(\tilde{p} + \psi_K)^*(-w_1, -w_2) + \tilde{q}^*(w_1, w_2)\right] \tag{4.23}$$

for some $w = (w_1, w_2)$ belonging to the space $Y = Y_1 \times Y_2$ paired with $X = X_1 \times X_2$. Since (4.20) yields

$$\tilde{q}^*(w) = \tilde{q}^*(w_1, w_2) = \begin{cases} q^*(w_2) & \text{if } w_1 = 0, \\ +\infty & \text{otherwise,} \end{cases}$$

and since $\mu > -\infty$, it follows from (4.23) that

$$w_1 = 0 \quad \text{and} \quad \tilde{q}^*(w_1, w_2) = q^*(w_2). \tag{4.24}$$

By [19, p. 179], we also have $(\tilde{p} + \psi_K)^*(0, -w_2) = \tilde{p}^* \square \psi_K^*(0, -w_2)$, where the inf-convolution on the right is attained, provided the hypothesis

$$\varnothing \neq \text{dom } \tilde{p} \cap \text{dom } \psi_K \tag{4.25}$$

is satisfied. This means, according to (4.23) and (4.24), that

$$\mu = -\left[\tilde{p}^*(z_1, z_2) + \psi_K^*(0 - z_1, -w_2 - z_2) + q^*(w_2)\right] \tag{4.26}$$

for some $z = (z_1, z_2)$ in Y. Since (4.19) yields

$$\tilde{p}^*(z) = \tilde{p}^*(z_1, z_2) = \begin{cases} p^*(z_1) & \text{if } z_2 = 0, \\ +\infty & \text{otherwise,} \end{cases}$$

and since $\mu > -\infty$, it follows from (4.26) that

$$z_2 = 0 \quad \text{and} \quad \tilde{p}^*(z_1, z_2) = p^*(z_1).$$

From this, (4.26) can be rewritten as

$$\mu = -\left[p^*(z_1) + q^*(w_2) + \psi_K^*(-z_1, -w_2)\right]. \tag{4.27}$$

Now notice from the additive separability in (4.18) that

$$f^*(y) = f^*(y_1, y_2) = p^*(y_1) + q^*(y_2).$$

Using this together with the fact that $\psi_K^* = \psi_{K^0} = \psi_{K^*}(-\cdot)$, we can rewrite (4.27) as

$$\mu = -\left[f^*(y) + \psi_{K^*}(y)\right], \quad y = (z_1, w_2).$$

Using $\mu > -\infty$ once more, we obtain $y \in K^*$ for this y, and so (4.17) is established. This argument used (4.22) and (4.25). Since $\text{dom}(\tilde{p} + \psi_K) = \text{dom } \tilde{p} \cap \text{dom } \psi_K$, it is clear that

$$\varnothing \neq \text{dom } \tilde{p} \cap \text{dom } \psi_K \cap \text{ri dom } \tilde{q} \tag{4.28}$$

is the hypothesis required for our argument. In the event that K is not polyhedral, we simply group the term ψ_K with $\tilde{q}$ rather than $\tilde{p}$ and then mimic the above argument. The hypotheses required for this are

$$\varnothing \neq \text{dom } \tilde{p} \cap \text{ri dom}(\psi_K + \tilde{q})$$

and

$$\varnothing \neq \text{ri dom } \psi_K \cap \text{ri dom } \tilde{q}.$$

By [**19**, Theorem 6.5], these are equivalent to the single hypothesis

$$\varnothing \neq \text{dom } \tilde{p} \cap \text{ri dom } \psi_K \cap \text{ri dom } \tilde{q}. \tag{4.29}$$

To finish the proof, it remains only to check that the condition assumed in the proposition amounts to either (4.28) or (4.29) according to whether P is or is not polyhedral. Since $K = P \times Q$, where Q is relatively open, $\text{dom } \psi_K = P \times Q$ and $\text{ri dom } \psi_K = \text{ri } P \times Q$. Now consider the remaining sets involved in (4.28) and (4.29). By formulas (4.18)–(4.20), together with the additive separability of p and q in terms of the functions f_0, ψ_i, σ_j, the issue reduces to determining formulas describing the sets $\text{dom } \psi_i$, $\text{dom } \sigma_j$ and also their relative interiors. But (2.1) yields $\text{dom } \psi_i = \{(x_i, \xi_i) | f_i(x_i) + \xi_i \leqslant 0\}$, and so by [**19**, Lemma 7.3 or Theorem 6.8] it follows that

$$\text{ri dom } \psi_i = \{(x_i, \xi_i) | x_i \in \text{ri dom } f_i, f_i(x_i) + \xi_i < 0\}.$$

Also, according to (4.8) and (4.9) we have

$$\text{dom } \sigma_j = \{(x_j, \xi_j) | \xi_j > 0 \text{ and } x_j \in \text{dom } f_j\xi_j, \text{ or } \xi_j = 0 \text{ and } x_j \in \text{dom } h_j\},$$

so by [**19**, Theorem 6.8] it follows that

$$\text{ri dom } \sigma_j = \{(x_j, \xi_j) | \xi_j > 0, x_j \in \xi_j \text{ dom } f_j\}.$$

From this information it is easy to see that the condition in the proposition is equivalent to either (4.28) or (4.29), depending on whether P is or is not polyhedral. This concludes the proof.

Alternative conditions which ensure (4.12) can be developed by combining [**21**, Theorem 17] with the various criteria given in [**21**, Theorem 18], much as in the proof of Proposition 5. This we leave to the interested reader.

Finally, we remark that there are other general results which hold for the present model by virtue of its fitting the general perturbational duality framework. We refer the reader to [**21**, especially §7] from which it is possible to deduce these further results concerning the symmetrized model.

5. Specialization to ordinary programming: comparison with the traditional Lagrangian model. We now analyze the symmetrized model of §4 in the most important case,

$$J = \varnothing, \quad P = \{(x_0, x') | x_0 = x_i, \forall i \in I\},$$

$$P^* = \left\{(y_0, y') | y_0 + \sum_I y_i = 0\right\}, \tag{5.1}$$

to see how it compares with the usual duality model for the ordinary convex

programming problem

$$\min\{ f_0(x_0) | f_i(x_0) \leqslant 0, \forall i \in I \}. \tag{P_0}$$

We shall see that the resulting symmetrized model is technically distinct from ordinary duality, yet closely related and in fact, under generally mild conditions, essentially equivalent. The symmetrized model thus provides an alternate means of obtaining facts concerning the ordinary convex program. In particular, we establish in this way a new, sharpened existence theorem for optimal Lagrange multipliers in the finite-dimensional case. We also derive (directly) under minimal assumptions the classical Kuhn-Tucker optimality conditions for the case of nondifferentiable, not everywhere finite functions in general spaces. As an illustration of the delicate interrelationships between the ordinary and the symmetrized models, we conclude with a discussion of the case of homogeneous constraint functions. This extends recent work of C. R. Glassey concerning "explicit duality" for such problems.

By the "ordinary duality model" we mean the trio of problems consisting of (P_0) together with its so-called Lagrangian dual problem

$$\max_{\eta' > 0} \left\{ \inf_{x_0 \in C} \left\{ f_0(x_0) + \sum_I \eta_i f_i(x_0) \right\} \right\} \tag{D_0}$$

and Lagrangian saddlepoint problem

$$\underset{x_0 \in C}{\text{minimax}}_{\eta' > 0} \left\{ f_0(x_0) + \sum_I \eta_i f_i(x_0) \right\}. \tag{L_0}$$

We use the notation

$$C = C_0 \cap \bigcap_I C_i, \quad C_0 = \operatorname{dom} f_0, \quad C_i = \operatorname{dom} f_i \;\; \forall i \in I \tag{5.2}$$

and make the trivial *nondegeneracy assumption* that $C \neq \varnothing$. In the case of functions not everywhere finite, it is essential to observe the restriction $x_0 \in C$ in both (D_0) and (L_0).

This trio arises from the general perturbational duality model by choosing

$$F_0(x_0, \mu') = \begin{cases} f_0(x_0) & \text{if } f_i(x_0) \leqslant \mu_i \; \forall i \in I, \\ +\infty & \text{otherwise.} \end{cases} \tag{5.3}$$

Indeed, $\min_{x_0}\{ F_0(x_0, 0)\}$ is exactly (P_0), and when we generate G_0 and H_0 from F_0 by means of the formulas

$$H_0(x_0, \eta') = \inf_{\mu'} \left\{ F_0(x_0, \mu') + \langle \mu', \eta' \rangle \right\}$$

and

$$G_0(\eta', v_0) = \inf_{x_0} \left\{ H_0(x_0, \eta') - \langle x_0, v_0 \rangle \right\}$$

prescribed in [21, see Equations (4.2) and (4.15)], we easily obtain by (5.3)

that

$$H_0\left(x_0, \eta^I\right) = \begin{cases} f_0\left(x_0\right) + \sum_I \eta_i f_i\left(x_0\right) & \text{if } x_0 \in C \text{ and } \eta^I \geqslant 0, \\ -\infty & \text{if } x_0 \in C \text{ and } \eta^I \not\geqslant 0, \\ +\infty & \text{if } x_0 \notin C, \end{cases} \quad (5.4)$$

and

$$G_0\left(\eta^I, v_0\right) = \begin{cases} \inf_{x_0 \in C} \left\{ f_0\left(x_0\right) + \sum_I \eta_i f_i\left(x_0\right) - \langle x_0, v_0\rangle \right\} & \text{if } \eta^I \geqslant 0, \\ -\infty & \text{otherwise.} \end{cases} \quad (5.5)$$

From these, it is clear that the problem $\max_{\eta^I}\{G_0(\eta^I, 0)\}$ coincides with (D$_0$) and the problem

$$\underset{x_0 \quad \eta^I}{\text{minimax}} \left\{ H_0\left(x_0, \eta^I\right) \right\}$$

coincides with (L$_0$).

Now consider the form taken by the symmetrized model of §4 under the particular choice (5.1). The primal problem (P) becomes

$$\min\left\{ f_0\left(x_0\right) \mid \left(x_0, x^I, \xi^I\right) \text{ such that } f_i\left(x_0\right) \leqslant 0 \ \forall i \in I, \right.$$
$$\left. \text{and } \xi^I = 0, x_0 = x_i \ \forall i \in I \right\}. \quad \text{(P}_s\text{)}$$

The extra variables (x^I, ξ^I) here are of course completely determined and could be suppressed without harm. Nevertheless, for the purpose of maintaining a clear distinction between the ordinary and the symmetrized model, it is useful to keep in mind, at least in a formal way, the presence of these additional variables. With the dual and Lagrangian problems (D$_s$) and (L$_s$), analogous distinctions actually can make a difference in terms of solvability.

The dual problem (D) when specialized according to (5.1) becomes

$$-\min\left\{ f_0^*\left(y_0\right) + \sum_I \left(f_i^*\eta_i\right)\left(y_i\right) \mid \eta^I \geqslant 0, y_0 + \sum_I y_i = 0 \right\},$$
$$\text{where } f_i^*\eta_i \text{ is } f_i^*0^+ \text{ whenever } \eta_i = 0. \quad \text{(D}_s^+\text{)}$$

Proposition 7(a) below shows that, under a usually harmless assumption, we can substitute for (D$_s^+$) a simpler dual problem not requiring the recession functions $f_i^*0^+$:

$$-\min\left\{ f_0^*\left(y_0\right) + \sum_I \left(f_i^*\eta_i\right)\left(y_i\right) \mid \eta^I \geqslant 0, y_0 + \sum_I y_i = 0 \right\},$$
$$\text{where } f_i^*\eta_i \text{ is } \psi_{\{0\}} \text{ whenever } \eta_i = 0. \quad \text{(D}_s\text{)}$$

In the course of the proof it is shown that the relations

$$\text{val}(D_s) \leqslant \text{val}(D_s^+) \leqslant \text{val}(D_0)$$

hold in general. Proposition 7(b^+) provides a relatively mild condition under which solving the symmetrized dual (D_s^+) is equivalent to solving the ordinary dual (D_0). Proposition 7(b) is a variant which provides the condition under which solving the simplified symmetrized dual (D_s) is equivalent to solving (D_0).

The reader will notice here and below that, in labeling results and equations pertaining to the two symmetrized dual problems, we try systematically to use the presence or absence of the superscript $^+$ to signal reference to problem (D_s^+) or (D_s), respectively. In other words, the $^+$ is employed in this way as a mnemonic device intended to signal the use of $f_i^*0^+$ or $\psi_{\{0\}}$ for the indices i corresponding to $\eta_i = 0$.

PROPOSITION 7. (a) *Assume that the effective domains of the functions satisfy*

$$C_0 \subset C_i, \quad \forall i \in I \tag{5.6}$$

(see (5.2)). Then the optimal values in (D_s^+) and (D_s) coincide. Moreover, any solution of (D_s) is a solution of (D_s^+), and conversely, if (y_0, y^I, η^I) solves (D_s^+) then ($\tilde{y}_0, \tilde{y}^I, \eta^I$) solves ($D_s$), where

$$\tilde{y}_0 = y_0 + \sum_{I_0} y_i \quad and \quad \tilde{y}_i = \begin{cases} 0, & i \in I_0, \\ y_i, & i \in I_+, \end{cases} \tag{5.7}$$

for $I_0 = \{i \in I | \eta_i = 0\}$, $I_+ = I \setminus I_0$.

(b^+) *Assume that for each $\eta^I \geqslant 0$ the functions satisfy the condition*

$$-\left(f_0 + \sum_I \eta_i f_i\right)^*(v_0) = \min\left\{ f_0^*(y_0) + \sum_I (f_i^*\eta_i)(y_i) | y_0 + \sum_I y_i = v_0 \right\},$$

where $\eta_i f_i$ is ψ_{C_i} and $f_i^*\eta_i$ is $f_i^*0^+$ whenever $\eta_i = 0$. (5.8$^+$)

Then the optimal values of (D_s^+) and (D_0) coincide. Moreover, if (y_0, y^I, η^I) solves (D_s^+) then η^I solves (D_0), and conversely, if η^I solves (D_0) then there exist y_0 and y^I such that (y_0, y^I, η^I) solves (D_s^+).

(b) *Assume that for each $\eta^I \geqslant 0$ the functions satisfy the condition*

$$-\left(f_0 + \sum_{I_+} \eta_i f_i\right)^*(v_0) = \min\left\{ f_0^*(y_0) + \sum_{I_+} (f_i^*\eta_i)(y_i) | y_0 + \sum_{I_+} y_i = v_0 \right\},$$

where $I_+ = \{i \in I | \eta_i > 0\}$, (5.8)

and assume that (5.6) holds as well. Then the optimal values of (D_s) and (D_0) coincide. Moreover, if (y_0, y^I, η^I) solves (D_s) then η^I solves (D_0), and conversely, if η^I solves (D_0) then there exist y_0 and y^I such that (y_0, y^I, η^I) solves (D_s).

PROOF. Observe first that any collection $\varphi_1, \ldots, \varphi_m$ of proper convex functions having lower semicontinuous hulls somewhere finite satisfies

$$\inf_{X_0} \left\{ \sum \varphi_i \right\} \geqslant \inf_{X_0} \left\{ \sum \mathrm{cl}\, \varphi_i \right\} = -\left(\sum \mathrm{cl}\, \varphi_i \right)^*(0)$$

$$= -\mathrm{cl}(\varphi_1^* \square \cdots \square \varphi_m^*)(0) \geqslant -(\varphi_1^* \square \cdots \square \varphi_m^*)(0). \quad (5.9)$$

The second equality here follows by the same argument as given in [**19**, p. 145]. Using (5.9) and then $\psi_{\{0\}} \geqslant f_i^* 0^+$, we have

$$\inf_C \left\{ f_0 + \sum_I \eta_i f_i \right\} = \inf_{X_0} \left\{ f_0 + \sum_{I_0} \psi_{C_i} + \sum_{I_+} \eta_i f_i \right\}$$

$$\geqslant -\inf \left\{ f_0^*(y_0) + \sum_{I_0} (f_i^* 0^+)(y_i) + \sum_{I_+} (f_i^* \eta_i)(y_i) \,\big|\, y_0 + \sum_I y_i = 0 \right\}$$

$$\geqslant -\inf \left\{ f_0^*(y_0) + \sum_{I_+} (f_i^* \eta_i)(y_i) \,\big|\, y_0 + \sum_{I_+} y_i = 0 \right\} \quad (5.10)$$

for each $\eta^I \geqslant 0$. This establishes the general relations

$$\mathrm{val}(D_s) \leqslant \mathrm{val}(D_s^+) \leqslant \mathrm{val}(D_0). \quad (5.11)$$

Now assume condition (5.6). Then $\psi_{C_i} \leqslant \psi_{C_0}$ and hence $f_i^* 0^+ \geqslant f_0^* 0^+$ by Lemma 1 in §3. If (y_0, y^I, η^I) satisfies $\eta^I \geqslant 0$, $y_0 + \sum_I y_i = 0$ and $y_0 \in \mathrm{dom}\, f_0^*$, then we can write

$$f_0^*(y_0) + \sum_{I_0} (f_i^* 0^+)(y_i) \geqslant f_0^*(y_0) + \sum_{I_0} (f_0^* 0^+)(y_i)$$

$$\geqslant f_0^*(y_0) + f_0^* 0^+ \left(\sum_{I_0} y_i \right) \geqslant f_0^*(\tilde{y}_0) = f_0^*(\tilde{y}_0) + \sum_{I_0} \psi_{\{0\}}(\tilde{y}_i). \quad (5.12)$$

The second inequality is by the subadditivity of $f_0^* 0^+$ (Lemma 1 shows it is a support function), while the third inequality follows from $y_0 \in \mathrm{dom}\, f_0^*$ and the definition of recession function (see ff. (2.2)). Since $\psi_{\{0\}} \geqslant f_i^* 0^+$, equality must hold throughout (5.12). Hence

$$f_0^*(y_0) + \sum_{I_0} (f_i^* 0^+)(y_i) + \sum_{I_+} (f_i^* \eta_i)(y_i)$$

$$= f_0^*(\tilde{y}_0) + \sum_{I_+} \psi_{\{0\}}(\tilde{y}_i) + \sum_{I_+} (f_i^* \eta_i)(\tilde{y}_i),$$

where $\tilde{y}_0 + \sum_I \tilde{y}_i = (y_0 + \sum_{I_0} y_i) + \sum_{I_+} y_i = 0$. This shows that the optimal values in (D_s^+) and (D_s) coincide and that if (y_0, y^I, η^I) solves (D_s^+) then the (y_0, y^I, η^I) induced as in (5.7) solves (D_s). On the other hand, suppose that (y_0, y^I, η^I) solves (D_s). If the (common) optimal value is $-\infty$, then trivially (y_0, y^I, η^I) also solves (D_s^+). So suppose $\mathrm{val}(D_s) > -\infty$. Then necessarily $y_i = 0$ for each $i \in I_0$, so that $\psi_{\{0\}}(y_i) = 0 = (f_i^* 0^+)(y_i)$ for such i's. Hence

(y_0, y^I, η^I) yields the same value in the objective function of (D_s^+) as it does in (D_s). Since the two optimal values agree, this vector must also be optimal for (D_s^+). This completes the proof of part (a). Now assume condition (5.8^+). This yields equality, with attainment by some y's, in the first inequality of calculation (5.10). The conclusions of part (b^+) are immediate from this. Finally, assume conditions (5.8) and (5.6). These yield

$$\inf_C \left\{ f_0 + \sum_I \eta_i f_i \right\} = \inf_{X_0} \left\{ f_0 + \sum_{I_0} \psi_{C_i} + \sum_{I_+} \eta_i f_i \right\} = \inf_{X_0} \left\{ f_0 + \sum_{I_+} \eta_i f_i \right\}$$

$$= -\min \left\{ f_0^*(y_0) + \sum_{I_+} (f_i^* \eta_i)(y_i) | y_0 + \sum_{I_+} y_i = 0 \right\}$$

for each $\eta^I \geqslant 0$. The conclusions of part (b) are immediate from this.

Condition (5.6) can always be met by simply redefining f_0 (if necessary) to be $+\infty$ outside C. The new f_0 will then satisfy $C_0 = C \subset C_i$, $\forall i \in I$. The only adverse effect likely from this redefinition is the possible complication of the formula for f_0^*. This can occur, as for example in the case of programming subject to homogeneous constraints discussed at the end of this section.

Conditions (5.8) and (5.8^+) are somewhat more restrictive hypotheses but still relatively harmless. Indeed, generally speaking, they are easier to satisfy (i.e. weaker) than a constraint qualification. For example, according to [21, Theorem 20(a)], (5.8) is satisfied in general spaces if there exists an x_0 such that each of the functions f_0, f_i $(i \in I)$ is finite at x_0 and all except possibly one of them are bounded above on some neighborhood of x_0. Alternatively, if $X_0 = R^n$, (5.8) is satisfied whenever

$$\varnothing \neq \text{ri } C_0 \cap \bigcap_I \text{ri } C_i,$$

where here the relative interior operation can be deleted for any and all of those functions f_0, f_i $(i \in I)$ which happen to be polyhedral [19, Theorems 16.4, 20.1]. Other conditions sufficient to guarantee validity of (5.8) and (5.8^+) in both finite- and infinite-dimensional contexts can be formulated using [21, Theorem 20], [6, §§4e and 9c] and [22, Theorem 5.6.2].

For a somewhat tangential yet related discussion of the consequences of (5.8), see [5].

We remark that in the presence of (5.6) it can be proved that condition (5.8) is equivalent to condition (5.8^+).

From Proposition 7(a) it follows that, when (5.6) can be arranged (without adverse effects from complicating f_0^*), we might as well deal with the simplified symmetrized dual (D_s) rather than the technically correct version (D_s^+). It should be noted, though, that in so substituting (D_s) for (D_s^+), we will in general be discarding certain unbounded portions of the solution set of

(D_s^+). But this is immaterial in terms of the actual solvability of the dual problem, so long as (5.6) holds (cf. (5.7)).

In connection with Proposition 7(a), we remark that even without condition (5.6) an asymptotic relationship holds between (D_s^+) and (D_s). Namely, by using [17, Theorem 3b(e)] one can show that to each feasible solution (y_0, y^I, η^I) of (D_s^+) there corresponds a net (y_0, y_α^I, η_α^I) of vectors, indexed by α, satisfying

$$\eta_\alpha^I > 0, \quad \eta^I = \lim_\alpha \eta_\alpha^I, \quad 0 = \lim_\alpha \left(y_0 + \sum_I y_{i,\alpha} \right)$$

and

$$-\left[f_0^*(y_0) + \sum_I (f_i^*\eta_i)(y_i) \right] \leqslant -\lim_\alpha \left[f_0^*(y_0) + \sum_I (f_i^*\eta_{i,\alpha})(y_{i,\alpha}) \right].$$

If (D_s) were known to be normal (i.e. if the perturbation function corresponding to (D_s) satisfied (usc γ)(0) $\leqslant \gamma$(0)), it would follow that val(D_s) = val(D_s^+) and hence that each solution of (D_s) is a solution of (D_s^+).

We turn our attention now towards a comparison of Kuhn-Tucker theories associated with the symmetrized and the ordinary models. For the ordinary model this involves the traditional Lagrangian function H_0 given in (5.4), while for the symmetrized model it involves specializing the function H given by (4.5) according to (5.1). This yields H_s, defined as follows:

$$H_s\left(x_0, x^I, \xi^I, y_0, y^I, \eta^I\right) = f_0(x_0) - \left\langle (x_0, x^I, \xi^I), (y_0, y^I, \eta^I) \right\rangle \quad (5.13)$$

if the "x-conditions" $x_0 \in \text{dom} f_0$, $f_i(x_i) + \xi_i \leqslant 0$ $\forall i \in I$ and the "y-conditions" $\eta^I \in R^I$, $y_0 + \Sigma_I y_i = 0$ are both satisfied; $H_s(x,y) = -\infty$ if these x-conditions are met but the y-conditions fail; and $H_s(x,y) = +\infty$ if the x-conditions fail. The Lagrangian function H_s plays the same role for (P_s) and (D_s^+) (and also (D_s) when (5.6) holds) as H_0 plays for (P_0) and (D_0). In particular, the pairs (x_0, x^I, ξ^I), (y_0, y^I, η^I) characterized in Proposition 9 below as the solutions to the "symmetrized" Kuhn-Tucker conditions are precisely the saddlepoints of the minimax problem (L_s) determined by H_s.

The *Kuhn-Tucker vectors for* (P_s) are, according to (5.1) and the general definition in §4, those (y_0, y^I, η^I) which satisfy

$$\eta^I \geqslant 0, \quad y_0 + \sum_I y_i = 0 \quad\quad (5.14a^+)$$

and

$$\text{val}(P_s) = -\left[f_0^*(y_0) + \sum_{I_0} (f_i^*0^+)(y_i) + \sum_{I_+} (f_i^*\eta_i)(y_i) \right] \in R, \quad (5.14b^+)$$

where as usual we write $I_0 = \{ i \in I \mid \eta_i = 0 \}$ and $I_+ = I \setminus I_0$. In particular, they are solutions of (D_s^+). Also of interest are the vectors which satisfy the

stronger conditions

$$\eta^I \geqslant 0, \quad y_0 + \sum_I y_i = 0, \quad y_i = 0 \ \forall i \in I_0 \tag{5.14a}$$

and

$$\text{val}(P_s) = -\left[f_0^*(y_0) + \sum_{I_+} (f_i^*\eta_i)(y_i) \right] \in R. \tag{5.14b}$$

Vectors (y_0, y^I, η^I) satisfying the latter pair of conditions will be called *strong Kuhn-Tucker vectors for* (P_s). They are contained in the solution set of the simplified dual (D_s). The *Kuhn-Tucker vectors for* (P_0) are the η^I which satisfy

$$\eta^I \geqslant 0, \quad \text{val}(P_0) = \inf_C \left\{ f_0 + \sum_I \eta_i f_i \right\} \in R. \tag{5.15}$$

These are also called optimal Lagrange multipliers for (P_0). The next result describes the interrelationships among these objects.

PROPOSITION 8. *Let* $\eta^I \geqslant 0$ *be given, and write* $I_+ = \{i \in I | \eta_i > 0\}$ *and* $I_0 = I \setminus I_+$. *The implications* (b) $\Rightarrow$ (b$^+$) $\Rightarrow$ (a) *hold among the conditions below. Furthermore, if* (5.8$^+$) *holds then* (b$^+$) *is equivalent to* (a), *and if* (5.6) *holds then* (b) *is equivalent to* (b$^+$). *If both* (5.8) *and* (5.6) *hold, the three conditions are mutually equivalent.*

(a) η^I *is a Kuhn-Tucker vector for* (P_0), *i.e.* (5.15) *holds*;

(b$^+$) *there exist* y_0 *and* y^I *such that* (y_0, y^I, η^I) *is a Kuhn-Tucker vector for* (P_s), *i.e.* (5.14$^+$) *holds.*

(b) *there exist* y_0 *and* y^I *such that* (y_0, y^I, η^I) *is a strong Kuhn-Tucker vector for* (P_s), *i.e.* (5.14) *holds.*

PROOF. Note first that to go along with the general relation (5.11) established above we also have

$$\text{val}(D_0) \leqslant \text{val}(P_0) = \text{val}(P_s). \tag{5.16}$$

The implication (b) $\Rightarrow$ (b$^+$) is trivial from the fact that $f_i^*0^+(0) = 0$. Suppose (y_0, y^I, η^I) satisfies (5.14$^+$). By the estimate (5.10) together with the general inequalities (5.11) and (5.16), it follows that η^I satisfies (5.15). Hence (b$^+$) $\Rightarrow$ (a). Next, suppose η^I satisfies (5.15) and that condition (5.8$^+$) holds. Then by Proposition 7(b$^+$), there exist y_0 and y^I such that (5.14$^+$) holds. Hence, (a) $\Rightarrow$ (b$^+$) in the presence of (5.8$^+$). Now suppose (y_0, y^I, η^I) satisfies (5.14$^+$) and that condition (5.6) holds. Define $\tilde{y}_0$ and $\tilde{y}^I$ as in (5.7). By Proposition 7(a) it follows that $(\tilde{y}_0, \tilde{y}^I, \eta^I)$ satisfies (5.14). Hence, (b$^+$) $\Rightarrow$ (b) in the presence of (5.6). The final assertion follows by combining what has already been proved, using the fact (remarked earlier) that (5.8$^+$) is equivalent to (5.8) in the presence of (5.6). Alternatively, apply Proposition 7(b). This completes the proof.

Propositions 3 and 4 and Corollary 4A from §4 can all be specialized according to (5.1), of course, to yield comparable assertions concerning the trio (P_s), (D_s^+), (L_s). We will not write all of this down, though. Instead, we focus on just that part of Proposition 4 characterizing the Kuhn-Tucker conditions, to see what they look like for (P_s).

PROPOSITION 9. *A pair of vectors* (x_0, x^I, ξ^I) *and* (y_0, y^I, η^I) *satisfies the Kuhn-Tucker conditions for* (P_s) *if and only if it satisfies the conditions*

$$x_0 = x_i \ \forall i \in I, \quad \xi^I = 0, \quad y_0 + \sum_I y_i = 0, \tag{5.17a}$$

$$y_0 \in \partial f_0(x_0), \tag{5.17b}$$

$$f_i(x_0) \leqslant 0 \leqslant \eta_i \quad and \quad f_i(x_0) \cdot \eta_i = 0 \ \forall i \in I, \tag{5.17c}$$

$$y_i \in \partial \psi_{\mathrm{dom}\, f_i}(x_0) \ \forall i \in I_0 \quad and \quad y_i \in \eta_i \partial f_i(x_0) \ \forall i \in I_+, \tag{5.17d$^+$}$$

where $I_0 = \{i \in I \,|\, \eta_i = 0\}$ *and* $I_+ = I \setminus I_0$. *When* (5.6) *holds, it is possible to satisfy the preceding conditions if and only if it is possible to satisfy the simplified conditions obtained by replacing* (5.17d$^+$) *with*

$$y_i = 0 \ \forall i \in I_0 \quad and \quad y_i \in \eta_i \partial f_i(x_0) \ \forall i \in I_+. \tag{5.17d}$$

PROOF. The first assertion is immediate from Proposition 4, in view of (5.1). Now observe that $\partial(\eta_i f_i)(x_0) = \eta_i \partial f_i(x_0)$ whenever $\eta_i > 0$ and $\partial f_i(x_0) \neq \emptyset$, and that $0 \in \partial \psi_{C_i}(x_0)$ whenever $x_0 \in C_i$. From these facts it is clear that any pair (x_0, x^I, ξ^I), (y_0, y^I, η^I) which satisfies (5.17) must also satisfy (5.17$^+$). Now suppose we have such a pair satisfying (5.17$^+$), and assume that condition (5.6) holds. Define $(\tilde{y}_0, \tilde{y}^I, \eta^I)$ as in (5.7). It is clear that the given (x_0, x^I, ξ^I) together with $(\tilde{y}_0, \tilde{y}^I, \eta^I)$ satisfy everything in (5.17) except possibly requirement (5.17d). But we have $x_0 \in C_0 \subset C_i$, from which $\partial \psi_{C_i}(x_0) \subset \partial \psi_{C_0}(x_0)$ follows easily. We also have the easy fact that $\partial f_0(x_0) + \partial \psi_{C_0}(x_0) \subset \partial f_0(x_0)$. Combining this information yields

$$\tilde{y}_0 = y_0 + \sum_{I_0} y_i \in \partial f_0(x_0) + \sum_{I_0} \partial \psi_{C_i}(x_0)$$

$$\subset \partial f_0(x_0) + \sum_{I_0} \partial \psi_{C_0}(x_0) \subset \partial f_0(x_0),$$

and so the proof is complete.

The extremality conditions just derived for the symmetrized model, at least in the simplified form (5.17), are extremely closely related to the well known, classical Kuhn-Tucker conditions. In order to make a precise comparison between the two in the present context, we now derive the classical Kuhn-Tucker conditions. These are obtained as part (b) of the following result, which extends [19, Theorem 28.3 (see 1972 edition)] to the general case.

PROPOSITION 10. (a) *In order that x_0 solve* (P_0) *and* η^I *be a Kuhn-Tucker vector for* (P_0), *it is necessary and sufficient that* (x_0, η^I) *solve* (L_0).

(b) *In order that* (x_0, η^I) *solve* (L_0) *it is sufficient that* (x_0, η^I) *satisfies the conditions*

$$f_i(x_0) \leqslant 0 \leqslant \eta_i \quad and \quad f_i(x_0) \cdot \eta_i = 0 \; \forall i \in I \tag{5.18a}$$

and

$$0 \in \partial f_0(x_0) + \sum_{I_+} \eta_i \partial f_i(x_0), \tag{5.18b}$$

where $I_+ = \{i \in I | \eta_i > 0\}$. *When* (5.6) *and* (5.8) *hold, these conditions are also necessary.*

PROOF. Most of part (a) follows immediately from the equivalence between (e) and (f) in [**21**, Theorem 15], since the function F_0 defined in (5.3), which underlies the ordinary duality model, is clearly closed convex in μ^I for each fixed x_0. The finiteness of the common optimal vlaue for the necessity half of (a) is built in to the definition of Kuhn-Tucker vector. Finiteness in the sufficiency half follows from the fact that if (x_0, η^I) is a saddlepoint of H_0 then $x_0 \in C$ and $\eta^I \geqslant 0$, in which case the saddlevalue is finite. This is easy to deduce directly from (5.4). (It requires our nondegeneracy assumption $C \neq \varnothing$.) Now let us establish (b). As just noted, if (x_0, η^I) is a saddlepoint of H_0 then necessarily $x_0 \in C$ and $\eta^I \geqslant 0$. From this it follows, using (5.4), that the saddlepoint condition is equivalent to the conditions ($x_0 \in C$, $\eta^I > 0$ and)

$$\sum_I (\tilde{\eta}_i - \eta_i) \cdot f_i(x_0) \leqslant 0, \quad \forall \tilde{\eta}^I \geqslant 0,$$

and

$$f_0(x_0) + \sum_I \eta_i f_i(x_0) \leqslant f_0(\tilde{x}_0) + \sum_I \eta_i f_i(\tilde{x}_0), \quad \forall \tilde{x}_0 \in C.$$

Now it is easily seen (in the presence of $x_0 \in C$, $\eta^I \geqslant 0$) that the first of these is equivalent to (5.18a) and that the second is equivalent to

$$0 \in \partial \left(f_0 + \sum_I \eta_i f_i \right)(x_0),$$

where $\eta_i f_i$ is interpreted as ψ_{C_i} whenever $\eta_i = 0$, or in other words, to the condition

$$0 \in \partial \left(f_0 + \sum_{I_0} \psi_{C_i} + \sum_{I_+} \eta_i f_i \right)(x_0), \tag{5.18b*}$$

where $I_0 = \{i \in I | \eta_i = 0\}$ and $I_+ = I \setminus I_0$. The remaining analysis concerns breaking (5.18b*) down further. Suppose (x_0, η^I) satisfies (5.18a) together

with (5.18b). Then

$$0 \in \partial f_0(x_0) + \sum_{I_0} \{0\} + \sum_{I_+} \eta_i \partial f_i(x_0)$$

$$\subset \partial f_0(x_0) + \sum_{I_0} \partial \psi_{C_i}(x_0) + \sum_{I_+} \eta_i \partial f_i(x_0)$$

$$\subset \partial \left(f_0 + \sum_{I_0} \partial \psi_{C_i} + \sum_{I_+} \eta_i f_i \right)(x_0)$$

shows that (5.18b*) holds. The first inclusion here follows trivially from the fact $0 \in \partial \psi_{C_i}(x_0)$, while the second is an elementary fact concerning the subdifferential of a sum (see, e.g. [19, proof of Theorem 23.8]). This establishes the sufficiency part of assertion (b). Now assume conversely that (x_0, η^I) satisfies (5.18a) together with (5.18b*) and that conditions (5.6) and (5.8) hold. By (5.6) we have that

$$f_0 + \sum_{I_0} \psi_{C_i} + \sum_{I_+} \eta_i f_i = f_0 + \sum_{I_+} \eta_i f_i,$$

and hence (5.18b*) simplifies to $0 \in \partial (f_0 + \sum_{I_+} \eta_i f_i)(x_0)$. From (5.8) it follows by an elementary argument that

$$\partial \left(f_0 + \sum_{I_+} \eta_i f_i \right)(x_0) = \partial f_0(x_0) + \sum_{I_+} \eta_i \partial f_i(x_0).$$

(Such an argument can be found in [5, §2].) Combining the last two facts, we obtain (5.18b), thus completing the proof.

It is clear that the ordinary Kuhn-Tucker conditions (5.18) are satisfiable if and only if the simplified conditions (5.17) corresponding to the symmetrized model are satisfiable. Similarly, conditions (5.17$^+$) are satisfiable if and only if conditions (5.18$^+$) are, where by (5.18$^+$) we mean (5.18a) together with

$$0 \in \partial f_0(x_0) + \sum_{I_0} \partial \psi_{C_i}(x_0) + \sum_{I_+} \eta_i \partial f_i(x_0). \qquad (5.18b^+)$$

COROLLARY 10A (KUHN-TUCKER THEOREM). *If conditions (5.6), (5.8) and the strong duality relation* $\inf(P_0) = \max(D_0)$ *hold, then in order that* x_0 *solve* (P_0) *it is necessary that there exists an* η^I *which together with* x_0 *satisfies the Kuhn-Tucker conditions (5.18). Conversely, if* (x_0, η^I) *satisfies (5.18), then* x_0 *solves* (P_0) *and* η^I *is a Kuhn-Tucker vector for* (P_0).

PROOF. The converse assertion is immediate from the proposition. Suppose now that $\inf(P_0) = \max(D_0) = \mu$. From (5.5) it is clear that no η^I can yield value $+\infty$ in (D_0), so since (D_0) has a solution we must have $\mu < +\infty$. (This uses our nondegeneracy assumption $C \neq \varnothing$; see (5.5).) Suppose (P_0) has a solution. Then $\mu > -\infty$, because (5.3) shows that no x_0 can yield value $-\infty$ in (P_0). Hence μ must be finite, and the hypothesized solution to (D_0) is

actually a Kuhn-Tucker vector for (P_0). Now apply the proposition.

Notice that Proposition 10 and Corollary 10A are the analogues, for the ordinary duality model, of Proposition 4 and Corollary 4A. For completeness, we provide the analogue of Proposition 3. This extends [**19**, Theorem 28.1] to the general case.

PROPOSITION 11. *Let η^I be a Kuhn-Tucker vector for* (P_0). *Then the optimal solutions to* (P_0), *if any, occur among the global minimizers of the function*

$$x_0 \to H_0\left(x_0, \eta^I\right) = \begin{cases} f_0(x_0) + \sum_I \eta_i f_i(x_0) & \text{if } x_0 \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

In particular, x_0 solves (P_0) *if and only if*

$$x_0 \quad \text{solves} \quad \inf_C \left\{ f_0 + \sum_I \eta_i f_i \right\} \tag{5.19}$$

and also satisfies

$$f_i(x_0) \leqslant 0 \leqslant \eta_i \quad \text{and} \quad f_i(x_0) \cdot \eta_i = 0, \quad \forall i \in I. \tag{5.20}$$

PROOF. We can obtain this result very succinctly as a corollary to Proposition 10, as follows. Since η^I is a Kuhn-Tucker vector, the proof of Proposition 10 shows that x_0 solves (P_0) if and only if (x_0, η^I) satisfies (5.18a) and (5.18b*). But these conditions are the same as (5.20) and (5.19), respectively. An elementary, direct proof can also be given. This we leave to the reader.

Next, we discuss constraint qualifications. This is the term usually given to any of a variety of conditions which guarantee the existence of a Kuhn-Tucker vector, or at least guarantee the strong duality relation (4.12) for the model under study. Propositions 8 and 7 imply that any result along these lines for the symmetrized model (involving either (D_s^+) or (D_s)) yields the corresponding result for the ordinary model. Sharp conditions under which the converse implications hold are provided also in Propositions 8 and 7. Notice in particular that when conditions (5.6) and (5.8) both hold, it makes no difference for which of the (three) models one establishes such results. That is, under these conditions it is immaterial whether one derives the result for (P_0) and (D_0), for (P_s) and (D_s^+), or for (P_s) and (D_s).

This raises the question of whether one can generate weaker (hence better) constraint qualifications by working with the symmetrized model. In general the answer is no, and the reason is as follows. Constraint qualifications are intimately tied up with the optimal value function's being bounded above on some neighborhood of the origin in U, the space of perturbations (see [**21**, §7]). Now the U involved in the symmetrized model is $X_0 \times X^I \times R^I$, since the model includes horizontal translations of the functions f_0, f_i ($i \in I$) as

well as vertical translations of the f_i's. On the other hand, U for the ordinary model is simply R^I, since only the vertical translations are involved. So if X_0 is infinite-dimensional, it is generally more difficult to ensure the boundedness property for φ_s, the optimal value function for (P_s), than it is for φ_0, the optimal value function for (P_0).

To illustrate, consider the form taken for (P_s) by the all-purpose constraint qualification presented in Proposition 5. It is that there should exist an x_0 satisfying both (1) f_0 is bounded above on some neighborhood of x_0, and (2) for each $i \in I$ there exists an $\alpha_i > 0$ such that $f_i \leqslant -\alpha_i$ on some neighborhood of x_0. This qualification is essentially based on applying [21, Theorem 18(a)] to F_s, as an inspection of the proof of Proposition 5 quickly reveals. It guarantees an optimal (y_0, y^I, η^I) for (D_s^+), where the component η^I is optimal for (D_0) by Proposition 7 (see the estimate (5.10)). By contrast, applying the same tool [21, Theorem 18(a)] to F_0 yields a considerably weaker (hence better) constraint qualification, namely, that there should exist an x_0 in dom f_0 such that $f_i(x_0) < 0$ for every $i \in I$ (the so-called Slater condition). The latter condition, though weaker, does however have the countervailing aspect of guaranteeing only an optimal η^I for (D_0). Lacking further assumptions, it is not enough to ensure in general the existence of an "optimal" pair y_0, y^I which, together with η^I, will solve (D_s^+); condition (5.8) or (5.8$^+$) would typically be required for that (see Proposition 7). In situations where not both (5.6) and (5.8) hold, there is then a certain tradeoff: constraint qualifications for the ordinary model are generally weaker, but the conclusions implied are not as strong as for the symmetrized model.

In the case $X_0 = R^n$, it turns out that the general difficulties mentioned above do not apply. This is due in part to the availability of the considerable arsenal of special facts concerning relative interiors. Combined with the additive separability structure which is the distinguishing feature of the symmetrized model, this permits the following particularly refined existence result. It extends slightly Rockafellar's theorem [19, Theorem 28.2], which already refined the Slater condition to handle affine functions.

PROPOSITION 12. *Assume* $X_0 = R^n$ *and* val(P_0) $> -\infty$. *In order that there exist a Kuhn-Tucker vector for* (P_s) *(and hence a fortiori a Kuhn-Tucker vector for* (P_0)*) it is sufficient that there exists an* x_0 *satisfying the conditions*

$$x_0 \in \text{ri dom } f_0 \tag{5.21}$$

and

$$x_0 \in \text{ri dom } f_i, \quad f_i(x_0) < 0 \tag{5.22_i}$$

for each $i \in I$. *Moreover, when any or all of the functions* f_0, f_i ($i \in I$) *are polyhedral, the corresponding conditions* (5.21), (5.22$_i$) *can be weakened as*

follows:

$$x_0 \in \text{dom } f_0, \tag{5.21'}$$

$$f_i(x_0) \leqslant 0. \tag{5.22$_i'$}$$

PROOF. Proposition 6 specialized according to (5.1) yields a Kuhn-Tucker vector (y_0, y^I, η^I) for (P_s). By the implication $(b^+) \Rightarrow (a)$ of Proposition 8, the component η^I is a Kuhn-Tucker vector for (P_0).

We conclude this section with a brief discussion of the case in which all of the constraint functions are (translates of) homogeneous functions. That is, we assume now that each f_i has the form

$$f_i(x_0) = h_i(x_0 - a_i) - \alpha_i, \tag{5.23}$$

where h_i is positively homogeneous of degree one and is assumed to be closed. Now it is easy to check that a closed proper convex function is homogeneous if and only if its conjugate is an indicator (i.e. assumes only the values 0 or $+\infty$). Hence, $h_i^* = \psi_{D_i}$ for some nonempty closed convex set D_i, and in fact one has

$$D_i = \{ y_0 \in Y_0 | \langle x_0, y_0 \rangle \leqslant h_i(x_0), \forall x_0 \in X_0 \} = \partial h_i(0)$$

(cf. [19, Corollary 13.2.1]). From this it is easy to compute that

$$(f_i^* \eta_i)(y_0) = \psi_{\eta_i D_i}(y_0) + \langle a_i, y_0 \rangle + \eta_i \alpha_i \tag{5.24}$$

for any $\eta_i > 0$. Here, for notational convenience we use the value $\eta_i = 0$ to represent the case of 0^+, i.e.

$$(f_i^* 0^+)(y_0) = \psi_{0^+ D_i}(y_0) + \langle a_i, y_0 \rangle.$$

Due to formulas (5.23) and (5.24), the various problems treated earlier in this section assume the following form:

$$\min\{ f_0(x_0) | h_i(x_0 - a_i) \leqslant \alpha_i, \forall i \in I \}; \tag{P_0}$$

$$\max_{\eta^I > 0} \left\{ \inf_{x_0 \in C} \left\{ f_0(x_0) + \sum_I \eta_i [h_i(x_0 - a_i) - \alpha_i] \right\} \right\}; \tag{D_0}$$

$$- \min \left\{ f_0^*(y_0) + \sum_I [\langle a_i, y_0 \rangle + \eta_i \alpha_i] | \eta^I \geqslant 0, \right.$$

$$\left. y_0 + \sum_I y_i = 0, y_i \in \eta_i D_i \ \forall i \in I \right\},$$

$$\text{where } \eta_i D_i \text{ is } 0^+ D_i \text{ whenever } \eta_i = 0; \tag{D_s^+}$$

$$\left. \begin{array}{l} \text{same as } (D_s^+), \text{ except for } \eta_i D_i \text{ being} \\ \text{interpreted as } 0 D_i = \{0\} \text{ whenever } \eta_i = 0. \end{array} \right\} \tag{D_s}$$

In a recent paper [2], C. R. Glassey essentially argued that solving (D_0) is equivalent to solving the projection of (D_s) onto R^I, i.e. solving for those

$\eta^I > 0$ satisfying

$$\text{val}(D_s) = -\inf\Big\{ f_0^*(y_0) + \sum_I [\langle a_i, y_0 \rangle + \eta_i \alpha_i] |$$

$$y_0 + \sum_I y_i = 0, y_i \in \eta_i D_i, \forall i \in I \Big\}.$$

The proof given, treating under certain assumptions the case in which $X_0 = R^n$, the a_i's are all zero, and f_0 is linear, has gaps however (see [2, p. 181]), and is conclusive only for functions h_i which are everywhere finite. (E.g. the formula derived just above equation (5a) requires $C = X_0$ in our terminology.) The difficulties which can arise when $C \neq X_0$, or more generally when (5.6) fails, are illustrated by the following examples.

EXAMPLE 1. Suppose $X_0 = R^2$ and $I = \{1, 2\}$, and let f_0, f_1, f_2 be as follows:

$$f_0(x_0) = \xi_2 \quad (\forall \xi_1 \in R),$$

$$f_1(x_0) = \begin{cases} -2(\xi_1\xi_2)^{1/2} & \text{if } x_0 \geq 0, \\ +\infty & \text{otherwise,} \end{cases}$$

$$f_2(x_0) = \begin{cases} 0 & \text{if } \xi_1 \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

This yields for the primal problem $\text{val}(P_0) = 0$, achieved on $\{x_0 | \xi_1 \geq 0 = \xi_2\}$, and for the ordinary dual (note that $C = R_+^2$ here), $\text{val}(D_0) = 0$, achieved on $\{\eta^I | \eta_1 = 0 \leq \eta_2\}$. Now it is not hard to verify that the conjugates of f_0, f_1, f_2 are the indicator functions

$$f_0^*(y_0) = \begin{cases} 0 & \text{if } y_0 = (0, 1), \\ +\infty & \text{otherwise,} \end{cases}$$

$$f_1^*(y_0) = \begin{cases} 0 & \text{if } \eta_1 < 0 \text{ and } \eta_1\eta_2 \geq 1, \\ +\infty & \text{otherwise,} \end{cases}$$

$$f_2^*(y_0) = \begin{cases} 0 & \text{if } \eta_1 \leq 0 \text{ and } \eta_2 = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

From this it follows that for the symmetrized dual, $\text{val}(D_s^+) = 0$, achieved on $\{(y_0, y^I, \eta^I) | \eta_1 = 0 \leq \eta_2, \ y_0 = (0, 1) = -y_1, y_2 = (0, 0)\}$, while for the simplified symmetrized dual, $\text{val}(D_s) = -\infty$ (infeasible). The latter means, of course, that the projection of (D_s) onto the space $R^I = R^2$ of multipliers η^I also has value $-\infty$ and is infeasible.

This example shows that (D_s), and its projection, can be hopelessly inadequate, with in fact an infinite gap between $\text{val}(D_s)$ and $\text{val}(D_s^+)$, even when both (5.8) and (5.8^+) are satisfied and there exists a Slater point. Condition (5.6) fails here. The next example is a slight variation, involving a

nonlinear f_0, illustrating a finite gap between val(D_s) and val(D_s^+), where both values are achieved.

EXAMPLE 2. Let everything be the same as in Example 1 except for replacing the f_0 there by $f_0(x_0) = e^{\xi_2}$ ($\forall \xi_1 \in R$), so that

$$f_0^*(y_0) = \begin{cases} \eta_2 \ln \eta_2 - \eta_2 & \text{if } \eta_1 = 0 \text{ and } \eta_2 > 0, \\ 0 & \text{if } \eta_1 = 0 = \eta_2, \\ +\infty & \text{otherwise.} \end{cases}$$

Then one can check that val(P_0) = 1, achieved on $\{x_0 | \xi_1 \geqslant 0 = \xi_2\}$, val($D_0$) = 1, achieved on $\{\eta' | \eta_1 = 0 \leqslant \eta_2\}$, and val($D_s^+$) = 1, achieved on $\{(y_0, y^1, \eta') | \eta_1 = 0 \leqslant \eta_2, y_0 = (0, 1) = -y_1, y_2 \in D_2\}$, whereas val($D_s$) = 0, achieved on $\{(y_0, y^1, \eta') | \eta_1 = 0 \leqslant \eta_2, y_0 = (0, 0) = y_1 = y_2\}$.

According to Proposition 7, if condition (5.6) holds then solving (D_s^+) is essentially equivalent to solving the simpler (D_s), and if in addition (5.8) holds then solving (D_s), or its projection, is essentially equivalent to solving (D_0). The pair of conditions (5.6), (5.8) is weaker than the conditions imposed in [2], and furthermore does not require linearity or even homogeneity of f_0. When (5.6) cannot be conveniently arranged, though (as for example when f_0 is linear and the h_i's are not everywhere finite), then according to Proposition 7(b$^+$) we could still use (D_s^+), or its projection onto R^I, as a satisfactory dual provided (5.8$^+$) is satisfied.

6. The ordinary model as the projection of the symmetrized model. In the last section we saw that the problems (P_0), (D_0), (L_0) in the ordinary duality model can each be regarded as essentially a projection of the corresponding problem (P_s), (D_s) (or (D_s^+)), (L_s) in the symmetrized duality model. We shall now show that in fact the entire ordinary problem trio collectively is the projection of the entire symmetrized problem trio. This we do by showing that the three projection transformations underlying the phenomena in §5 are interrelated in a certain well-prescribed way, much as are the three functions F, G, H which characterize the three problems in a perturbational duality model.

The projection transformations involved are linear transformations which are not everywhere defined. We find it appropriate to view them as oriented convex processes (see Rockafellar [19, §39] for definitions). Specifically, we shall exhibit three oriented convex processes, call them M, N, L, which are interrelated by the adjoint operation of [19] and which satisfy the relations

$$MF = F_0, \quad NG = G_0, \quad LH = H_0. \tag{6.1}$$

The left-hand sides of these relations represent certain operations of forming images of (convex, concave, or convex-concave) functions under (variously oriented) convex processes; these operations will be explained as we go. Here and throughout, the functions F_0, G_0, H_0 are those of (5.3), (5.5), (5.4) while

the functions F, G, H are those of (4.1), (4.3), (4.5) specialized according to (5.1). Each of the three relations in (6.1) expresses one of the individual projection phenomena of §5. It is the further fact that M, N, L are interrelated by means of the adjoint operation which corresponds to our assertion that the whole symmetrized trio projects onto the whole ordinary trio. Viewed another way, the results to follow establish under mild conditions the commutativity of certain constructions involving taking conjugates of functions, adjoints of processes, and images of functions under processes.

Let $M: X \times U \to X_0 \times R^I$ be the sup-oriented convex process given by

$$M(x, u) = \begin{cases} \{(x_0, \mu^I)\} & \text{if } u_0 = 0 \text{ and } u^I = 0, \\ \varnothing & \text{otherwise.} \end{cases} \tag{6.2}$$

Since F is convex and M is sup-oriented, the function $MF: X_0 \times R^I \to [-\infty, +\infty]$, called the image of F under M, is defined [19] by

$$(MF)(x_0, \mu^I) = \inf\{F(x, u) | (x, u) \in M^{-1}(x_0, \mu^I)\}. \tag{6.3}$$

It is not hard to show that this construction results in a convex function. What we wish to note is the following.

PROPOSITION 13. *The identity* $MF = F_0$ *holds without any additional assumption.*

PROOF. We can simply compute that

$$(MF)(x_0, \mu^I) = \inf_{x^I, \xi^I} \left\{ f(x_0, x^I, \xi^I) + \psi_K(x_0, x^I, \xi^I + \mu^I) \right\}$$

$$= \inf\left\{ f_0(x_0) + \sum_I \psi_i(x_i, \xi_i) | x_i = x_0 \text{ and } \xi_i = -\mu_i, \forall i \in I \right\}$$

$$= f_0(x_0) + \sum_I \psi_i(x_0, -\mu_i) = F_0(x_0, \mu^I).$$

Here we have used (6.3), (6.2), (4.1), (5.1), (2.1) and (5.3).

It is not hard to calculate from the definitions in [19] that the inverse of the adjoint of M is the sup-oriented convex process $M^{*-1}: V \times Y \to V_0 \times R^I$ given by

$$M^{*-1}(v, y) = \begin{cases} \{(v_0, \eta^I)\} & \text{if } v^I = 0 \text{ and } v^I = 0, \\ \varnothing & \text{otherwise.} \end{cases}$$

Now define an inf-oriented convex process $N: Y \times V \to R^I \times V_0$ by switching the order of variables in M^{*-1} and reversing its orientation:

$$N(y, v) = \begin{cases} \{(\eta^I, v_0)\} & \text{if } v^I = 0 \text{ and } v^I = 0, \\ \varnothing & \text{otherwise.} \end{cases} \tag{6.4}$$

Since G is concave and N is inf-oriented, the function $NG: R^I \times V_0 \to$

$[-\infty, +\infty]$, called the image of G under N, is defined [19] by

$$(NG)(\eta', v_0) = \sup\{G(y, v) | (y, v) \in N^{-1}(\eta', v_0)\}. \tag{6.5}$$

Again, it is not hard to show that this construction results in a concave function. What is of interest here is the following.

PROPOSITION 14. *If* (5.8^+) *holds, then* $NG = G_0$. *Moreover, this identity is equivalent to the identity* $(MF)^* = M^{*-1}F^*$.

PROOF. Using (6.5), (6.4), (4.3) and (2.6) yields

$$(NG)(\eta', v_0) = \sup_{y_0, y'} \left\{ -\psi_{K^*}(y_0, y', \eta') - f^*(y_0 + v_0, y', \eta') \right\}$$

$$= \sup_{y_0 + \Sigma_I y_i = 0} \left\{ -f_0^*(y_0 + v_0) - \sum_I \psi_i^*(y_i, \eta_i) \right\}.$$

By (2.8) this equals $-\infty$ when $\eta' \geqslant 0$ fails, while if $\eta' \geqslant 0$ holds it equals

$$\sup_{y_0 + \Sigma_I y_i = 0} \left\{ -f_0^*(y_0 + v_0) - \sum_I (f_i^*\eta_i)(y_i) \right\},$$

where $f_i^*\eta_i$ is $f_i^*0^+$ when $\eta_i = 0$. (It can be shown, incidentally, that this quantity is at most $G_0(\eta', v_0)$ even without (5.8^+).) Assuming that (5.8^+) holds, we therefore have that, for $\eta' \geqslant 0$,

$$(NG)(\eta', v_0) = - \inf_{y_0 + \Sigma_I y_i = v_0} \left\{ f_0^*(y_0) + \sum_I (f_i^*\eta_i)(y_i) \right\}$$

$$= -\left(f_0 + \sum_I \eta_i f_i \right)^*(v_0)$$

$$= \inf_{x_0 \in C} \left\{ f_0(x_0) + \sum_I \eta_i f_i(x_0) - \langle x_0, v_0 \rangle \right\} = G_0(\eta', v_0),$$

where the function $\eta_i f_i$ is ψ_{C_i} when $\eta_i = 0$ and we have used (5.5). For the equivalence assertion, note first that

$$- G_0(-\eta', v_0) = F_0^*(v_0, \eta') = (MF)^*(v_0, \eta')$$

by [21, equation (4.17)] and Proposition 13. Using [21, equation (4.17)] again, we have

$$- (NG)(-\eta', v_0) = -\sup\{G(y, v) | (y, v) \in N^{-1}(-\eta', v_0)\}$$

$$= -\sup\{G(y_0, y', -\eta', v_0, 0, 0) | y_0, y'\}$$

$$= -\sup\{G(-y_0, -y', -\eta', v_0, 0, 0) | y_0, y'\}$$

$$= \inf\{F^*(v_0, 0, 0, y_0, y', \eta') | y_0, y'\}$$

$$= \inf\{F^*(v, y) | (v, y) \in M^*(v_0, \eta')\} = (M^{*-1}F^*)(v_0, \eta').$$

The equivalence of the identities is immediate from these facts.

Propositions 13 and 14 show that when (5.8^+) holds the following diagram commutes:

$$
\begin{array}{ccc}
F & \xrightarrow{\quad * \text{ in } (x,\,u)\quad} & G \\[2pt]
M\downarrow & & N\downarrow \\[2pt]
F_0 & \xrightarrow{\quad * \text{ in } (x_0,\,\mu')\quad} & G_0
\end{array}
$$

The $*$ here denotes conjugacy, modulo minus signs (see [21, equation (4.17)]). Note also that N is just the inverse adjoint of M, modulo orientation.

It remains to establish the third relation in (6.1). For this, consider the product transformation $L: X \times Y \to X_0 \times R^I$ given by

$$L = L_1 \times L_2, \tag{6.6}$$

where $L_1: X \to X_0$ is the sup-oriented convex process

$$L_1(x) = \{x_0\} \tag{6.7}$$

and $L_2: Y \to R^I$ is the inf-oriented convex process

$$L_2(y) = \{\eta^I\}. \tag{6.8}$$

We have used the singleton set notation here to emphasize that we are regarding these (ordinary projection) transformations as convex processes. This is conceptually helpful later, when we consider the (not everywhere defined) inverse adjoints and corresponding orientations.

What we wish to do now is to form the image of H under L, much as we did previously with M and F in (6.3) and with N and G in (6.5). There is now, however, the ambiguity of whether we should take

$$
J_1(x_0, \eta') = \inf_{x \in L_1^{-1}x_0} \; \sup_{y \in L_2^{-1}\eta'} H(x,y)
$$
$$
= \inf\left\{\sup\{H(x,y)\,|\,y \in L_2^{-1}\eta'\}\,|\,x \in L_1^{-1}x_0\right\}
$$

or

$$
J_2(x_0, \eta') = \sup_{y \in L_2^{-1}\eta'} \; \inf_{x \in L_1^{-1}x_0} H(x,y)
$$
$$
= \sup\left\{\inf\{H(x,y)\,|\,x \in L_1^{-1}x_0\}\,|\,y \in L_2^{-1}\eta'\right\}.
$$

Our final result shows that it usually does not matter in which order these extrema are taken, and that moreover the result is H_0.

PROPOSITION 15. *The identity $J_1 = H_0$ holds without any additional assumption. If (5.6) holds and each f_i for $i \in I$ is closed, then also $J_2 = H_0$ holds.*

PROOF. Using (6.7) and (6.8), we have

$$J_1(x_0, \eta') = \inf_{x', \xi'} \sup_{y_0, y'} \{ H(x_0, x', \xi', y_0, y', \eta') \}.$$

By (4.5) this equals $+\infty$ when $x_0 \notin C_0$, while if $x_0 \in C_0$ we can continue:

$$= \inf_{\substack{\xi_i < -f_i(x_i) \\ \text{each } i \in I}} \sup_{y_0, y'} \{ f_0(x_0) - \psi_{K^*}(y) - \langle x, y \rangle \}$$

$$= \inf_{\substack{\xi_i < -f_i(x_i) \\ \text{each } i \in I}} \sup_{y_0 + \Sigma_I y_i = 0} \left\{ f_0(x_0) - \left[\langle x_0, y_0 \rangle + \sum_I \langle x_i, y_i \rangle \right] - \sum_I \xi_i \eta_i \right\}.$$

For any choice of x_i's which fails the condition $x_i = x_0, \forall i \in I$, the term in brackets can be used to drive the supremum to $+\infty$. Hence we can continue:

$$= \inf \left\{ f_0(x_0) - \sum_I \xi_i \eta_i | \xi_i < -f_i(x_0), \forall i \in I \right\}.$$

Now this infimum equals $+\infty$ vacuously when $x_0 \notin C_i$ for some $i \in I$; when $x_0 \in C$, it equals $-\infty$ if any $\eta_i < 0$, and it equals $f_0(x_0) + \Sigma_I \eta_i f_i(x_0)$ if $\eta' > 0$. Summarizing all this, we have

$$J_1(x_0, \eta') = \begin{cases} f_0(x_0) + \sum_I \eta_i f_i(x_0) & \text{if } x_0 \in C \text{ and } \eta' > 0, \\ -\infty & \text{if } x_0 \in C \text{ and } \eta' \ngtr 0, \\ +\infty & \text{if } x_0 \notin C, \end{cases}$$

$$= H_0(x_0, \eta'),$$

where the last equality is by (5.4). On the other hand,

$$J_2(x_0, \eta') = \sup_{y_0, y'} \inf_{x', \xi'} \{ H(x_0, x', \xi', y_0, y', \eta') \}.$$

By (4.5) this equals $+\infty$ when $x_0 \notin C_0$, while if $x_0 \in C_0$ we can continue

$$= \sup_{y_0, y'} \inf_{\substack{\xi_i < -f_i(x_i) \\ \text{each } i \in I}} \{ f_0(x_0) - \psi_{K^*}(y) - \langle x, y \rangle \}$$

$$= \sup_{y_0 + \Sigma_I y_i = 0} \inf_{\substack{\xi_i < -f_i(x_i) \\ \text{each } i \in I}} \left\{ f_0(x_0) - \left[\langle x_0, y_0 \rangle + \sum_I \langle x_i, y_i \rangle \right] - \sum_I \xi_i \eta_i \right\}.$$

If any $\eta_i < 0$, then for all y_0, y' the term $-\Sigma_I \xi_i \eta_i$ can be used to drive the infimum to $-\infty$, so that the overall value is $-\infty$ in this case. If $\eta' > 0$, we

can continue:

$$= \sup_{y_0 + \Sigma_i y_i = 0} \ \inf_{x_i \in C_i} \left\{ f_0(x_0) - \langle x_0, y_0 \rangle - \sum_I \langle x_i, y_i \rangle + \sum_I \eta_i f_i(x_i) \right\}$$

$$= \sup_{y_0 + \Sigma_i y_i = 0} \left\{ f_0(x_0) - \langle x_0, y_0 \rangle - \sum_I (f_i^* \eta_i)(y_i) \right\}$$

$$= \sup_{y_i} \left\{ f_0(x_0) + \sum_I \langle x_0, y_i \rangle - \sum_I (f_i^* \eta_i)(y_i) \right\}$$

$$= f_0(x_0) + \sum_I (f_i^* \eta_i)^*(x_0) = f_0(x_0) + \sum_I (\eta_i \, \mathrm{cl} \, f_i)(x_0).$$

In this calculation, for $\eta_i = 0$ we interpret $f_i^* \eta_i$ as $f_i^* 0^+$ and $\eta_i \, \mathrm{cl} \, f_i$ as $\psi_{\mathrm{cl} \, C_i}$ (by Lemma 1 of §3). This information is summarized by

$$J_2(x_0, \eta^I) = \begin{cases} f_0(x_0) + \Sigma_I(\eta_i \, \mathrm{cl} \, f_i)(x_0) & \text{if } x_0 \in C_0 \text{ and } \eta^I \geqslant 0, \\ -\infty & \text{if } x_0 \in C_0 \text{ and } \eta^I \not\geqslant 0, \\ +\infty & \text{if } x_0 \notin C_0, \end{cases}$$

where $\eta_i \, \mathrm{cl} \, f_i$ is $\psi_{\mathrm{cl} \, C_i}$ when $\eta_i = 0$. If (5.6) holds, then $C_0 = C$ and so $\psi_{\mathrm{cl} \, C_i}$ is zero on C_0. If also each f_i is closed, it follows from (5.4) that J_2 coincides with H_0.

Proposition 15 justifies the relation $LH = H_0$, where it is immaterial in which order the extrema are taken, provided (5.6) holds and the f_i's are closed. This completes the remaining part of (6.1).

On the strength of Propositions 13, 14 and 15, we have that the following diagram commutes:

$$\begin{array}{ccccc} F & \overset{\text{* in } u}{\rightarrow} & H & \overset{\text{* in } x}{\rightarrow} & G \\ M\downarrow & & L\downarrow & & N\downarrow \\ F_0 & \overset{\text{* in } \mu^I}{\rightarrow} & H_0 & \overset{\text{* in } x_0}{\rightarrow} & G_0 \end{array}$$

The * here denotes partial conjugacy in the arguments indicated, modulo minus signs (see [21, equations (4.2) and (4.15)]). In fact, there is among M, L, N a counterpart to the partial conjugacy relations among F, H, G. To see this, note from (6.7) that the sup-oriented convex process L_1^{*-1} is given by

$$L_1^{*-1}(v) = \begin{cases} \{v_0\} & \text{if } v^I = 0 \text{ and } \nu^I = 0, \\ \varnothing & \text{otherwise,} \end{cases} \tag{6.9}$$

and from (6.8) that the inf-oriented convex process L_2^{*-1} is given by

$$L_2^{*-1}(u) = \begin{cases} \{\mu^I\} & \text{if } u_0 = 0 \text{ and } u^I = 0, \\ \varnothing & \text{otherwise.} \end{cases} \tag{6.10}$$

Now reversing the orientations of both L_1^{*-1} and L_2^{*-1}, we have from (6.2), (6.7) and (6.10) that

$$M = L_1 \times L_2^{*-1} \tag{6.11}$$

and from (6.4), (6.8) and (6.9) that

$$N = L_2 \times L_1^{*-1}. \tag{6.12}$$

It follows from (6.11), (6.6) and (6.12) that the inverse adjoint operation on a process is the analogue of taking the partial conjugate of a function, and orientation reversal of a process is the counterpart of placing a minus sign before a function.

Another illustration of the phenomenon of an entire duality model projecting onto another appears in [4, §6].

REFERENCES

1. N. Bourbaki, *Espaces vectoriels topologiques*, Hermann, Paris, 1953.

2. C. R. Glassey, *Explicit duality for convex homogeneous programs*, Math. Programming **10** (1976), 176–191.

3. J. L. Kelley, I. Namioka, et al., *Linear topological spaces*, Van Nostrand, Princeton, N.J., 1963.

4. L. McLinden, *Symmetric duality for structured convex programs*, Trans. Amer. Math. Soc. **245** (1978), 147–181.

5. _____, *Affine minorants minimizing the sum of convex functions*, J. Optimization Theory Appl. **24** (1978), 569–583.

6. J.-J. Moreau, *Fonctionnelles convexes*, Mimeographed lecture notes, Collège de France, Paris, 1967.

7. E. L. Peterson, *Generalization and symmetrization of duality in geometric programming*, Discussion paper (Northwestern University, October 1972).

8. _____, *Geometric programming and some of its extensions*, Optimization and design (Internat. Summer School on the Impact of Optimization Theory on Technological Design, Louvain, 1971), Prentice-Hall, Englewood Cliffs, N. J., 1973, pp. 228–289.

9. _____, *Geometric programming*, SIAM Rev. **18** (1976), 1–51.

10. _____, *Constrained duality via unconstrained duality in generalized geometric programming*, J. Optimization Theory Appl. **26** (1978), 43–50.

11. _____, *The duality between suboptimization and parameter deletion*, Math. Oper. Res. **2** (1977), 311–319.

12. _____, *Saddle points and duality in generalized geometric programming*, J. Optimization Theory Appl. **26** (1978),

13. _____, *Optimality conditions in generalized geometric programming*, J. Optimization Theory Appl. **26** (1978), no. 1.

14. _____, *Ordinary duality vis-a-vis geometric duality*, (submitted for publication).

15. _____, *Geometric duality via Rockafellar duality*, (submitted for publication).

16. R. T. Rockafellar, *Duality theorems for convex functions* , Bull. Amer. Math. Soc. **70** (1964), 189–192.

17. _____, *Level sets and continuity of conjugate convex functions*, Trans. Amer. Math. Soc. **123** (1966), 46–63.

18. _____, *Extension of Fenchel's duality theorem for convex functions*, Duke Math. J. **33** (1966), 81–89.

19. _____, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.

20. _____, *Some convex programs whose duals are linearly constrained*, Nonlinear programming (Proc. Sympos., Univ. of Wisconsin, Madison, Wisconsin, 1970), Academic Press, New York, 1970, pp. 293–322.

21. _____, *Conjugate duality and optimization*, CBMS Regional Conf. Ser. Appl. Math., no. 16, SIAM, Philadelphia, 1974.

22. J. Stoer and C. Witzgall, *Convexity and optimization in finite dimensions*. I, Springer-Verlag, Berlin, 1970.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, ILLINOIS 61801